Numerical Comparisons of Path-Following Strategies for a Basic Interior-Point Method for Nonlinear Programming

M. Argaez, R.A. Tapia, and L. Velazquez

CRPC-TR97777-S
Revised August 1998

Center for Research on Parallel Computation
Rice University
6100 South Main Street
CRPC - MS 41
Houston, TX 77005

Numerical Comparisons of Path–Following Strategies for a Basic Interior–Point Method for Nonlinear Programming

M. Argáez,* R. A. Tapia† L. Velázquez‡

Abstract

An important research activity in interior–point methodology for general nonlinear program-
ming is to determine effective path–following strategies and its implementations using merit
function technology. The objective of this work is to present numerical comparisons of several
globalization strategies for the local interior–point Newton's method described by El-Bakry,
Tapia, Tsuchiya, and Zhang. We conduct a numerical experimentation between three notions
of proximity to the central path or a relaxation of it, with three merit functions to obtain an
optimal solution. We present which strategy works best, as well as some interesting comments
about the rest of the strategies.

Keywords: interior-point method, path-following strategy, merit function.

Abbreviated Title: Numerical Comparisons.

1   Introduction

In 1991, El-Bakry, Tapia, Tsuchiya, and Zhang [5] extended the primal-dual interior-point New-
ton formulation from linear programming to general nonlinear programming, and established local
and Q-quadratic convergence under standard assumptions for Newton's method. Moreover, they
presented a globalization strategy which consists in a linesearch procedure using the ℓ_2-norm resid-
ual function of the Karush-Kuhn-Tucker (KKT) conditions associated to the problem as a merit
function, and as centrality condition a standard measure used in feasible linear programming.

In 1995, Argáez and Tapia [2] proposed a path-following primal-dual interior-point method for
nonlinear programming (NLP) to globalize the strategy of El-Bakry at el [5]. This method intro-
duces a new centrality region, that is a relaxation of the central path, and a new merit function

which can be classified as a modified augmented Lagrangian function. The new centrality region, so called quasi central–path, removes the gradient with respect to the primal variables of the Lagrangian function of the perturbed KKT conditions. Therefore the Lagrange multipliers associated with the equality constraints are neglected as primary variables in their formulation. The new merit function modifies the augmented Lagrangian function associated with the equality constrained optimization problem by adding to its penalty term a potential reduction function utilized in linear programming to handle the perturbed complementarity condition. Also, they introduced a new centrality condition that is used as a measure of proximity to the quasi central–path. The method consists in applying a linesearch strategy that uses the new merit function to reach their centrality condition for a fixed perturbation parameter.

Due to the promising numerical results of Argáez and Tapia method, Parada and Tapia [11] defined in 1997 another modified augmented Lagrangian function and a new centrality condition as a measure of proximity to the quasi central–path. This new merit function differs from the Argáez–Tapia merit function in the second term of the penalty term by replacing the potential reduction function with the norm squared of the perturbed complementarity condition, and they also obtain encouraging numerical results.

The determination of an effective path–following strategy is an open research area in general nonlinear programming from both the theoretical and computational point of view. Therefore the objective of this work is to conduct a numerical comparison of different options that can be taken to reach an optimal solution between three notions of proximity to either the central path or quasi central-path, with the three merit functions that were used by El-Bakry et al [5], Argáez and Tapia[2] Parada and Tapia [11] in their work.

We present which centrality condition works best, and which merit function is the best tool to reach this centrality condition. We also present some important comments about the other options implemented.

2 Problem Formulation

We consider the general nonlinear program in the form

$$
\begin{aligned}
\text{minimize} \quad & f(x) \\
\text{subject to} \quad & h(x) = 0 \\
& x \geq 0,
\end{aligned} \tag{1}
$$

where $h(x) = (h_1(x), \ldots, h_m(x))^T$ and $f, h_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, .., m$, $(n \geq m)$ are twice continuously differentiable functions.

A point $x \in \mathbb{R}^n$ is said to be strictly feasible for problem (1) if $h(x) = 0$ and $x > 0$.

The Lagrangian function associated with problem (1) is

$$
\ell(x, y, z) = f(x) + h(x)^T y - x^T z \tag{2}
$$

where $y \in \mathbb{R}^m$ and $z \geq 0 \in \mathbb{R}^n$ are Lagrange multipliers associated with the equality and inequality constraints, respectively.

2.1 Perturbed KKT Conditions

For $\mu > 0$, the perturbed KKT conditions associated with problem (1) are

$$F_\mu(x, y, z) \equiv \begin{pmatrix} \nabla f(x) + \nabla h(x)y - z \\ h(x) \\ XZe - \mu e \end{pmatrix} = 0, \tag{3}$$

$$(x, z) > 0$$

where $X = \text{diag}(x), Z = \text{diag}(z), e = (1, \ldots, 1)^T \in \mathbb{R}^n$.

The perturbation parameter μ only affects the complementarity condition $XZe = 0$. For $\mu = 0$, these conditions are merely the KKT conditions associated with problem (1).

Comment 1 *Perturbed KKT conditions (3) are equivalent to the necessary conditions for the logarithmic barrier subproblem associated with problem (1). These conditions keep the iterates away from the boundaries and are not necessarily an ill–conditioned system of equations. See Sections 2.2 and 2.3 of Argáez [3].*

We base our numerical experimentation using the perturbed KKT conditions as a central framework to promote global convergence.

Definition 2.1 Centrality Region. *We define a centrality region associated with problem (1) as any collection of points that satisfy at least the perturbed complementarity condition with $(x, z) > 0$, parameterized by μ.*

3 Path–Following Strategies

In this section we present our philosophy of a path-following interior–point Newton's method as a globalization strategy to obtain an optimal solution of problem (1).

Definition 3.1 Path-Following Strategy. *For $\mu > 0$, and working from the interior ($(x, z) > 0$), apply a linesearch (trust region) Newton's method to the perturbed KKT conditions (or an equivalent system) until the iterate arrives to a specified proximity to the centrality region. Then decrease μ, specify a new proximity, and repeat the process.*

Remark 3.1 *Under appropriate conditions, an optimal solution will be obtained as μ approaches zero.*

To implement this strategy, we need to state three primary issues among others. First a centrality region, secondly how close we want to move about a centrality region, and third what merit function needs to be used to obtain the proximity. We address these issues in the next subsections.

3.1 Centrality Regions

The choice of a centrality region plays an important role in the implementation of a path–following strategy. We present two centrality regions that we implement in the path–following strategy.

3.1.1   Central Path

In their globalization strategy, El-Bakry et al [5] use the perturbed KKT conditions as a notion of a centrality region. This notion is defined as the collection of solutions, $v_\mu^* = (x_\mu^*, y_\mu^*, z_\mu^*)$, of the perturbed KKT conditions parameterized by $\mu > 0$, i.e.,

$$\begin{bmatrix} \nabla f(x) + \nabla h(x)y - z \\ h(x) \\ XZe - \mu e \end{bmatrix} = 0 \tag{4}$$

$$(x, z) > 0.$$

This notion is an extension of the central path from linear programming to nonlinear programming that inherits the following advantages:

1. Keeps the iterates away from the boundaries.

2. Promotes global convergence with non ill-conditioned systems of equations.

In nonlinear programming, this notion has the following disadvantages:

1. Far from the solution, a point in the central path for a particular μ may not exist.

2. Requiring iterates to be closed to the central path makes the algorithm quite expensive.

3.1.2   Quasi Central-Path

To overcome the disadvantages of the central–path, Argáez and Tapia [2] presented a relaxed notion of the central path as a centrality region. In their globalization strategy, they defined the notion of quasi central-path as the collection of points parameterized by $\mu > 0$ such that

$$\begin{bmatrix} h(x) \\ XZe - \mu e \end{bmatrix} = 0$$

$$(x, z) > 0.$$

Proposition 3.1 *The quasi central-path is equivalent to the region of strictly feasible points.*

Proof. For $\mu > 0$, $h(x) = 0$, $XZe = \mu e$, and $x, z > 0$ is trivially equivalent to $h(x) = 0$ and $x > 0$. This centrality region has two main advantages:

1. Near the solution, where the central path is well defined, the quasi central-path takes advantage in finding a point in this region since it is less restrictive than the central path.

2. Far from the solution, where we don't know if the central path is well defined, the quasi central-path presents a dramatic advantage since by Proposition 3.1 a point on the quasi central-path is just a strictly feasible point.

3.2 Proximity to the Centrality Region

Since by following the central path or the quasi central-path to reach an optimal solution can make the strategy quite expensive, we present some measures of how close an interior-point is from satisfying these centrality regions. Specifically, we use three different notions of proximity to the centrality region that El-Bakry et al [5], Argáez and Tapia [2], and Parada and Tapia [11] used in their globalization strategies.

3.2.1 Proximity to the Central Path

El-Bakry et al [5] used the following inequality

$$\min(XZe) \geq \sigma \frac{x^T z}{n}, \quad \sigma \in (0, 1) \quad \text{(C1)}$$

as a measure of proximity to the central path. This inequality is well known in the area of linear programming. We denote this centrality condition by C1.

3.2.2 Proximity to the Quasi Central-Path

Argáez and Tapia [2] and Parada and Tapia [11] in their globalization strategy present the following inequality

$$\|h(x)\|^2 + \|W^{-1}(XZe - \mu e)\|^2 \leq \gamma \mu, \quad \gamma \in (0, 1)$$

as a measure of proximity to the quasi central-path, where the weighting matrix W is given by

$$W = \begin{cases} (XZ)^{\frac{1}{2}}; & \text{Argáez-Tapia} \quad \text{(C2)} \\ I; & \text{Parada-Tapia} \quad \text{(C3)}. \end{cases}$$

The Argaez-Tapia is a weighted proximity measure and is denoted as C2, and the Parada-Tapia proximity measure is an unweighted centrality condition denoted by C3.

3.3 Merit Functions

Now we present some merit functions used as tools to reach centrality conditions C1, C2, and C3.

3.3.1 KKT Residual Function

To obtain centrality condition C1, El-Bakry et al [5] used the KKT residual function given by

$$\Phi(v) = \frac{1}{2}\|F(v)\|^2 \qquad \text{(M1)}$$

where $F(v) = 0$ are the KKT conditions for problem (1). This function is denoted by M1. This merit function has the following important property:

(P1) For any interior-point $v = (x, y, z)$, the Newton step Δv is a descent direction for the residual function $\Phi(v)$ for specific choices of μ, i.e.,

$$\nabla \Phi(v)^T \Delta v < 0 \quad \text{iff} \quad \mu < \frac{\|F(v)\|^2}{x^T z}.$$

Proof. See El-Bakry et al [5].

3.3.2 $\mathcal{M}_\mu$ Argáez-Tapia Function

To obtain centrality condition C2, Argáez and Tapia [2] present a modification of the augmented Lagrangian function as a merit function to be used in their globalization strategy, and is defined by

$$\mathcal{M}_\mu(x, z; y, \rho) = \ell(x, y, z) + \rho \, \Phi_\mu(x, z) \qquad \text{(M2)}$$

where $\ell(x, y, z)$ is the Lagrangian function associated with problem (1), ρ is a nonnegative penalty parameter, and $\Phi_\mu(x, z)$ is the penalty term. The penalty term is given by

$$\Phi_\mu(x, z) = \frac{1}{2} h(x)^T h(x) + x^T z - \mu \sum_{i=1}^{n} \ln(x_i z_i),$$

which is composed of the square norm of the equality constraints and a known potential reduction function. This merit function is denoted by M2.

3.3.3 $\mathcal{L}_\mu$ Parada-Tapia Function

To obtain centrality condition C3, Parada and Tapia [11] present another modified augmented Lagrangian function given by

$$\mathcal{L}_\mu(x, y, z; \rho) = \ell(x, y, z) + \rho \, \mathcal{C}_\mu(x, z) \qquad \text{(M3)}$$

where the penalty term is

$$\mathcal{C}_\mu(x, z) = \|h(x)\|^2 + \|XZe - \mu e\|^2.$$

The penalty term is composed of the square norm of the equality constraints and the square norm of the perturbed complementarity condition. Observe that the difference between M2 and M3 is the second term of its corresponding penalty terms. Merit function M3 replaces the potential reduction function in M2 with the norm squared of the perturbed complementarity condition.

The modified augmented Lagrangian functions M2 and M3 have two important properties:

(P1) For $\mu > 0$ and fixed. If $v_\mu^* = (x_\mu^*, y_\mu^*, z_\mu^*)$ satisfies the perturbed KKT conditions, then for sufficiently large ρ the primal solution x_μ^* satisfies

$$x_\mu^* = \operatorname{argmin} \ \mathcal{M}_\mu(x \, ; y_\mu^*, z_\mu^*, \rho),$$

$$x_\mu^* = \operatorname{argmin} \ \mathcal{L}_\mu(x \, ; y_\mu^*, z_\mu^*, \rho).$$

Proof. See Argáez and Tapia [2] and Parada and Tapia [11].

(P2) For $\mu > 0$ and fixed. The Newton step, Δv, obtained from the perturbed KKT conditions at an interior-point $v = (x, y, z)$ is a descent direction for ρ sufficiently large if and only if the point v is not on the quasi central-path, i.e.,

$$\nabla \mathcal{M}_\mu(v; \rho)^T \Delta v < 0,$$

$$\nabla \mathcal{L}_\mu(v; \rho)^T \Delta v < 0.$$

Proof. See Argáez-Tapia [2] and Parada-Tapia [11].

4 Path–Following Algorithms

In this section we present two global algorithms of the primal–dual interior–point algorithm presented by El–Bakry et al [5] which is a pure extension of the well known primal–dual interior–point algorithm for linear programming, presented in 1987 by Kojima, Mizuno and Yoshise [9], to nonlinear programming. Both algorithms are a damped Newton's method with linesearch applied to the perturbed KKT conditions. The first one is a linesearch globalization of the local primal–dual interior–point method for nonlinear programming introduced by El–Bakry et al [5] which we are implementing using merit functions M1, M2, and M3. The second one is the algorithm presented by Argáez and Tapia [2] which we are implementing with centrality condition C2 and C3 and merit functions M1, M2, and M3. Both algorithms are path–following strategies since they satisfy Definition (3.1).

Algorithm 1

Step 1. *Select a merit function M1, M2, or M3. Consider an initial interior-point*
$v_o = (x_o, y_o, z_o)$. *Choose* β, p, γ.
Step 2. *For k=0,1,2,....until convergence do*
Step 3. *Choose* $\sigma_k \in (0, 1), \mu_k = \sigma_k(\frac{(x^T z)}{n})$.
Step 4. *Solve the linear system*
$$F'(v_k)\Delta v_k = -F_{\mu_k}(v_k).$$

Step 5. *(Maintain x and z positive). Choose* $\tau_k \in (0, 1)$ *and set* $\alpha = min(1, \tau_k \bar{\alpha})$
where
$$\bar{\alpha} = min\left\{ \frac{-1}{min(X_k^{-1}\Delta x_k, -1)} , \frac{-1}{min(Z_k^{-1}\Delta z_k, -1)} \right\}.$$

Step 6. *(Steplength selection). Compute* $\tilde{\alpha} \in (0, \alpha]$ *such that*
$$min(X(\tilde{\alpha})Z(\tilde{\alpha})e) \geq \gamma(\frac{x(\tilde{\alpha})^T z(\tilde{\alpha})}{n}).$$
Step 7. *(Sufficient decrease). Find* $\alpha_k = p^t \tilde{\alpha}$ *where t is the smallest positive integer such that* α_k *satisfies*

$$Mj_{\mu_k}(v_k + \alpha_k \Delta v_k) \leq Mj_{\mu_k}(v_k) + \alpha_k \beta \nabla Mj_{\mu_k}(v_k)^T \Delta v_k. \text{ (See Comment 2)}$$

7

Step 8. *Set $v_k = v_k + \alpha_k \Delta v_k$.*

Comment 2 *For $j = 1$, the algorithm is the primal–dual interior–point Newton's method introduced by El-Bakry et al [5]. For $j = 2, 3$ the algorithm is a modification of El-Bakry et al's formulation. In these cases the algorithm uses the modified augmented Lagrangian merit functions presented by Argáez-Tapia and Parada-Tapia, respectively. The choice of the penalty parameter ρ, used to force a descent direction, is calculated as in Argáez-Tapia [2].*

Now we present Algorithm 2 which follows the quasi central-path defined in 3.1.2 as a centrality region.

Algorithm 2

Step 1. *Select a merit function M1, M2 or M3. Select centrality condition C2 or C3. Consider an initial interior-point $v_o = (x_o, y_o, z_o)$. Choose β, p, γ.*

Step 2. *For $k = 0, 1, 2, \ldots$ until convergence do*

Step 3. *Choose $\sigma \in (0, 1), \mu_k > 0$.*

Step 4. *Repeat*

    4.1 Solve the linear system $\qquad F^{'}(v_k)\Delta v_k = -F_{\mu_k}(v_k)$.

    4.2 (Maintain x and z positive). Choose $\tau_k \in (0, 1)$ and set $\tilde{\alpha} = min(1, \tau_k \bar{\alpha})$ where

$$\bar{\alpha} = min\left\{ \frac{-1}{min(X_k^{-1}\Delta x_k, -1)} \ , \ \frac{-1}{min(Z_k^{-1}\Delta z_k, -1)} \right\}.$$

    4.3 (Force a descent direction). Calculate ρ_k to ensure a Newton descent direction for a given merit function Mj. (See Comment 3).

    4.4 (Sufficient decrease). Find $\alpha_k = p^t \tilde{\alpha}_k$ where t is the smallest positive integer such that α_k satisfies

$$Mj_{\mu_k}(v_k + \alpha_k \Delta v_k; \rho_k) \leq Mj_{\mu_k}(v_k; \rho_k) + c_k \alpha_k \beta \nabla Mj_{\mu_k}(v_k)^T \Delta v_k$$

    4.5 Set $v_k = v_k + \alpha_k \Delta v_k$.

Step 5. *(Proximity to the quasi central-path)*

    5.1 If $v_k \notin Ci$, (See Comment 4)

$$go \ to \ Step \ 3$$

    5.2Else $\qquad\qquad\qquad\qquad go \ to \ Step \ 1$

Comment 3 *For $j = 1$, the Newton step is a descent direction for merit function M1, and therefore Substep 4.3 is not needed. For $j = 2, 3$ the penalty parameter ρ is chosen as in Argáez–Tapia [2].*

Comment 4 *The condition $v_k \notin Ci$ means that the updated point v_k does not satisfy the centrality condition Ci for $i = 2, 3$.*

Remark 4.1 *For a $\mu > 0$, a fundamental difference between the Algorithm 1 and Algorithm 2 is that for Algorithm 1 one iteration is enough to reach centrality condition C1, see El–Bakry et al [5], meanwhile for Algorithm 2 one iteration may not be enough to reach centrality condition C2, and C3, see Argáez-Tapia [2], Parada- Tapia[11].*

5 Numerical Experimentation

The first objective of this work is to analyze which of the nine strategies performs numerically better in obtaining an optimal solution for a set of test problems. These 9 options are summarized in the following table.

	C1	C2	C3
M1	$TR92 - 40$	Present Work	Present Work
M2	Present Work	$TR95 - 29$	Present Work
M3	Present Work	Present Work	$TR97 - 12$

The ordered couple (Mi, Cj) represents a path-following strategy to reach an optimal solution for problem (1). It is important to remark that even though the strategies (M1, C1), (M2, C2), and (M3, C3) are presented in technical reports [5],[2],[11], we have done our own implementation. The strategies (M2,C1), (M3,C1), (M1,C2), (M3,C2), (M1,C3) and (M2,C3) are studied for the first time in this work.

The second objective is to compare the numerical behavior of centrality conditions C1, C2 and C3 as measures of proximity to either the central path or quasi central-path.

The numerical experiments were performed on a set of test problems of Hock and Schittkowski [8], and Schittkowski [12] used by El-Bakry et al [5]. [8] and [12]. The numerical results are presented in Tables 1-3 and summarized in Graphs 1 and 2.

5.1 Implementation of the Algorithms

The numerical experiments are done on a SPARC station 4 running the SOLARIS Operating-System with 64 Megabytes of memory. The programs are written in MATLAB version 5a. We use a finite difference approximation to the Hessian of the Lagrangian function. In all cases, the number of Newton iterations reported are the number of times each algorithm solves the linear system associated with the problem until it generates a point that satisfies the following stopping criterion

$$\frac{\|F(v_k)\|_2}{1 + \|v_k\|_2} \le 10^{-8}.$$

In the implementation of the algorithms, the critical parameters are τ_k (percentage of movement to the boundary), σ_k (centering parameter), and μ_k (perturbation parameter). The parameters τ_k and σ_k, that depend on the choice of a merit function, are updated and initialized as follow:

- For merit function M1:
  $\tau_k = \max(.8, 1 - 100x_k^T z_k)$, $\sigma_k = \min(.2, 100x_k^T z_k)$, $k = 0, 1, \ldots..$

- For merit function M2:
  $\tau = .99995$, $\sigma_k = \min(.2, 100x_k^T z_k)$, $k = 0, 1, \ldots.$

- For merit function M3:
  $\tau_k = \max(.8, 1 - 100x_k^T z_k)$, $\sigma_k = \min(.2, 100x_k^T z_k)$, $k = 0, 1, \ldots.$

9

The perturbation parameter μ_k is initialized and updated depending on which centrality condition is implemented:

- Centrality condition C1:
  $\mu_k = \sigma_k \frac{x_k^T z_k}{n}, k = 0, 1, \ldots$.

- Centrality condition C2:
  Initialize $\mu_o = \sigma_o \frac{x_o^T z_o}{n}$, and set $W = (XZ)^{\frac{1}{2}}$.

  if $(\|h(x)\|^2 + \|(W)^{-1}(XZe - \mu e)\|^2) \leq (.5\mu_k)$,

  $$\mu_{k+1} = 10^{-2}\mu_k,$$

  else

  $$\mu_{k+1} = \mu_k.$$

- Centrality condition C3:
  Initialize $\mu_o = \frac{1}{w} * \min(.5, \|F(v_o)\|) * \|F(v_o)\|$ where v_o is the initial point, and w is the dimension of the problem.

  if $(\|h(x)\|^2 + \|XZe - \mu e\|^2) \leq (.8\mu_k)$,

  $$\mu_{k+1} = \frac{1}{w} * \min(.5, \|F_\mu(v_k)\|) * \|F_\mu(v_k)\|,$$

  else

  $$\mu_{k+1} = \mu_k.$$

5.2   Table Notation and Graphs

In Tables 1-3, we present the numerical results obtained for solving the 72 problems tested from [8] and [12]. We denote the merit functions and centrality conditions by the following abbreviations:

- M1 : ℓ_2-norm residual function,

- M2 : Argáez-Tapia modified augmented Lagrangian function,

- M3 : Parada-Tapia modified augmented Lagrangian function,

- C1 : Centrality condition number 1,

- C2 : Centrality condition number 2,

- C3 : Centrality condition number 3.

In each table, the central heading identifies the merit function used with the three different centrality conditions. The first column contains the problem number tested of [8] and [12], and the next three columns contain the number of Newton iterations necessary to obtain an optimal solution of the problem using centrality conditions C1, C2, and C3, respectively. Each solution obtained from the different strategies was verified, and it agrees with the solution reported in Hock and Schitwoski and Schitwoski books. The symbol (-) means that the problem tested did not converge within 100 iterations. We summarize the numerical results presented in Tables 1-3 in Graph 1 by adding the number of Newton iterations needed to solve the set of problems tested for each of the nine strategies. This is done only for problems that converge in all the options.

In particular, we present the number of Newton iterations that are needed to solve problem 13 from [8] in Graph 2. We compare the numerical results obtained using merit functions M1, M2 and M3 combined with centrality conditions C2 and C3.

5.3 Summary of Numerical Results

Now we present our interpretation of the numerical results shown in Graphs 1 and 2 for solving the set of test problems.

First, Graph 1 shows that C2, the weighted proximity measure to the quasi central-path, is the best centrality condition with any choice of merit function since it is more robust and requires fewer number of Newton iterations than C1 and C3. Centrality condition C3 has a better numerical behavior than centrality condition C1 in terms of number of problems that converge, but it requires more Newton iterations for solving the test problems. We observe that some problems do not converge with option C1, the proximity measure to the central path, because it takes shorter steps exceeding the allowable number of iterations (e.g. prob. 45,106), or when close to the solution the matrix becomes ill-conditioned (prob. 13). In the contrary, option C2 takes longer steps, and none of the problems failed due to an ill-conditioned matrix. Option C3, the unweighted proximity measure to the quasi central-path, has a similar behavior in robustness than option C2, but is quite expensive as option C1. This observations show that following the quasi central-path is an effective choice as centrality region for solving nonlinear programming problems.

Second, merit function M2 has an outstanding numerical behavior in terms of the number of Newton iterations than merit function M1 and M3. We also observe that merit function M1 is a more robust choice than the other two options. We investigated further why some of the problems failed with options M2 and M3. In the case of merit function M2, the starting point makes a difference because this function involves a logarithmic function. We observe that with a starting point that is not interior, any strategy using this merit function has problems to converge (prob. 74, 75). This difficulty is overcomed using an initial interior-point. This is not an issue for merit functions M1 and M3. For merit function M3, some of the problems did not converge because of large values of the penalty parameter ρ (prob. 2,26). We observe this is not a problem for merit function M2 because the range of values of this parameter is between 0 and 3.6.

Finally, Graph 1 shows that the best path-following strategy is (M2,C2) since it has a better numerically performance for solving the test problems in fewer number of Newton iterations. In terms of robustness, the strategy (M1,C2) represents a better choice, but strategy (M2,C2) becomes a competitive choice once some of the starting points are changed to be interior-points.

11

Now for problem 13, where strict complementarity does not hold, (i.e. the Jacobian is singular at the solution) Graph 2 shows that merit functions M2 and M3 with centrality condition C2 and C3 have an outstanding behavior in terms of number of Newton iterations necessary to reach an optimal solution. El-Bakry et al [5] stopped their algorithm in 100 iterations without reaching a solution of the problem. Yamashita [13] in 172 iterations reported a good approximation to the primal and dual solutions, but the norm of the KKT conditions associated to the problem was not close to zero.

6 Concluding Remarks

In this work we have implemented nine globalization strategies for solving nonlinear programming problems by combining any of the merit functions and centrality conditions presented by El-Bakry et al, Argáez-Tapia and Parada-Tapia (see [5], [2], [11]). Six of these strategies are presented and compared for the first time in this study. A general MATLAB code was written to solve nonlinear programming problems by using each of the nine strategies. Our numerical experimentation was done on the set of test problems presented by El-Bakry et al[5].

From our numerical experimentation, the best option to solve nonlinear programming problems using a path-following strategy given by Definition 3.1 is the strategy (M2,C2). Merit function M2 promotes good numerical behavior by making progress to the quasi central-path and by expecting a decrease in the objective function with moderate values of the penalty parameter. Furthermore, centrality condition C2 performs numerically better than the other two options with any combination of merit function in terms of robustness and total number of Newton iterations. Also, from our numerical results we observe that centrality conditions C2 and C3 have better numerical performance than centrality condition C1. Therefore we recommend to choose the quasi central-path as a centrality region and centrality condition C2 as an effective measure of proximity to this region. It is important to mention that using centrality condition C2, none of the problems failed due an ill-conditioned matrix, and we conjecture it should be used to avoid the boundaries close to the solution.

Further numerical and theoretical research is needed to establish the role that the quasi central-path and the centrality condition C2 play for solving nonlinear programming problems.

7 Acknowledgments

M1 vs C1, C2, C3											
Prob	C1	C2	C3	Prob	C1	C2	C3	Prob	C1	C2	C3
1	49	22	68	32	13	13	13	76	11	11	10
2	38	17	65	33	12	25	22	80	7	7	7
3	6	5	5	34	11	11	12	81	12	11	12
4	9	7	7	35	11	11	10	83	19	21	24
5	11	10	9	36	14	12	15	84	31	31	130
10	10	8	8	37	14	12	15	86	21	22	24
11	8	6	6	38	-	16	24	93	13	14	15
12	14	11	12	41	7	7	7	100	15	15	17
14	6	5	5	42	12	11	11	104	12	14	13
15	15	15	18	43	16	15	16	106	-	32	32
16	22	22	25	44	13	13	12	226	10	10	9
17	13	14	14	45	-	9	8	227	9	9	8
18	30	30	30	53	11	10	10	231	42	39	95
19	13	15	24	60	12	12	11	233	45	29	62
20	13	13	15	62	10	10	14	250	14	13	15
21	14	16	17	63	12	12	12	251	14	12	15
22	5	5	5	64	24	23	36	262	11	12	11
23	21	22	22	65	24	25	24	325	11	11	10
24	8	8	7	66	10	11	12	339	13	13	13
25	7	7	7	71	13	13	15	340	9	7	7
26	15	16	13	72	12	19	27	341	12	11	10
29	10	10	11	73	18	18	21	342	24	23	43
30	12	17	16	74	20	19	21	353	13	13	14
31	13	13	14	75	21	18	20	354	16	15	19

Table 1:

M2 vs C1, C2, C3											
Prob	C1	C2	C3	Prob	C1	C2	C3	Prob	C1	C2	C3
1	22	13	43	32	11	9	9	76	8	8	8
2	16	9	29	33	14	16	52	80	5	5	5
3	4	2	2	34	8	7	7	81	8	7	7
4	6	4	3	35	10	7	7	83	18	18	19
5	9	7	7	36	10	9	10	84	27	22	26
10	9	8	8	37	9	7	8	86	16	20	38
11	9	6	6	38	18	32	35	93	-	12	13
12	6	7	7	41	5	5	5	100	10	13	13
14	5	5	5	42	11	11	43	104	8	11	10
15	12	9	13	43	23	9	12	106	35	33	33
16	18	14	14	44	8	8	9	226	10	9	8
17	9	10	10	45	-	7	-	227	5	7	6
18	26	42	42	53	7	4	5	231	21	20	37
19	-	11	16	60	10	9	9	233	21	16	17
20	10	7	9	62	7	6	16	250	9	9	10
21	13	15	14	63	16	23	19	251	9	8	8
22	3	3	3	64	22	18	30	262	8	9	8
23	22	18	18	65	24	41	40	325	9	7	7
24	5	5	5	66	8	7	7	339	16	10	10
25	8	9	7	71	9	9	12	340	5	5	5
26	19	16	43	72	13	11	12	341	9	7	7
29	8	8	8	73	16	14	14	342	16	13	21
30	10	12	11	74	43	65	155	353	11	10	11
31	9	9	6	75	32	39	-	354	11	12	14

Table 2:

14

M3 vs C1, C2, C3											
Prob	C1	C2	C3	Prob	C1	C2	C3	Prob	C1	C2	C3
1	16	21	29	32	9	9	9	76	6	7	7
2	-	21	24	33	26	21	24	80	5	7	7
3	2	2	2	34	7	10	10	81	6	7	7
4	3	4	3	35	27	6	53	83	16	18	19
5	5	6	6	36	7	9	9	84	19	22	30
10	8	15	15	37	6	8	8	86	38	36	25
11	6	7	7	38	11	9	10	93	-	-	-
12	10	10	10	41	5	5	5	100	10	12	13
14	4	5	5	42	5	6	6	104	7	10	9
15	8	13	14	43	12	14	12	106	30	21	21
16	12	14	15	44	6	8	7	226	6	7	6
17	8	10	10	45	-	13	-	227	6	7	6
18	20	27	27	53	4	5	5	231	34	44	60
19	-	-	-	60	13	14	12	233	13	38	42
20	7	8	9	62	5	7	9	250	7	9	9
21	10	16	16	63	12	11	18	251	6	8	8
22	5	6	6	64	27	20	45	262	7	9	8
23	20	20	20	65	17	19	19	325	10	9	9
24	4	6	4	66	7	8	8	339	13	8	9
25	23	47	15	71	9	12	15	340	17	44	8
26	-	-	-	72	8	12	12	341	6	7	7
29	5	7	7	73	11	14	15	342	12	16	20
30	6	43	43	74	15	18	16	353	10	12	12
31	6	9	9	75	12	13	13	354	9	12	14

Table 3:

15
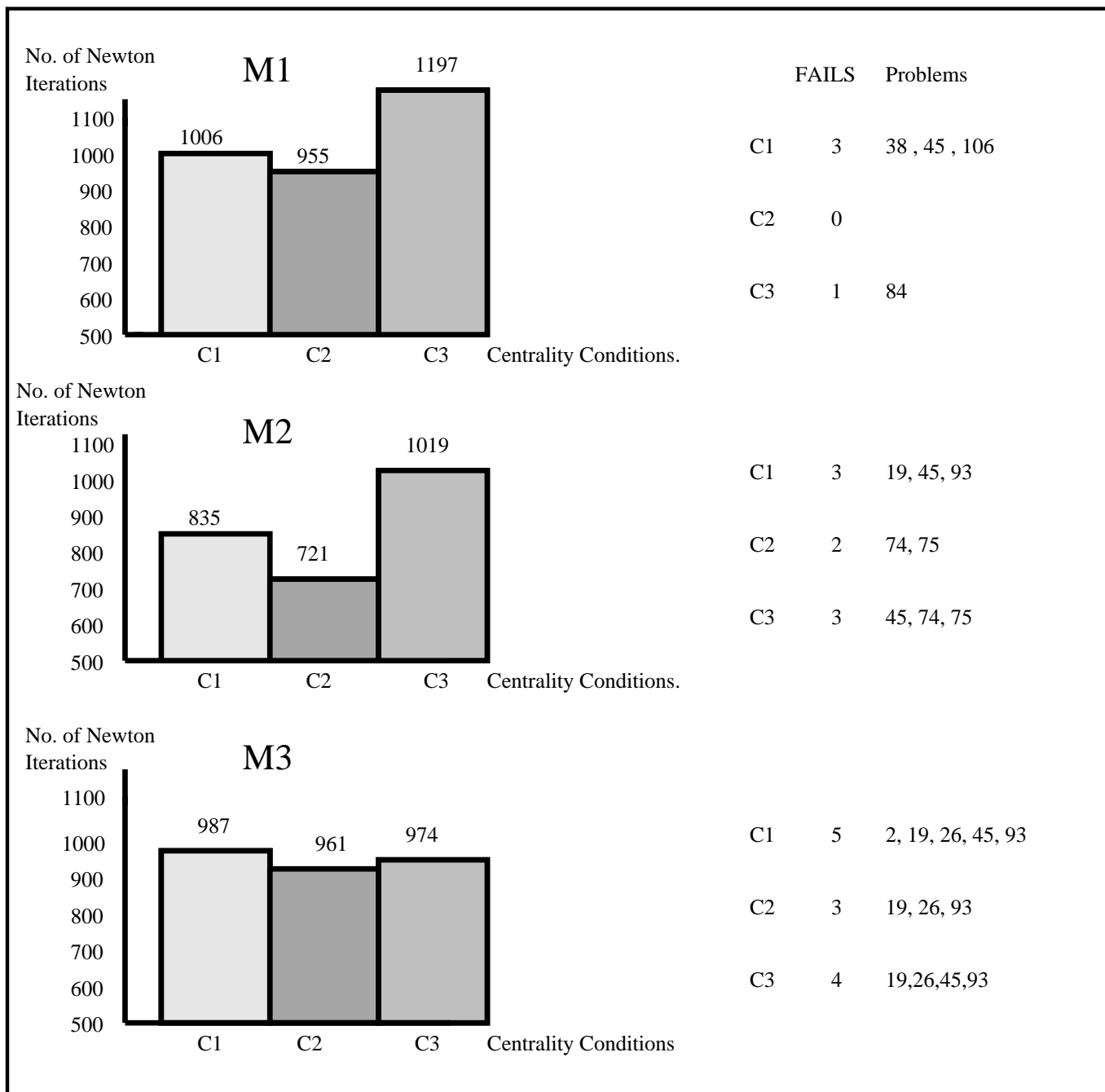
Figure 1: Graph 1

Problem 13		
	C2	C3
M1	65	64
M2	13	15
M3	13	16

Table 4:

Problem 13   (A Difficult Problem)
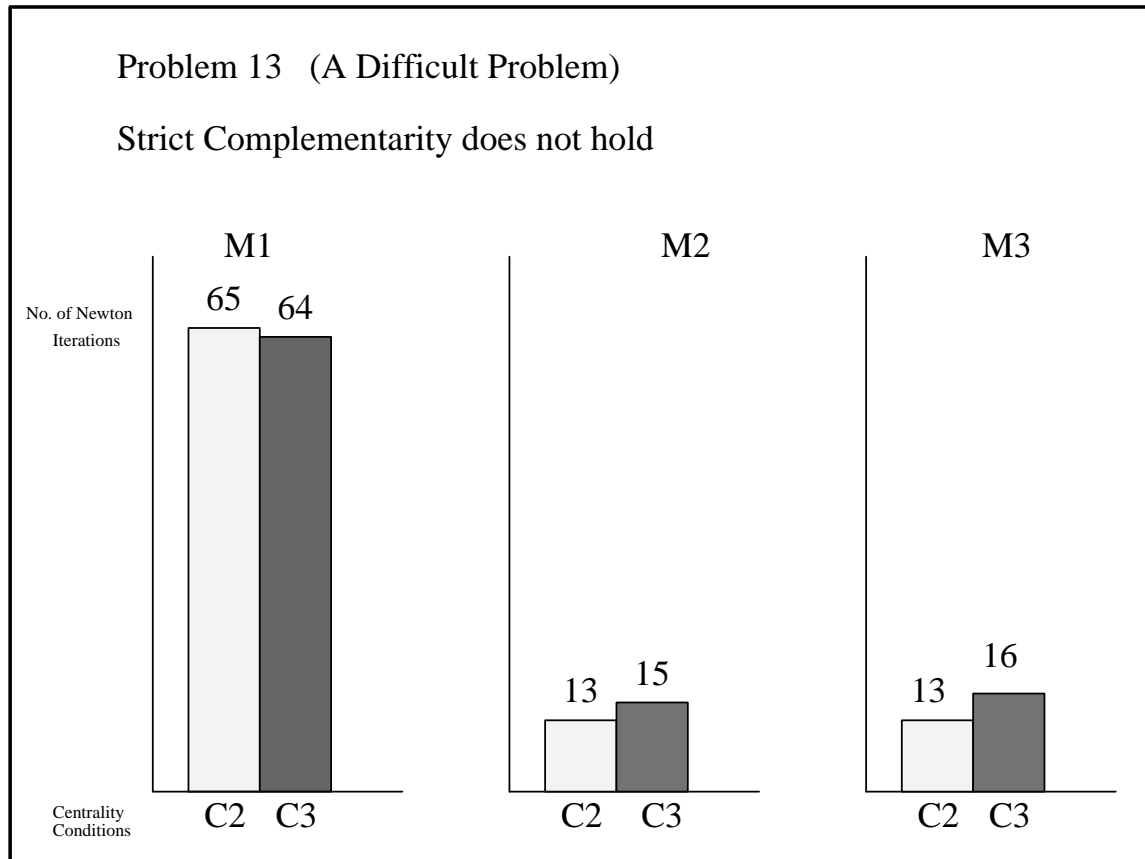
Strict Complementarity does not hold

Figure 2: Graph 2

References

[1] K. M. Anstreicher and J. P. Vial, *On the convergence of an infeasible primal-dual interior-point method for convex programming*, Tech. Rep. 93-34, Faculty of Technical Mathematics and Informatics, Delft University of Technology, The Netherlands.

[2] M. Argáez, and R. A. Tapia, *On the global convergence of a modified augmented Lagrangian linesearch interior point Newton method for nonlinear programming*, Tech. Rep. 95-29, Dept. of Computational and Applied Mathematics, Rice University, Houston, TX.

[3] M. Argáez, *Exact and inexact newton linesearch interior-point method algorithms for nonlinear programming problems*, Ph.D. thesis, Tech. Rep. 97-13, Dept. of Computational and Applied Mathematics, Rice University, Houston, TX.

[4] J. E. Dennis Jr. and R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice-Hall, Englewood Cliffs, New Jersey, (1983).

[5] A. S. El-Bakry, R. A. Tapia, T. Tsuchiya, and Y. Zhang, *On the formulation of the primal-dual interior point method for nonlinear programming*, J. Optim. Theory Appl., Vol. 89 (1996), pp. 507-541.

[6] A. V. Fiacco, G. P. McCormick, *Nonlinear programming: sequential unconstrained minimization techniques*, John Wiler, New York, (1968).

[7] M. González-Lima, R. A. Tapia, and F. Potra, *On effectively computating the analytic center of the solution set by primal-dual interior-point methods*, Tech. Rep. 95-38, Dept. of Computational and Applied Mathematics, Rice University, Houston, TX. To appear in SIAM Journal on Optimization.

[8] W. Hock, and K. Schittkowski, *Test examples for nonlinear programming code*, Springer Verlag, New York, (1981).

[9] Kojima, M. Mizuno, S. Yoshise, *A primal-dual interior-point method for linear programming*, Progress in Mathematical Programming Interior-Point and Related Methods, Nimrod Megiddo editor, Springer Verlag, New York, (1989), pp 29-47.

[10] H. J. Martinez, Z. Parada, and R. A. Tapia, *On the characterization of Q-superlinear convergence of quasi-Newton interior-point methods for nonlinear programming*, Bol. Soc. Mat. Mexicana Vol.1 (1995), pp. 137-148.

[11] Z. Parada, *Computational experience with a modified augmented Lagrangian merit function in a primal-dual interior- point method*, Ph.D. thesis, Tech. Rep. 97-12, Dept. of Computational and Applied Mathematics, Rice University, Houston, TX.

[12] K. Schittkowski, *More test examples for nonlinear programming codes*, Springer-Verlag, New York, (1987).

[13] H. Yamashita, *A globally convergent primal-dual interior point method for constrained optimization*, Tech. Rep., Mathematical System Institute, Inc.. Japan, (1992).