Perturbed Path Following Interior Point Algorithms

J. Frederic Bonnans, Cécilia Pola, Raja Rébaï

10
10
8cmcsc8

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Perturbed path following interior point algorithms

J. Frédéric Bonnans, Cecilia Pola, Raja Rebaï

N 2745

Décembre 1995

PROGRAMME 5

Rapport de recherche

ISSN 0249-6399

Perturbed path following interior point algorithms*

J. Frédéric Bonnans**, Cecilia Pola***, Raja Rebaï****

Programme 5 — Traitement du signal, automatique et productique
Projet Programmation mathématique

Rapport de recherche n° 2745 — Décembre 1995 — 24 pages

Abstract: The path following algorithms of predictor corrector type have proved to be very effective for solving linear optimization problems. However, the assumption that the Newton direction (corresponding to a centering or affine step) is computed exactly is unrealistic. Indeed, for large scale problems, one may need to use iterative algorithms for computing the Newton step.

In this paper, we study algorithms in which the computed direction is the solution of the usual linear system with an error in the right-hand-side. We give precise and explicit estimates of the error under which the computational complexity is the same as for the standard case. We also give explicit estimates that guarantee an asymptotic linear convergence at an arbitrary rate. Finally, we present some encouraging numerical results.

Because our results are in the framework of monotone linear complementarity problems, our results apply to convex quadratic optimization as well.

Key-words: Linear programming, linear complementarity problems, perturbation, decomposition, parallel computation, large scale problems, interior point methods, polynomial complexity, predictor corrector algorithm, infeasible algorithms.

(Résumé : tsvp)

* Supported by Human Capital and Mobility Program (European Community) CHRX-CT94-0471
** INRIA, B.P. 105, 78153 Rocquencourt, France. Email: Frederic.Bonnans@inria.fr.
*** Dpto Mat., Est. y Comp., Univ. Cantabria, Spain. Email:pola@matsun3.unican.es
**** INRIA, B.P. 105, 78153 Rocquencourt, France. Email: Raja.Rebai@inria.fr.

Algorithmes de points intérieurs par suivi de chemin avec perturbations

Résumé : Les algorithmes de suivi de chemin de type prédicteur correcteur se sont avérés très efficaces pour résoudre des problèmes d'optimisation linéaire. Cependant, l'hypothèse du calcul exact de la direction de Newton (correspondant à un pas affine ou de centrage) est peu réaliste. En effet, dans le cas de problèmes de grande taille, il peut être nécessaire d'utiliser des algorithmes itératifs.

L'article présente une étude de ces algorithmes prenant en compte une erreur dans le membre de droite. On donne des estimations précises et explicites de l'erreur permettant de préserver la complexité algorithmique du cas non perturbé. On calcule aussi des estimations explicites permettant de garantir un taux de convergence linéaire donné. Finalement, nous présentons des résultats numériques encourageants.

Les résultats sont obtenus dans le cadre de problèmes de complémentarité linéaire monotone. Ils s'appliquent donc aussi à l'optimisation quadratique convexe.

Mots-clé : Programmation linéaire, complémentarité linéaire monotone, perturbation, décomposition, calcul parallèle, problèmes de grande taille, méthodes de points intérieurs, complexité polynômiale, algorithme prédicteur correcteur algorithmes non réalisables.

1. Introduction.

In the last decade, a new generation of polynomial algorithms, based on the idea of computing a sequence of interior points, brought a revolution in the field of linear optimization [3, 4]. In the past few years, some algorithms that follow the central path, using Newton directions on the equation of the central path, focused the attention of the community because they both reach the optimal complexity known until now, while converging quadratically [7, 6, 8]. Most implementations actually available of interior point methods are of this type.

For very large scale problems it may be useful to compute the Newton step by an iterative algorithm. It is often observed that iterative algorithms compute at a small cost a rough approximation of the solution, while it may be much more expensive to obtain a precise value of the solution. Therefore a question arises: will the good convergence properties of path following algorithms remain if the Newton direction is computed approximately? This question is meaningful in the general framework of linear complementarity problems, in which linear optimization problems can be embedded (this also allows us to embed quadratic programming). Our concern is when the linearization of the complementarity condition is solved approximately, that is to say we consider the usual linear equations corresponding to the complementarity condition with a perturbation in the right-hand-side; we assume that the equations of the Newton step corresponding to the linear equations of the complementarity problem (i.e., in the case of linear optimization, the primal and dual linear feasibility constraints) themselves are not perturbed, but they can have a non-null right-hand-side if the starting point is infeasible.

In that framework, we are able to give explicit estimates on the precision with which the Newton step is computed, in order to keep complexity at the same order as for the non-perturbed case, or an asymptotic linear convergence rate at an arbitrary rate. We give such estimates for two algorithms of predictor corrector type, in small and large neighborhoods, respectively.

The paper is structured as follows. Section 2 presents the main results for the perturbed predictor corrector algorithm in small and large neighborhoods. The proofs are given in Section 3. Finally some numerical results are reported in the last section.

Conventions Given a vector $x \in I\!R^n$ the relation $x > 0$ is equivalent to $x_i > 0,\ i = 1, 2, \cdots, n$, while $x \geq 0$ means $x_i \geq 0,\ i = 1, 2, \cdots, n$. We denote $I\!R^n_+ = \{x \in I\!R^n : x \geq 0\}$ and $I\!R^n_{++} = \{x \in I\!R^n : x > 0\}$. We write $||.||$ instead of $||.||_2$. Whenever we use other norms like $||.||_\infty$ we use the corresponding symbol.

Given a vector x, the corresponding upper case symbol denotes as usual the diagonal matrix X defined by the vector. The symbol 1 represents the vector of all ones, with dimension given by the context.

We denote component-wise operations on vectors by the usual notations for real numbers. Thus, given two vectors u, v of the same dimension, uv, u/v, etc. will denote the vectors with components $u_i v_i$, u_i/v_i, etc. We denote the null space and range space of a matrix M by $\mathcal{N}(M)$ and $\mathcal{R}(M)$ respectively.

The notation $x^k = O(\mu_k)$ means that there is a constant K (dependent on problem data) such that for every $k \in \mathcal{N}$, $\|x^k\| \leq K\mu_k$. Similarly, if $x^k > 0$, $x^k = \Omega(\mu_k)$ means that $(x^k)^{-1} = O(1/\mu_k)$. Finally, $x^k \approx \mu_k$ means that $x^k = O(\mu_k)$ and $x^k = \Omega(\mu_k)$.

We use the same notations for a point x in a set parameterized by μ, say $\mathcal{E}_\mu$. We say that $x = O(\mu)$ (resp. $x = \Omega(\mu)$, $x \approx \mu$) whenever there is a constant K such that $\|x\| \leq K\mu$ (resp. $x^{-1} = O(1/\mu)$, $x = O(\mu)$ and $x = \Omega(\mu)$) for all $x \in \mathcal{E}_\mu$, and all small enough μ. In particular, $x \approx 1$ in $\mathcal{E}_\mu$ means that there are constants $K_2 > K_1 > 0$, such that any $x \in \mathcal{E}_\mu$ satisfies $K_1 \leq x_i \leq K_2$, $i = 1, \cdots, n$.

Given two vector functions x and y, $x \approx y$ means that $x_i \approx y_i$ for $i = 1, \cdots, n$, for small enough μ.

2. Main results.

The monotone horizontal linear complementarity problem (LCP) is as follows: to find $(x, s) \in I\!R^n \times I\!R^n$ satisfying

$$\begin{array}{rcl} xs & = & 0, \\ Qx + Rs & = & h, \\ x, s & \geq & 0, \end{array} \qquad (LCP)$$

where $h \in I\!R^n$, and $Q, R \in I\!R^{n \times n}$ are such that for any $u, v \in I\!R^n$,

$$Qu + Rv = 0 \text{ implies } u^T v \geq 0. \qquad (1)$$

We denote the feasible set, and the set of solutions of (LCP) as

$$\mathcal{F} := \{(x, s) \in I\!R^n_+ \times I\!R^n_+;\ Qx + Rs = h\}, \qquad (2)$$

$$\mathcal{S} := \{(x, s) \in \mathcal{F}; xs = 0\}. \qquad (3)$$

Also, we denote the set of strictly complementary solutions by

$$\mathcal{S}^0 := \{(x, s) \in \mathcal{S}; x + s > 0\}. \qquad (4)$$

It is well known that if (LCP) represents a linear programming problem, then if $\mathcal{S}$ is nonempty so is $\mathcal{S}^0$. This is not true in general, since it is easy to construct a quadratic program with nonempty $\mathcal{S}$ and empty $\mathcal{S}^0$. The existence of a strictly complementary solution will be an essential assumption in points 3.2, 3.3.2 and 3.4.2.

Let $(x^0, s^0, \mu_0) \in I\!R^n_+ \times I\!R^n_+ \times I\!R_{++}$ be given. Set

$$g = (h - Qx^0 - Rs^0)/\mu_0.$$

Then (x^0, s^0, μ_0) is an element of the set

$$\mathcal{F}^g := \{(x, s, \mu) \in I\!R^n_+ \times I\!R^n_+ \times I\!R_{++};\ Qx + Rs = h - \mu g\}. \qquad (5)$$

If in addition $x^0s^0 = \mu_0 1$, then (x^0, s^0, μ_0) belongs to the *infeasible central path pinned on* g, defined as

$$\begin{array}{rcl} xs & = & \mu 1, \\ Qx + Rs & = & h - \mu g, \end{array} \tag{6}$$

We consider algorithms for solving (LCP) that follow approximately this infeasible central path, which whenever $g = 0$ (feasible case) coincide with the (usual) central path. We assume that the algorithms under consideration generate points (x, s, μ) belonging to a "small" neighborhood of the form

$$\mathcal{V}_\alpha^g := \{(x, s, \mu) \in \mathcal{F}^g;\ \|\frac{xs}{\mu} - 1\| \le \alpha,\ \mu \le \mu_0\},$$

where $\alpha > 0$ and $\mu_0 > 0$ are given constants, or to a "large" neighborhood

$$\mathcal{N}_\nu^g := \{(x, s, \mu) \in \mathcal{F}^g;\ \nu 1 \le \frac{xs}{\mu} \le \nu^{-1} 1,\ \mu \le \mu_0\}$$

where $0 < \nu < 1$ and $\mu_0 > 0$ are again given constants. It is easily seen that

$$\mathcal{V}_\alpha^g \subset \mathcal{N}_\nu^g \quad \text{for all } 0 < \nu \le 1 - \alpha, \tag{7}$$

$$\mathcal{N}_\nu^g \subset \mathcal{V}_\alpha^g \quad \text{for all } \alpha \ge \sqrt{n}\left(\frac{1}{\nu} - 1\right). \tag{8}$$

The algorithms studied in this paper start each iteration from data $(x, s, \mu) \in \mathcal{V}_\alpha^g$ (or $\mathcal{N}_\nu^g$) and then compute an approximate solution of (6) with μ replaced by $\gamma\mu$ where $\gamma \in [0, 1]$. The Newton direction (u, v) is solution of the following perturbed linear system:

$$\begin{array}{rcl} su + xv & = & -xs + \gamma\mu 1 + \mu\eta, \\ Qu + Rv & = & (1 - \gamma)\mu g, \end{array} \tag{9}$$

where η is a perturbation taking into account the error in the computation of the Newton step. (As the unperturbed right-hand side is of order μ, we may expect to deal with perturbations of order μ, i.e. η of order 1.) Note that the second equation is not perturbed. The new point, obtained by taking a steplength $\theta \in (0, 1]$ along this direction is

$$x^\sharp = x + \theta u,\ s^\sharp = s + \theta v,\ \mu_\sharp := (1 - \theta + \theta\gamma)\mu, \tag{10}$$

where $\mu_\sharp$ is such that $\mu_\sharp \le \mu \le \mu_0$, and the new point $(x^\sharp, s^\sharp, \mu_\sharp)$ belongs to $\mathcal{V}_\alpha^g$ (or to $\mathcal{N}_\nu^g$), provided the "centering parameter" $\gamma \in [0, 1]$ and the steplength $\theta \in (0, 1]$ are chosen such that the new point belongs to a certain neighborhood of the central path.

Let us note that under the monotonicity assumption (1) the system (9) has a unique solution. Note also that the system (9) can be written as

$$
\begin{array}{rcl} su + xv & = & (1-\gamma)(-xs + \mu\eta) + \gamma(-xs + \mu 1 + \mu\eta), \\ Qu + Rv & = & (1-\gamma)\mu g, \end{array} \tag{11}
$$

and therefore

$$
(u, v) \quad = \quad \gamma(u^c, v^c) + (1-\gamma)(u^a, v^a)
$$

where (u^c, v^c) is the solution of (9) with $\gamma = 1$, called centering step, and (u^a, v^a) is the solution of (9) with $\gamma = 0$, called affine-scaling step.

Now we are ready to state a generic perturbed predictor corrector algorithm (**GPC**), in which the set $\mathcal{G}$ denotes $\mathcal{V}^g_\alpha$ or $\mathcal{N}^g_\nu$ depending on the algorithm:

Algorithm GPC Data: $\mu_\infty > 0$, $(x^0, s^0, \mu_0) \in \mathcal{G}$. $k := 0$

REPEAT

- $x := x^k$, $s := s^k$, $\mu := \mu_k$;
- Corrector step: Compute (u^c, v^c) solution of (9) with $\gamma = 1$.
  $x(\theta) := x + \theta u^c$, $s(\theta) := s + \theta v^c$.
  Compute $\theta^c \in]0, 1]$ such that $(x(\theta^c), s(\theta^c), \mu) \in \mathcal{G}$.
  $x := x(\theta^c)$, $s := s(\theta^c)$.
- Predictor step: Set $\gamma = 0$. Compute (u^a, v^a) solution of (9) with $\gamma = 0$.
  $x(\theta) := x + \theta u^a$, $s(\theta) := s + \theta v^a$, $\mu(\theta) := (1-\theta)\mu$.
  Compute θ^a, the largest value in $]0, 1[$ such that
  $(x(\theta), s(\theta), \mu(\theta)) \in \mathcal{G}$, $\forall\, \theta \leq \theta^a$.
  $x^{k+1} := x(\theta^a)$, $s^{k+1} := s(\theta^a)$, $\mu_{k+1} := (1-\theta^a)\mu_k$.
- $k := k + 1$.

UNTIL $\mu_k < \mu_\infty$.

In the following we consider two particular algorithms of the above general scheme and we state our main results in both cases.

The *perturbed predictor corrector algorithm in small neighborhoods* is defined as follows:

Algorithm SPC

Fix $\epsilon_c > 0$, $\epsilon_a > 0$, $\alpha \in]0, 1/2]$ and $\mathcal{G} = \mathcal{V}^g_\alpha$.
Specialize Algorithm **GPC** to this case with $\|\eta\| \leq \epsilon\alpha$, where $\epsilon \leq \epsilon_c$ during the corrector step and $\epsilon \leq \epsilon_a$ during the predictor step, and $\theta^c = 1$ at each iteration.

Note that in **GPC** we require that the point obtained after a restoration step is in the neighborhood $\mathcal{G}$, whereas we fix here $\theta^c = 1$. We check below that when ϵ_c is small enough, then the point obtained in a corrector step with $\theta^c = 1$ belongs to $\mathcal{V}^g_{\alpha/2}$. In addition, we give an explicit estimate of the size of errors for which the complexity of $O(\sqrt{n}L)$ for the feasible case (resp. $O(nL)$ for the infeasible case) is preserved. Assuming the strict complementarity hypothesis, we also compute a tighter bound under which a given asymptotic linear rate may be observed.

We say that (x^0, s^0) *dominates* (x^*, s^*) if $x^0 \geq x^*$ and $s^0 \geq s^*$. Similarly, we say that (x^0, s^0) *dominates* a solution of (LCP) if there exists $(x^*, s^*) \in \mathcal{S}$ which is dominated by (x^0, s^0).

THEOREM 2.1. *Let $L := \log_2(\mu_0/\mu_\infty)$. We assume that $\epsilon_c \leq 1/12$ ($\epsilon_c \leq 1/4$ whenever $\alpha \leq 1/4$), and that (x^0, s^0) dominates a solution of (LCP) whenever it is not feasible. Then Algorithm **SPC** has the following properties:*

(i) (Feasible case) Assume (x^0, s^0) to be feasible. If $\epsilon_a \leq 1/4$, then the algorithm finds a feasible point such that $\mu_k \leq \mu_\infty$ in at most $3\sqrt{n/\alpha}L$ iterations.

(ii) (Infeasible case) If $\epsilon_a = O(n)$, then the algorithm finds a point such that $\mu_k \leq \mu_\infty$ in at most $O(nL)$ iterations.

(iii) (Asymptotic rate of convergence) Assume $\mathcal{S}^0 \neq \emptyset$. Let $\beta \in (0,1)$ and ϵ_a be such that

$$\epsilon_a \left(1 + (1-\beta)\sqrt{4 + \frac{\epsilon_a^2}{2}}\right) < \frac{\beta}{2(1-\beta)}. \tag{12}$$

Then $\limsup \mu_{k+1}/\mu_k \leq \beta$. In particular, if $\epsilon_a \leq 1/5$, then $\mu_{k+1} \leq \mu_k/2$ for k large enough.

We now state the *perturbed predictor corrector algorithm in large neighborhoods.*

Algorithm LPC

Fix $\epsilon_c > 0$, $\epsilon_a > 0$, $\nu \in]0, 1/2]$ and $\mathcal{G} = \mathcal{N}_\nu^g$.
Specialize Algorithm **GPC** to this case with at each iteration $||\eta||_\infty \leq \epsilon_c$ when computing the centering direction and $||\eta||_\infty \leq \epsilon_a$ when computing the affine direction, and

$$\theta^c = \min\left\{1, \frac{\mu}{2||u^c v^c||_\infty}\left(\frac{1}{2} - ||\eta||_\infty\right)\right\}. \tag{13}$$

This choice of θ^c implies, as we will see, that the point obtained after a corrector step belongs to the interior of the large neighborhood, so that the algorithm is well defined. The theorem below also gives an explicit estimate of the size of errors for which the known complexity of the algorithm with $\eta = 0$, which is now $O(nL)$ for the feasible case and $O(n\sqrt{n}L)$ for the infeasible case, is preserved. Assuming the strict complementarity hypothesis, we also compute a tighter bound under which a given asymptotic linear rate may be observed.

THEOREM 2.2. *Let $L = \log_2(\mu_0/\mu_\infty)$. Assume that $\epsilon_c \leq 1/4$ and that (x^0, s^0) dominates a solution of (LCP) whenever it is not feasible. Then Algorithm **LPC** has the following properties:*

(i) (Feasible case) Assume (x^0, s^0) to be feasible. If $\epsilon_a \leq 1/4$, the algorithm finds a feasible point satisfying $\mu_k \leq \mu_\infty$ in at most $16nL/\nu^3$ iterations.

(ii) (Infeasible case) If $\epsilon_a = O(\sqrt{n})$, then the algorithm finds a point satisfying $\mu_k \leq \mu_\infty$ in $O(n\sqrt{n}L)$ iterations.

(iii) (Asymptotic rate of convergence) Assume $\mathcal{S}^0 \neq \emptyset$. *Let* $\beta \in (0,1)$ *and* ϵ_a *be such that*

$$\epsilon_a \leq \frac{\nu^4 \beta}{40\sqrt{2}(1-\beta)n}.$$

Then $\limsup \mu_{k+1}/\mu_k \leq \beta$.

3. Proof of main results.

3.1. Preliminary results. In this section we start by recalling some known technical results and deriving an easy consequence of them. These results will be used for the analysis of both algorithms. Set

$$\phi := \sqrt{\frac{xs}{\mu}};\ d := \sqrt{\frac{\mu x}{s}};\ \overline{u} := d^{-1}u;\ \overline{v} := \frac{dv}{\mu};\ \overline{Q} := QD;\ \overline{R} := \mu R D^{-1}.$$

Multiplying the first equation in (9) by $1/\sqrt{\mu x s}$, we obtain

$$\sqrt{\frac{s}{\mu x}}u + \frac{1}{\mu}\sqrt{\frac{\mu x}{s}}v = -\sqrt{\frac{xs}{\mu}} + \gamma\sqrt{\frac{\mu}{xs}} + \sqrt{\frac{\mu}{xs}}\eta.$$

Therefore $(\overline{u}, \overline{v})$ is solution of the scaled equation

$$\left\{ \begin{array}{lcl} \overline{u} + \overline{v} & = & -\phi + \gamma\phi^{-1} + \phi^{-1}\eta, \\ \overline{Q}\overline{u} + \overline{R}\overline{v} & = & (1-\gamma)\mu g. \end{array} \right. \tag{14}$$

From (9) and (10) it is easily seen that

$$\frac{x^\sharp s^\sharp}{\mu_\sharp} = \frac{1-\theta}{1-\theta+\theta\gamma}\left(\frac{xs}{\mu} - 1\right) + 1 + \frac{\theta}{1-\theta+\theta\gamma}\eta + \frac{\theta^2}{1-\theta+\theta\gamma}\frac{uv}{\mu}. \tag{15}$$

The first statement of the lemma below is due to Mizuno ([5], Lemma 1), and the second statement is due to Mizuno, Jarre and Stoer ([6], Corollary 1).

LEMMA 3.1.
(i) If $y, z \in I\!R^n$ *satisfies* $y^T z \geq 0$, *then* $\|yz\| \leq \frac{1}{\sqrt{8}}\|y+z\|^2$.
(ii) Let $(\hat{u}, \hat{v})$ *be solution of*

$$\begin{array}{rcl} \hat{u} + \hat{v} & = & \hat{f}, \\ Q\hat{u} + R\hat{v} & = & \hat{g}, \end{array}$$

and $\hat{x}, \hat{s}$ *such that* $Q\hat{x} + R\hat{s} = \hat{g}$. *Then*

$$\left\{ \begin{array}{l} \|\hat{u}\| \leq \|\hat{f}\| + \|\hat{x}\| + \|\hat{s}\|, \\[2ex] \|\hat{v}\| \leq \|\hat{f}\| + \|\hat{x}\| + \|\hat{s}\|. \end{array} \right.$$

The following technical estimates will be useful in the sequel.

LEMMA 3.2. *Let* (x^0, s^0, μ_0) *and* (x, s, μ) *be two elements of* $\mathcal{N}_\nu^g$ *such that* (x^0, s^0) *dominates a solution* (x^*, s^*) *of* (LCP). *Set* $(\tilde{x}, \tilde{s}) = (x^* - x^0, s^* - s^0)/\mu^0$. *Then*

$$\begin{aligned} x^T s^0 + s^T x^0 &\leq 4\mu_0 \frac{n}{\nu}, \\ Q\tilde{x} + R\tilde{s} &= g, \\ \|d^{-1}\tilde{x}\| + \|\frac{d\tilde{s}}{\mu}\| &\leq \frac{4n}{\nu\sqrt{\nu}\mu}. \end{aligned}$$

Proof. The first statement follows from Theorem 2.2 of [2], while the second is obvious. Let us prove the last one. We have

$$\|d^{-1}\tilde{x}\| = \|\frac{\phi^{-1}}{\mu} s\tilde{x}\| \leq \frac{\|\phi^{-1}\|_\infty}{\mu}\|s\tilde{x}\| \leq \frac{\|\phi^{-1}\|_\infty}{\mu}\|s\tilde{x}\|_1 = \frac{\|\phi^{-1}\|_\infty}{\mu} s^T|\tilde{x}|,$$

and similarly

$$\|\frac{d\tilde{s}}{\mu}\| = \|\frac{\phi^{-1}}{\mu} x\tilde{s}\| \leq \frac{\|\phi^{-1}\|_\infty}{\mu}\|x\tilde{s}\| \leq \frac{\|\phi^{-1}\|_\infty}{\mu}\|x\tilde{s}\|_1 = \frac{\|\phi^{-1}\|_\infty}{\mu} x^T|\tilde{s}|.$$

As (x^0, s^0) dominates (x^*, s^*), we obtain

$$\|d^{-1}\tilde{x}\| + \|\frac{d\tilde{s}}{\mu}\| \leq \frac{\|\phi^{-1}\|_\infty}{\mu}\left[s^T\frac{(x^0 - x^*)}{\mu_0} + x^T\frac{(s^0 - s^*)}{\mu_0}\right] \leq \frac{\|\phi^{-1}\|_\infty}{\mu\mu_0}(s^T x^0 + x^T s^0).$$

Combining with $\|\phi^{-1}\|_\infty \leq 1/\sqrt{\nu}$ and the first statement, the result follows. □

The next lemma gives upper bounds for the term $\|uv\|/\mu$ of (15). Part **i** will be used in the centering step and in the feasible affine-scaling step, while part **ii** will be useful in the infeasible affine-scaling step.

LEMMA 3.3. *Let* $(x, s, \mu) \in \mathcal{N}_\nu^g$ *and* (u, v) *be solution of (9). Then*

i. *If* $u^T v \geq 0$, *then*

$$\frac{\|uv\|}{\mu} \leq \frac{1}{\nu\sqrt{8}}\| - \phi^2 + \gamma 1 + \eta\|^2. \tag{16}$$

ii. *If* (x^0, s^0) *dominates a solution of* (LCP) *and* $\mu \leq \mu_0$, *then the affine direction satisfies*

$$\frac{\|u^a v^a\|}{\mu} \leq \left(\frac{\|\eta\|}{\nu} + \frac{5n}{\nu\sqrt{\nu}}\right)^2. \tag{17}$$

Proof. **i.** Using Lemma 3.1**i** and (14) we get

$$\frac{\|uv\|}{\mu} = \|\overline{uv}\| \leq \frac{1}{\sqrt{8}} \|\overline{u} + \overline{v}\|^2 = \frac{1}{\sqrt{8}} \| - \phi + \gamma\phi^{-1} + \phi^{-1}\eta\|^2,$$

from which the result follows.

ii. Applying Lemma 3.1**ii** and Lemma 3.2 to the scaled equation (14), and using $\|\phi^{-1}\|_\infty \leq 1/\sqrt{\nu}$ and $\|\phi\| \leq \sqrt{n/\nu}$, we have whenever $\gamma = 0$

$$\begin{aligned} \|\overline{u}\| &\leq \| - \phi + \phi^{-1}\eta\| + \mu\left(\|d^{-1}\tilde{x}\| + \|\frac{d\tilde{s}}{\mu}\|\right), \\ &\leq \frac{1}{\sqrt{\nu}}(\sqrt{n} + \|\eta\|) + 4\frac{n}{\nu\sqrt{\nu}} \leq \frac{\|\eta\|}{\sqrt{\nu}} + \frac{5n}{\nu\sqrt{\nu}}. \end{aligned}$$

The same estimate holds for $\overline{v}$. Using $\|\overline{uv}\| \leq \|\overline{u}\|\|\overline{v}\|$, the result follows. □

3.2. Asymptotic analysis. We now prove a general result that will be used for the study of the rapid asymptotic linear convergence in Section 3.3.2 and 3.4.2. Here, as in those sections, we assume (LCP) has a strictly complementary solution, i.e. $\mathcal{S}^0 \neq \emptyset$. It is well known that in this case there is a unique partition

$$B \cup N = \{1, 2, \cdots, n\}, \quad B \cap N = \emptyset,$$

such that for any $(x, s) \in \mathcal{S}^0$ we have $x_B > 0, s_B = 0, x_N = 0$ and $s_N > 0$.

Following Bonnans and Gonzaga [1] we rename the variables in the following way

$$x \leftarrow (x_B, s_N) \quad \text{and} \quad s \leftarrow (x_N, s_B).$$

Algorithm **GPC** is invariant with respect to the permutation, whose advantage is just to simplify the analysis by assuming that $N = \emptyset$. Therefore in points 3.3.2 and 3.4.2 we will always refer to x as the vector of large variables and to s as the vector of small variables. Of course this change of variables can only be done in the analysis, it cannot be used by algorithms since B and N are unknown.

The following lemma indicates what are the order of magnitudes in the vicinity of the infeasible central path.

LEMMA 3.4. *([2], Lemma 4.1) If $(x, s, \mu) \in \mathcal{G}$, then $x \approx 1$, $s \approx \mu$ and $d \approx 1$.*

Define the scaled variables

$$\overline{x} := d^{-1}x \ , \ \overline{s} := \frac{d}{\mu}s.$$

This scaling transfers x and s to the same vector $\overline{x} = d^{-1}x = \phi = ds/\mu = \overline{s}$. This scaling allows us to represent the solution of (9) as $O(\mu)$ perturbations of quantities that are easily analyzed. In order to do so we represent the solution of the scaled equations (14) in terms of orthogonal projections. Given $M \in IR^{m\times n}$, $q \in IR^m$ and the affine space defined by $Mx = q$, we define the projections operators $P_{M,q}$ and P_M by

$$x \to P_{M,q}x = argmin\{\|w - x\| : Mw + q = 0\},$$

and $P_M := P_{M,0}$. Similarly, we denote $\tilde{P}_M = I - P_M$ the orthogonal projection on $\mathcal{R}(M^T)$. It is easily checked that $P_{M,q}x = P_M x + P_{M,q}0$ for any $x \in IR^n$ and $q \in IR^m$.

LEMMA 3.5. *Let $(x, s, \mu) \in \mathcal{G}$.*

i. (Feasible case) If (x, s) is feasible, then the solution of the scaled system (14) satisfies

$$\begin{aligned} \overline{u} &= \gamma P_{\overline{Q}}\phi^{-1} + P_{\overline{Q}}(\phi^{-1}\eta) + O(\mu). \\ \overline{v} &= -\phi + \gamma\tilde{P}_{\overline{Q}}\phi^{-1} + \tilde{P}_{\overline{Q}}(\phi^{-1}\eta) + O(\mu). \end{aligned}$$

ii. (Infeasible case) Let $(\tilde{x}, \tilde{s})$ be such that $Q\tilde{x} + R\tilde{s} = g$. Then the solution of the scaled system (14) satisfies

$$\begin{aligned} \overline{u} &= \gamma P_{\overline{Q}}(\phi^{-1} + d\tilde{s}) + P_{\overline{Q}}(\phi^{-1}\eta) + O(\mu). \\ \overline{v} &= -\phi + \gamma\tilde{P}_{\overline{Q}}(\phi^{-1} + P_{\overline{Q}}(d\tilde{s})) + \tilde{P}_{\overline{Q}}(\phi^{-1}\eta) + O(\mu). \end{aligned}$$

Proof. Let us first check that

$$Q\hat{u} + R\hat{v} = 0 \Rightarrow \hat{v} \in \mathcal{R}(Q^T). \tag{18}$$

Indeed, if $(\hat{u}, \hat{v})$ satisfies the above relation, then $Q(\hat{u} + u') + R\hat{v} = 0$ for an arbitrary $u' \in \mathcal{N}(Q)$. Therefore $(\hat{u} + u')^T\hat{v} \geq 0$. As u' is an arbitrary element of the vector space $\mathcal{N}(Q)$, it follows that $(u')^T\hat{v} = 0$, i.e. $\hat{v} \in \mathcal{N}(Q)^\perp = \mathcal{R}(Q^T)$, as was to be proved.

i. In the feasible case we have $Qu + Rv = 0$. So, using (18), it follows that $\overline{v} = dv/\mu \in \mathcal{R}(\overline{Q}^T)$. Set

$$f := -\phi + \gamma\phi^{-1} + \phi^{-1}\eta.$$

Then

$$\begin{aligned} \overline{u} &\in f + \mathcal{R}(\overline{Q}^T), \\ \overline{Q}\overline{u} + \overline{R}\overline{v} &= 0. \end{aligned} \tag{19}$$

This is the optimality system characterizing the orthogonal projection of f over the set defined by the second relation. Therefore

$$\overline{u} = P_{\overline{Q},Rv}f = P_{\overline{Q}}f + P_{\overline{Q},Rv}0.$$

This expression may be simplified noticing that $P_{\overline{Q}}\phi = 0$. Indeed, let $(x^*, s^*) \in \mathcal{S}$, then $s^* = 0$ and $Q(x - x^*) + Rs = 0$; and therefore, using (18), $s \in \mathcal{R}(Q^T)$ and $\phi = ds/\mu \in \mathcal{R}(\overline{Q}^T)$. Hence, we deduce that

$$\overline{u} = P_{\overline{Q}}(\gamma\phi^{-1} + \phi^{-1}\eta) + P_{\overline{Q},Rv}0. \tag{20}$$

We obtain the desired expression for $\overline{u}$ by checking that

$$P_{\overline{Q},Rv}0 = O(\mu). \tag{21}$$

Indeed, using $\eta = O(1)$, $\|\overline{u}+\overline{v}\| = \|f\| = O(1)$ and, as $\overline{u}^T\overline{v} \geq 0$, we deduce that $\|\overline{v}\| \leq \|f\| = O(1)$, whence, using Lemma 3.4, $\|v\| = \mu\|d^{-1}\overline{v}\| = O(\mu)$. Now, let Q^- be a right inverse for Q. Then $\overline{Q}(D^{-1}Q^- Rv) = Rv$, whence

$$\|P_{\overline{Q},Rv}0\| \leq \|D^{-1}Q^- Rv\| = O(\|v\|) = O(\mu).$$

This proves the formula for $\overline{u}$, from which we deduce that

$$\overline{v} = f - \overline{u} = -\phi + \gamma\phi^{-1} + \phi^{-1}\eta - \gamma P_{\overline{Q}}\phi^{-1} - P_{\overline{Q}}(\phi^{-1}\eta) + O(\mu),$$

and so the last relation of part **i** follows.

ii. Now let us deal with the infeasible case. Set

$$\begin{aligned} \hat{u} &:= \overline{u} - (1-\gamma)\mu d^{-1}\tilde{x}, \\ \hat{v} &:= \overline{v} - (1-\gamma)\mu d\frac{\tilde{s}}{\mu}, \\ \hat{f} &:= -(1-\gamma)\mu(d^{-1}\tilde{x} + d\frac{\tilde{s}}{\mu}). \end{aligned}$$

Then, by using (14), we get

$$\left\{ \begin{array}{lcl} \hat{u} + \hat{v} & = & f + \hat{f}, \\ \overline{Q}\hat{u} + \overline{R}\hat{v} & = & 0. \end{array} \right. \tag{22}$$

Hence, as in the first part of the proof, we have

$$\begin{aligned} \hat{u} &= P_{\overline{Q}}(f+\hat{f}) + O(\mu), \\ &= P_{\overline{Q}}\left(-\phi - d\tilde{s} + \gamma(\phi^{-1} + d\tilde{s}) + \phi^{-1}\eta - \mu(1-\gamma)d^{-1}\tilde{x}\right) + O(\mu). \end{aligned}$$

Let us check that $P_{\overline{Q}}(\phi + d\tilde{s}) = 0$. Let $(x^*, s^*) \in \mathcal{S}$, $s^* = 0$, then $Q(x+\mu\tilde{x}-x^*)+R(s+\mu\tilde{s}) = 0$ and from (18) we deduce that $s+\mu\tilde{s} \in \mathcal{R}(Q^T)$. Therefore $\phi + d\tilde{s} = (d/\mu)(s+\mu\tilde{s}) \in \mathcal{R}(\overline{Q}^T) = \mathcal{N}(\overline{Q})^\perp$. Combining with the above display and Lemma 3.4 we deduce

$$\hat{u} = \gamma P_{\overline{Q}}(\phi^{-1} + d\tilde{s}) + P_{\overline{Q}}(\phi^{-1}\eta) + O(\mu).$$

Using $\overline{u} = \hat{u} + O(\mu)$, we obtain the formula for $\overline{u}$, from which we deduce the formula for $\overline{v} = f - \overline{u}$.
□

We will use the above lemma for the affine-scaling step in the rapid asymptotic linear convergence analysis.

3.3. The perturbed predictor corrector algorithm in small neighborhoods. Let us define the *centrality measure* as the mapping

$$\delta(x,s,\mu) = \|\frac{xs}{\mu} - 1\|.$$

If $(x,s,\mu) \in \mathcal{F}^g$ and $\delta(x,s,\mu) = 0$, then (x,s) is the *infeasible central point* associated with the parameter value μ.

3.3.1. Polynomial convergence.

Centering step. We start by describing the effect of a centering step on the centrality measure.

LEMMA 3.6. *Let (x,s,μ) be such that $\delta(x,s,\mu) \leq \alpha$. If $\epsilon_c > 0$ is so small that*

$$\epsilon_c + \frac{\alpha}{(1-\alpha)\sqrt{8}}(1+\epsilon_c)^2 \leq \frac{1}{2}, \tag{23}$$

then

$$\delta_c := \left\| \frac{(x+u^c)(s+v^c)}{\mu} - 1 \right\| \leq \frac{\alpha}{2}.$$

This holds in particular if $\epsilon_c \leq 1/12$ whenever $\alpha \leq 1/2$, and $\epsilon_c \leq 1/4$ whenever $\alpha \leq 1/4$.

Proof. From (15) with $\gamma = \theta = 1$ we get

$$\frac{(x+u^c)(s+v^c)}{\mu} - 1 = \eta + \frac{u^c v^c}{\mu}. \tag{24}$$

From (7) and (16), using $\delta(x,s,\mu) \leq \alpha$ and $\gamma = 1$, we have

$$\frac{\|u^c v^c\|}{\mu} \leq \frac{1}{(1-\alpha)\sqrt{8}}(\alpha + \|\eta\|)^2 \leq \frac{1}{(1-\alpha)\sqrt{8}}(1+\epsilon_c)^2\alpha^2. \tag{25}$$

Combining (25) and (24) and $\|\eta\| \leq \epsilon_c \alpha$ we deduce

$$\left\| \frac{(x+u^c)(s+v^c)}{\mu} - 1 \right\| \leq \epsilon_c\alpha + \frac{1}{(1-\alpha)\sqrt{8}}(1+\epsilon_c)^2\alpha^2.$$

From (23) we obtain the conclusion. □

Affine-scaling step. We now analyse the effect of an affine-scaling step on the centrality measure, beginning by considering an upper bound for this measure. From (15) with $\gamma = 0$, we get

$$\begin{aligned} \left\| \frac{(x+\theta u^a)(s+\theta v^a)}{\mu_\sharp} - 1 \right\| &= \left\| \frac{xs}{\mu} - 1 + \frac{\theta}{1-\theta}\eta + \frac{\theta^2}{1-\theta}\frac{u^a v^a}{\mu} \right\|, \\ &\leq \left\| \frac{xs}{\mu} - 1 \right\| + \frac{\theta}{1-\theta}\|\eta\| + \frac{\theta^2}{1-\theta}\frac{\|u^a v^a\|}{\mu}. \end{aligned}$$

By Lemma 3.6, the centrality measure after a centering step is at most $\alpha/2$. Hence we deduce that the point obtained with a value θ of the steplength along the direction (u^a, v^a) belongs to the small neighborhood whenever

$$\frac{\theta}{1-\theta}\|\eta\| + \frac{\theta^2}{1-\theta}\frac{\|u^a v^a\|}{\mu} \leq \frac{\alpha}{2}. \tag{26}$$

Therefore the main point is to estimate $\|u^a v^a\|$.

LEMMA 3.7. *Let* $(x, s, \mu) \in \mathcal{V}^g_{\alpha/2}$. *Then*

i. *(Feasible case) If (x,s) is feasible, we have*

$$\frac{\|u^a v^a\|}{\mu} \quad \leq \quad n\frac{(1+\alpha(1/2+\epsilon_a))^2}{(1-\alpha/2)\sqrt{8}}. \tag{27}$$

If in addition $0 < \epsilon_a \leq 1/2$ and $\hat\theta \leq 1/2$ verifies

$$2n\hat\theta^2\frac{(1+\alpha(1/2+\epsilon_a))^2}{(1-\alpha/2)\sqrt{8}} \leq \left(\frac{1}{2}-\epsilon_a\right)\alpha, \tag{28}$$

then $\delta(x+\theta u^a, s+\theta v^a, (1-\theta)\mu) \leq \alpha, \quad \forall\, \theta \in (0, \hat\theta]$. In particular, if $\epsilon_a \leq 1/4$ then $\theta^a \geq \frac{1}{3}\sqrt{\alpha/n}$.

ii. *(Infeasible case) Let (x^0, s^0) dominate a solution of (LCP). If $\epsilon_a = O(n)$, then $\theta^a = \Omega(1/n)$.*

Proof.

i. Remembering that $\delta(x,s,\mu) \leq \alpha/2$, and using $(u^a)^T v^a \geq 0$, we deduce from (7) and (16) with $\gamma = 0$, that

$$\frac{\|u^a v^a\|}{\mu} \leq \frac{1}{(1-\alpha/2)\sqrt{8}}\left\|\phi^2 - \eta\right\|^2.$$

Using $\left\|\phi^2\right\| \leq \sqrt{n}\left\|\phi^2\right\|_\infty \leq \sqrt{n}(1+\alpha/2)$, and $\|\eta\| \leq \epsilon_a\alpha$, we obtain

$$\frac{\|u^a v^a\|}{\mu} \leq \frac{[\sqrt{n}+\alpha(\epsilon_a+\sqrt{n}/2)]^2}{(1-\alpha/2)\sqrt{8}} \leq n\frac{(1+\alpha(1/2+\epsilon_a))^2}{(1-\alpha/2)\sqrt{8}}.$$

This proves (27).

Using $\delta(x,s,\mu) \leq \alpha/2$, (26) and $\|\eta\| \leq \epsilon_a\alpha$, we deduce that θ is feasible whenever

$$\frac{\theta}{1-\theta}\epsilon_a\alpha + \frac{\theta^2}{1-\theta}\frac{\|u^a v^a\|}{\mu} \leq \alpha/2.$$

As the function $f : (0,1) \longrightarrow IR$ defined by

$$f(\theta) = \frac{\theta}{1-\theta}\epsilon_a\alpha + \frac{\theta^2}{1-\theta}\frac{\|u^a v^a\|}{\mu}$$

is increasing, for obtaining the desired inequality it suffices to show that $f(\hat{\theta}) \leq \alpha/2$. Using $\hat{\theta} \leq \frac{1}{2}$ and $\frac{1}{1-\hat{\theta}} \leq 2$, we get

$$f(\hat{\theta}) \leq \epsilon_a \alpha + 2\hat{\theta}^2 \frac{\|u^a v^a\|}{\mu}.$$

Hence using part **i** and (28) we obtain that after a displacement step of value $\hat{\theta}$, the new point belongs to $\mathcal{V}_\alpha^g$.
For the proof of the last statement, observe that if $\theta^a \geq 1/2$, the result is obvious. Otherwise it suffices to prove that $\hat{\theta} := \frac{1}{3}\sqrt{\alpha/n}$ satisfies (28). In fact, as $\alpha \leq 1/2$ and $0 < \epsilon_a \leq 1/4$, we have

$$2n\hat{\theta}^2 \frac{(1+\alpha(1/2+\epsilon_a))^2}{(1-\alpha/2)\sqrt{8}} \leq n\hat{\theta}^2 \frac{(11/8)^2}{\frac{3}{4}\sqrt{2}} < \frac{1}{4}\alpha \leq \left(\frac{1}{2} - \epsilon_a\right)\alpha.$$

ii. Combining with (17) and (26), we deduce that θ is feasible whenever

$$\theta\|\eta\| + \theta^2 \left(\frac{\|\eta\|}{(1-\alpha)} + \frac{5n}{(1-\alpha)^{3/2}}\right)^2 \leq (1-\theta)\frac{\alpha}{2}.$$

If $\theta^a \geq 1/2$, the conclusion is obtained. Otherwise, the right-hand-side is greater than $\alpha/4$. Assuming $\|\eta\| \leq cn$, we see that θ is feasible whenever

$$(n\theta)c + \left(\frac{c}{(1-\alpha)} + \frac{5}{(1-\alpha)^{3/2}}\right)^2 (n\theta)^2 \leq \frac{\alpha}{4}.$$

The left-hand-side is null when $\theta = 0$. Therefore, the inequality holds whenever $n\theta$ is small enough. The result follows.

□

Using the above results we will can easily prove in Subsection 3.3.3 that Algorithm **SPC** has polynomial convergence.

3.3.2. Rapid asymptotic linear convergence.

LEMMA 3.8. *Let (x, s, μ) be such that $\delta(x, s, \mu) \leq \alpha$. Then*

$$\frac{\|u^a v^a\|}{\mu} \leq \epsilon_a \alpha \sqrt{4 + \frac{\epsilon_a^2}{2}} + O(\mu).$$

Proof. Define $\overline{u}^a := d^{-1}u^a, \overline{v}^a := d\frac{v^a}{\mu}$. From Lemma 3.5 with $\gamma = 0$ we get

$$\frac{u^a v^a}{\mu} = \overline{u}^a \overline{v}^a = (P_{\overline{Q}}\phi^{-1}\eta)(-\phi + \tilde{P}_{\overline{Q}}\phi^{-1}\eta) + O(\mu). \tag{29}$$

Using Lemma 3.1 **i** , $\|\phi^{-1}\|_\infty \leq \dfrac{1}{\sqrt{1-\alpha}}$, $\|\phi\|_\infty \leq \sqrt{1+\alpha}$ and $\|\eta\| \leq \epsilon_a\alpha$, and the fact that $P_{\overline{Q}}$ and $\tilde{P}_{\overline{Q}}$ are contractions, we have

$$\begin{aligned}
\|(P_{\overline{Q}}\phi^{-1}\eta)(-\phi+\tilde{P}_{\overline{Q}}\phi^{-1}\eta)\| &\leq \|(P_{\overline{Q}}\phi^{-1}\eta)(\tilde{P}_{\overline{Q}}\phi^{-1}\eta)\| + \|(P_{\overline{Q}}\phi^{-1}\eta)\phi\| \\
&\leq \frac{1}{\sqrt{8}}\|\phi^{-1}\eta\|^2 + \|\phi^{-1}\eta\|\|\phi\|_\infty \\
&= \|\phi^{-1}\eta\|\left(\frac{\|\phi^{-1}\eta\|}{\sqrt{8}} + \|\phi\|_\infty\right) \\
&\leq \|\phi^{-1}\|_\infty\|\eta\|\left(\frac{\|\phi^{-1}\|_\infty\|\eta\|}{\sqrt{8}} + \|\phi\|_\infty\right) \\
&\leq \frac{\epsilon_a\alpha}{\sqrt{1-\alpha}}\left(\frac{\epsilon_a\alpha}{\sqrt{8}\sqrt{1-\alpha}} + \sqrt{1+\alpha}\right) \\
&= \frac{\epsilon_a\alpha}{1-\alpha}\left(\frac{\epsilon_a\alpha}{\sqrt{8}} + \sqrt{1-\alpha^2}\right).
\end{aligned}$$

As the function $\alpha \to \dfrac{\epsilon_a\alpha}{\sqrt{8}} + \sqrt{1-\alpha^2}$ attains its maximum on $[0,1/2]$ at $\dfrac{\epsilon_a}{\sqrt{\epsilon_a^2+8}}$, we have, using $(1-\alpha)^{-1} \leq 2$:

$$\begin{aligned}
\|(P_{\overline{Q}}\phi^{-1}\eta)(-\phi+\tilde{P}_{\overline{Q}}\phi^{-1}\eta)\| &\leq 2\epsilon_a\alpha\left(\frac{\epsilon_a^2}{\sqrt{8}\sqrt{\epsilon_a^2+8}} + \sqrt{\frac{8}{\epsilon_a^2+8}}\right) \\
&= 2\epsilon_a\alpha\frac{\epsilon_a^2+8}{\sqrt{8}\sqrt{\epsilon_a^2+8}} = \epsilon_a\alpha\frac{\sqrt{\epsilon_a^2+8}}{\sqrt{2}} = \epsilon_a\alpha\sqrt{4+\frac{\epsilon_a^2}{2}}.
\end{aligned}$$

Combining the above inequality and (29) we obtain the conclusion. □

LEMMA 3.9. *Let $\beta \in (0,1)$ and (x,s,μ) such that $\delta(x,s,\mu) \leq \alpha/2$. If there exists a strictly complementary solution and (12) is satisfied, then $\theta^a \geq 1-\beta$ for k large enough.*

Proof. Using the same function f as in proof of Lemma 3.7, it suffices to show $f(1-\beta) \leq \alpha/2$ to obtain the conclusion. From Lemma 3.8, we get

$$\begin{aligned}
f(1-\beta) &= \frac{1-\beta}{\beta}\epsilon_a\alpha + \frac{(1-\beta)^2}{\beta}\frac{\|u^a v^a\|}{\mu}, \\
&\leq \frac{1-\beta}{\beta}\epsilon_a\alpha + \epsilon_a\alpha\frac{(1-\beta)^2}{\beta}\sqrt{4+\frac{\epsilon_a^2}{2}} + O(\mu), \\
&= \alpha\epsilon_a\frac{1-\beta}{\beta}\left(1+(1-\beta)\sqrt{4+\frac{\epsilon_a^2}{2}}\right) + O(\mu).
\end{aligned}$$

Hence, using (12) the result follows. □

In the next subsection we will use the previous results for proving the asymptotic rate of convergence of Algorithm LPC.

3.3.3. Proof of Theorem 2.1 .

i. By Lemma 3.6, we know that the centrality measure after a centering step is at most $\alpha/2$. In the feasible case, from Lemma 3.7 **i** the steplength of Algorithm **SPC** satisfies $\theta^a \geq \frac{1}{3}\sqrt{\alpha/n}$. Then $\mu_k \leq \left(1-\frac{1}{3}\sqrt{\frac{\alpha}{n}}\right)^k \mu_0$. Using $\left|log_2\left(1-\frac{1}{3}\sqrt{\alpha/n}\right)\right| \geq \frac{1}{3}\sqrt{\alpha/n}$ we obtain the conclusion.

ii. Similarly, in the infeasible case, from $\theta^a = \Omega(1/n)$, obtained in part **ii** of Lemma 3.7, we deduce that no more than $O(nL)$ iterations is necessary.

iii. This is a simple consequence of Lemmas 3.6 and 3.9.

3.4. The perturbed predictor corrector algorithm using large neighborhoods. In order to measure how interior is a point (x, s, μ) w.r.t. $\mathcal{N}_\nu^g$, we will use the distance of xs/μ to the boundary of the set

$$\mathcal{T}^\nu = \{z \in I\!R^n;\ \nu 1 \leq z \leq \nu^{-1}1\}.$$

3.4.1. Polynomial convergence.

Centering step. Let us denote

$$x^c = x + \theta^c u^c, \quad s^c = s + \theta^c v^c.$$

LEMMA 3.10. *Let $(x, s, \mu) \in \mathcal{N}_\nu^g$ and $\nu \in (0, 1/2]$. If $\epsilon_c \leq 1/4$, then $(x^c, s^c, \mu) \in \mathcal{N}_\nu^g$ and*

$$dist\left(\frac{x^c s^c}{\mu}, \partial\mathcal{T}^\nu\right) \geq \frac{\nu^3\sqrt{2}}{32n}. \tag{30}$$

Proof. From (15) with $\gamma = 1$

$$\begin{aligned} \frac{x^c s^c}{\mu} &= (1-\theta^c)\left(\frac{xs}{\mu}-1\right) + 1 + \theta^c\eta + (\theta^c)^2\frac{u^c v^c}{\mu} \\ &= (1-\theta^c)\frac{xs}{\mu} + \theta^c 1 + \theta^c\eta + (\theta^c)^2\frac{u^c v^c}{\mu}. \end{aligned} \tag{31}$$

Using $(x, s, \mu) \in \mathcal{N}_\nu^g$ and $\theta^c \leq 1$ we get

$$\begin{aligned} (\nu + \theta^c(1-\nu))1 = (1-\theta^c)\nu 1 + \theta^c 1 \leq (1-\theta^c)\frac{xs}{\mu} + \theta^c 1 &\leq \left((1-\theta^c)\frac{1}{\nu} + \theta^c\right)1, \\ &= \left(\frac{1}{\nu} + \theta^c\left(1-\frac{1}{\nu}\right)\right)1. \end{aligned}$$

As

$$\theta^c\left|1-\frac{1}{\nu}\right| = \theta^c\frac{|\nu-1|}{\nu} \geq \theta^c(1-\nu),$$

we have

$$dist((1-\theta^c)\frac{xs}{\mu}+\theta^c 1, \partial\mathcal{T}^\nu) \geq \theta^c(1-\nu) \geq \frac{\theta^c}{2}.$$

Using (31), the above inequality and $\theta^c > 0$, we get

$$\begin{aligned} dist(\frac{x^c s^c}{\mu}, \partial\mathcal{T}^\nu) &\geq dist\left((1-\theta^c)\frac{xs}{\mu}+\theta^c 1, \partial\mathcal{T}^\nu\right) - \theta^c\|\eta\|_\infty - (\theta^c)^2\frac{\|u^c v^c\|_\infty}{\mu}, \\ &\geq \theta^c(\frac{1}{2} - \|\eta\|_\infty) - (\theta^c)^2\frac{\|u^c v^c\|_\infty}{\mu}. \end{aligned}$$

Let us first consider the case when $\theta^c = 1$. Then, from (13), $\dfrac{\|u^c v^c\|_\infty}{\mu} \leq (\dfrac{1}{2} - \|\eta\|_\infty)/2$. Using $\|\eta\|_\infty \leq 1/4$, we deduce that $dist(\dfrac{x^c s^c}{\mu}, \partial\mathcal{T}^\nu) \geq (\dfrac{1}{2} - \|\eta\|_\infty)/2 \geq 1/8$, and the result follows. Otherwise

$$dist(\frac{x^c s^c}{\mu}, \partial\mathcal{T}^\nu) \geq (\frac{1}{2} - \|\eta\|_\infty)^2\frac{\mu}{4\|u^c v^c\|_\infty}.$$

Using (16), $(x, s, \mu) \in \mathcal{N}_\nu^g$ and $\|\eta\|_\infty \leq 1$, we obtain

$$\begin{aligned} \frac{\|u^c v^c\|_\infty}{\mu} &\leq \frac{\|u^c v^c\|}{\mu} \leq \frac{1}{\nu\sqrt{8}}\left\|\frac{xs}{\mu} - 1 + \eta\right\|^2 \leq \frac{n}{\nu\sqrt{8}}\left(\left\|\frac{xs}{\mu} - 1\right\|_\infty + \|\eta\|_\infty\right)^2, \\ &\leq \frac{n}{\nu\sqrt{8}}\left(\left(\frac{1}{\nu} - 1\right) + \|\eta\|_\infty\right)^2 \leq \frac{n}{\nu^3\sqrt{8}}. \end{aligned}$$

Combining with the previous inequality and using $\|\eta\|_\infty \leq 1/4$, we obtain the conclusion. □

Affine-scaling step. From (15) with $\gamma = 0$, denoting by $(x^\sharp, s^\sharp, \mu_\sharp)$ the point obtained after a step of value θ, we get

$$\frac{x^\sharp s^\sharp}{\mu_\sharp} = \frac{xs}{\mu} + \frac{\theta}{1-\theta}\eta + \frac{\theta^2}{1-\theta}\frac{u^a v^a}{\mu} \tag{32}$$

and therefore, by Lemma 3.10,

$$dist(\frac{x^\sharp s^\sharp}{\mu_\sharp}, \partial\mathcal{T}^\nu) \geq \frac{\nu^3\sqrt{2}}{32n} - \frac{\theta}{1-\theta}\|\eta\|_\infty - \frac{\theta^2}{1-\theta}\frac{\|u^a v^a\|_\infty}{\mu}. \tag{33}$$

The following technical lemma is used in the proof of Theorem 2.2.

Lemma 3.11. *Let* $(x, s, \mu) \in \mathcal{N}_\nu^g$.

i. *Assume that (x,s) is feasible. If $\epsilon_a \le 1/4$, then*

$$\frac{\|u^a v^a\|_\infty}{\mu} \quad \le \quad \frac{n}{\nu^3}. \tag{34}$$

Moreover if $dist(\frac{xs}{\mu}, \partial\mathcal{T}^\nu) \ge \frac{\nu^3\sqrt{2}}{32n}$ then $\overline{\theta} := \frac{\nu^3}{16n}$ satisfies

$$(x+\theta u^a, s+\theta v^a, (1-\theta)\mu) \in \mathcal{N}^g_\nu, \quad \forall\, \theta \in (0, \overline{\theta}].$$

ii. *If (x^0, s^0) dominate a solution of (LCP), $dist(\frac{xs}{\mu}, \partial\mathcal{T}^\nu) \ge \frac{\nu^3\sqrt{2}}{32n}$ and $\varepsilon_a = O(\sqrt{n})$, then $\theta^a = \Omega(1/(n\sqrt{n})$.*

Proof.

i. Whenever (x,s) is feasible, we have $(u^a)^T v^a \ge 0$. Hence from (16) with $\gamma = 0$, using $(x,s,\mu) \in \mathcal{N}^g_\nu$, $\varepsilon_a \le 1/4$ and $\nu \le 1$, we get

$$\begin{aligned} \frac{\|u^a v^a\|}{\mu} \quad &\le \quad \frac{1}{\nu\sqrt{8}} \left\| \frac{xs}{\mu} - \eta \right\|^2 \le \frac{n}{\nu\sqrt{8}} \left(\left\| \frac{xs}{\mu} \right\|_\infty + \|\eta\|_\infty \right)^2, \\ &\le \quad \frac{n}{\nu\sqrt{8}} \left(\frac{1}{\nu} + \frac{1}{4} \right)^2 \le \frac{n}{\nu\sqrt{8}} \frac{25}{16\nu^2} \le \frac{n}{\nu^3}. \end{aligned}$$

This proves (34).
It is easily checked that $\overline{\theta}$ satisfies

$$2\frac{n}{\nu^3}\overline{\theta}^2 + \frac{\overline{\theta}}{2} \le \frac{\nu^3\sqrt{2}}{32n}. \tag{35}$$

Using (33) and (34) we have whenever $\theta \le 1/2$ and $\|\eta\|_\infty \le 1/4$

$$dist(\frac{x^\sharp s^\sharp}{\mu_\sharp}, \partial\mathcal{T}^\nu) \quad \ge \quad \frac{\nu^3\sqrt{2}}{32n} - \frac{\theta}{2} - 2\theta^2\frac{n}{\nu^3}.$$

In view of (35), the right-hand-side is positive whenever $\theta \in (0, \overline{\theta}]$. The conclusion follows.

ii. If $\theta^a \ge 1/2$, the conclusion is obtained. Otherwise with (33) and (17), we get

$$dist(\frac{x^\sharp s^\sharp}{\mu_\sharp}, \partial\mathcal{T}^\nu) \quad \ge \quad \frac{\nu^3\sqrt{2}}{32n} - 2\theta\|\eta\|_\infty - 2\theta^2 \left(\frac{\sqrt{n}\|\eta\|_\infty}{\nu} + \frac{5n}{\nu\sqrt{\nu}} \right)^2.$$

Assuming $\|\eta\|_\infty \le c\sqrt{n}$, we see that θ is feasible whenever

$$\frac{\nu^3\sqrt{2}}{32} - 2\theta cn\sqrt{n} - 2\theta^2 n \left(\frac{cn}{\nu} + \frac{5n}{\nu\sqrt{\nu}} \right)^2 \ge 0.$$

Therefore θ is feasible whenever

$$\frac{\nu^3\sqrt{2}}{64} \geq c(\theta n\sqrt{n}) + (\theta n\sqrt{n})^2 \left(\frac{c}{\nu} + \frac{5}{\nu\sqrt{\nu}}\right)^2 .$$

The right-hand-side is null when $\theta = 0$. Therefore, the inequality holds whenever $\theta n\sqrt{n}$ is small enough. The result follows.

□

Using the above results we will prove in Subsection 3.4.3 that Algorithm **LPC** has polynomial convergence.

3.4.2. Asymptotic rate of convergence.

LEMMA 3.12. *Let $(x, s, \mu) \in \mathcal{N}_\nu^g$. If there exists a strictly complementary solution and $\epsilon_a \leq 1$ then*

$$\frac{\|u^a v^a\|_\infty}{\mu} \leq \frac{2}{\nu}\epsilon_a + O(\mu).$$

Proof. Using Lemma 3.5 with $\gamma = 0$, we get

$$\frac{u^a v^a}{\mu} = \overline{u}^a \overline{v}^a = (P_{\overline{Q}}(\phi^{-1}\eta))((-\phi + \tilde{P}_{\overline{Q}}(\phi^{-1}\eta))) + O(\mu). \tag{36}$$

Using $\|\phi^{-1}\|_\infty \leq 1/\sqrt{\nu}$, $\|\phi\|_\infty \leq 1/\sqrt{\nu}$ and the fact that a projection is a contraction, we have

$$\begin{aligned}
\|(P_{\overline{Q}}\phi^{-1}\eta)(-\phi + \tilde{P}_{\overline{Q}}\phi^{-1}\eta)\|_\infty &\leq \|P_{\overline{Q}}\phi^{-1}\eta\|_\infty \left(\|\phi\|_\infty + \|\tilde{P}_{\overline{Q}}\phi^{-1}\eta\|_\infty\right), \\
&\leq \|\phi^{-1}\|_\infty\|\eta\|_\infty \left(\|\phi\|_\infty + \|\phi^{-1}\|_\infty\|\eta\|_\infty\right), \\
&\leq \frac{\epsilon_a}{\sqrt{\nu}}\left(\frac{1}{\sqrt{\nu}} + \frac{\epsilon_a}{\sqrt{\nu}}\right) \leq \frac{2}{\nu}\epsilon_a.
\end{aligned}$$

Combining the above inequality and (36) we obtain the conclusion. □

LEMMA 3.13. *Let $\beta \in (0,1)$, $0 \leq \nu \leq 1/2$ and $(x, s, \mu) \in \mathcal{N}_\nu^g$ such that $dist(\frac{xs}{\mu}, \partial\mathcal{T}^\nu) \geq \frac{\nu^3\sqrt{2}}{32n}$. If there exists a strictly complementary solution and $\epsilon_a \leq \frac{\nu^4\beta}{40\sqrt{2}(1-\beta)n}$, then $\theta^a \geq 1 - \beta$ for k large enough.*

Proof. It suffices to show that the right-hand-side of (33) is positive whenever $\theta \leq 1 - \beta$, i.e.

$$\frac{1-\beta}{\beta}\epsilon_a + \frac{(1-\beta)^2}{\beta}\frac{\|u^a v^a\|_\infty}{\mu} \leq \frac{\nu^3\sqrt{2}}{32n}.$$

By Lemma 3.12, this will be satisfied if

$$\frac{1-\beta+\frac{2}{\nu}(1-\beta)^2}{\beta}\epsilon_a < \frac{\nu^3\sqrt{2}}{32n}.$$

i.e.

$$\epsilon_a \quad < \quad \frac{\beta}{1-\beta+\frac{2}{\nu}(1-\beta)^2}\frac{\nu^3\sqrt{2}}{32n} = \frac{1}{\nu+2(1-\beta)}\frac{\nu^4\beta}{(1-\beta)n}\frac{\sqrt{2}}{32}.$$

The conclusion follows. □

The results of this subsection will be useful for obtaining the asymptotic rate of convergence of Algorithm LPC.

3.4.3. Proof of Theorem 2.2 .

i. From Lemma 3.10, we get

$$dist(\frac{x^c s^c}{\mu}, \partial\mathcal{T}^\nu) \geq \frac{\nu^3\sqrt{2}}{32n};$$

hence from Lemma 3.11 **i** the steplength of Algorithm **LPC** satisfies $\theta^a \geq \nu^3/16n$. We obtain the conclusion following the same argument as in proof of part **i** of Theorem 2.1.

ii. Similarly, in the infeasible case, from $\theta^a = \Omega(\frac{1}{n\sqrt{n}})$, obtained in part **ii** of Lemma 3.11, we deduce that no more than $O(n\sqrt{n}L)$ iterations are necessary.

iii. This is an immediate consequence of Lemma 3.13.

4. Numerical experiments. In this section we present some numerical results that strongly support the theoretical estimates obtained in the preceding sections. These experiments are limited to the large neighborhood predictor corrector algorithm, choosing the size of the neighborhood $\nu = 0.01$. The algorithm is the one described in this paper, in which the stepsize for the centering displacement is as follows: starting from a unit step, we divide the step by two until the new point belongs to the large neighborhood. This is a rather rough linear search. Because it proved to be efficient, we content of it. We apply this algorithm to a family of linear programming problems in standard form, i.e.

$$\underset{x}{Min}\, c^T x \ ; \ Ax = b \ ; \ x \geq 0, \tag{37}$$

where $x \in I\!R^n$ and $b \in I\!R^p$. The necessary and sufficient optimality conditions for this problem are

$$\begin{array}{rcl} xs &=& 0, \\ Ax &=& b, \\ c + A^T\lambda &=& s, \\ x \geq 0, s &\geq& 0, \end{array}$$

and may be formally reduced to the linear complementarity format by writing the third condition in the equivalent form $B^T c = B^T s$, where B is a matrix whose colums form a basis of the kernel of A. The associated perturbed Newton step $(u, v, \delta\lambda)$ is solution of

$$\begin{array}{rcl} su + xv & = & -xs + \gamma\mu 1 + \mu\eta, \\ Au & = & (1-\gamma)\mu g_P, \\ A^T\delta\lambda - v & = & (1-\gamma)\mu g_D, \end{array}$$

where g_P and g_D depend on the starting point.

In order to test numerically the robustness of the Newton direction, we choose to generate some perturbations of the right-hand-side of the corresponding linear system in a random way. Then we compare the number of iterations to the one obtained without perturbation.

We generate the data of the problem to be solved as follows. Given an even value for n, we fix $p = n/2$. Then the matrix A is randomly generated (we performed all computations and generations of random numbers using the standard MATLAB functions). We randomly generate two nonnegative vectors $\hat{x}$ and $\hat{s}$. Then we take $b = A\hat{x}$ and $c = \hat{s}$. In that way, $\hat{x}$ and $(\hat{s}, \hat{\lambda} = 0)$ are primal and dual feasible, and therefore the linear problem has solutions. The starting point is $x^0 = s^0 = 1$ and $\mu_0 = 1$. The algorithm stops when $\mu < 10^{-10}$.

The perturbation η is chosen so that it satisfies at each iteration

$$\mu\|\eta\| = \epsilon\|f\|,$$

where $f = -xs + \gamma\mu 1$ is the right-hand-side of the equations of the Newton step corresponding to the linearization of the complementarity condition. That is, the perturbation parameter ϵ is the ratio (measured in the L^2 norm) between the perturbation and the right hand side. It varies between 0 and 0.25. Note that $\eta = \epsilon\mu^{-1}\|f\|z$, where z belongs to the unit sphere of $I\!R^n$. One difficulty is that we do not have a direct mean for generating an element of the unit sphere of $I\!R^n$ with uniform probability. Therefore we propose two ways for generating z.

The first method consists in generating at random each component of a vector $\hat{z}$. Then we obtain z by changing to 0 half of the components of $\hat{z}$, i.e. we have either $z_i = \hat{z}_i$ or $z_i = 0$, $i = 1, \ldots, n$.

Our results are a mean over ten runs except for $n = 10000$ where we perform only three runs. Using this first method for generating z, we obtain the results of tables 1 and 2.

n/ϵ	0	0.05	0.1	0.15	0.2	0.25
10	9.3	12.7	15.5	16.5	21.4	24.9
30	10.9	14.9	16.7	19	22	24.3
100	15	17.1	19.2	21.2	23.7	26.7
300	17.5	19.2	21.8	23.7	26.1	28.8
1000	17.9	19.9	21.8	24.7	27.6	30.4
3000	20	22.9	25	27.4	30.1	33.3
10000	27	29	29	32.67	36	37.67

Table 1: mean value of the number of iterations

n/ϵ	0	0.05	0.1	0.15	0.2	0.25
10	10	15	18	21	25	33
30	13	17	18	20	23	26
100	17	19	20	22	24	27
300	19	20	23	27	27	30
1000	21	22	23	29	31	32
3000	24	24	28	30	33	37
10000	29	31	30	34	44	39

Table 2: number of iterations in the worse case

We now consider a second method of generating the direction z where we choose to concentrate the perturbation in one component; i.e., all components of z are 0 except one, taken at random. We test this perturbation method only in the case $n = 10000$. We perform three runs for each value of ϵ. We obtain the following tables:

ϵ	0.05	0.1	0.15	0.2	0.25
	54.34	62.67	64.67	68	52.67

Table 3: mean value of the number of iterations

ϵ	0.05	0.1	0.15	0.2	0.25
	62	83	76	74	74

Table 4: number of iterations in the worse case

From these results, we may conclude that, at least for randomly generated problems, the large neighborhood predictor corrector algorithm described in this paper is both rapid and very robust with respect to perturbations in the computation of the Newton step.

REFERENCES

[1] J.F. Bonnans, C.C. Gonzaga : *Convergence of interior point algorithm for the monotone linear complementarity problem.* Mathematics of Operations Research, to appear.

[2] J.F. Bonnans, F.A. Potra : *Infeasible path following algorithms for linear complementarity problems.* Rapport de Recherche INRIA 2455, 1994. Mathematics of Operations Research, submitted.

[3] C.C. Gonzaga : *Path following methods for linear programming.* SIAM Review 34(1992), 167-227.

[4] M. Kojima, N. Megiddo, T. Noma, A. Yoshise : *A unified approach to interior point algorithms for linear complementarity problems*, Lecture Notes in Computer Science, 538, Springer Verlag, Berlin, (1991).

[5] S. Mizuno : *A new polynomial time method for a linear complementarity problem*, Mathematical Programming, 56 (1992), 31–43.

[6] S. Mizuno, F. Jarre, J. Stoer : *A unified approach to infeasible-interior-point algorithms via geometrical linear complementarity problems.* Preprint 213(1994), Math. Inst. Univ. Würzburg, 97074 Würzburg, Germany.

[7] S. Mizuno, M.J. Todd, Y. Ye : *On adaptive step primal-dual interior-point algorithms for linear programming.* Mathematics of Operations Research 18(1993), 964-981.

[8] F.A. Potra : *An $O(nL)$ infeasible interior point algorithm for LCP with quadratic convergence.* Reports on Computational Mathematics 50, Department of Mathematics, The University of Iowa, Iowa City, A52242, USA, 1994.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399