On Generalized Primal-Dual Interior-Point Methods with Non-uniform Complementarity Perturbations for Quadratic Programming

Altuğ Bitlislioğlu and Colin N. Jones

Abstract

This technical note discusses convergence conditions of a generalized variant of primal-dual interior point methods. The generalization arises due to the permitted case of having a non-uniform complementarity perturbation vector, which is equivalent to having different barrier parameters for each constraint instead of a global barrier parameter. Widely used prediction-correction methods and recently developed coordinated schemes can be considered as specific cases of the non-uniform perturbation framework. For convex quadratic programs, the polynomial complexity result of the standard feasible path following method with uniform perturbation vector is extended to the generalized case by imposing safeguarding conditions that keep the iterates close to the central-path.

I. Introduction

We consider the following convex quadratic optimization problem:

$$\begin{aligned} \min_{x,s}. \quad & \frac{1}{2}x^T H x + q^T x \\ \text{s. t.} \quad & Ax + b + s = 0 \ , \\ & s \geq 0 \ , \qquad\qquad (\lambda) \end{aligned} \tag{1}$$

where λ is the dual variable associated to the positivity constraint. This problem can be efficiently solved with primal-dual interior point methods that enjoy worst-case polynomial complexity [6].

In standard primal-dual methods, the problem is solved with the help of a smoothing on the KKT system. The smoothing is obtained by perturbing the complementarity conditions with a uniform and positive vector. In practice, certain correction terms are added to the perturbation vector [4], [7], to account for the non-linearity of the complementarity equation. This type of correction can be considered as a 'non-uniform' perturbation vector since the 'corrected' perturbation parameter for each constraint can be different. Another motivation for using a non-uniform perturbation arises from the implementation of primal-dual methods in a multi-agent setting [1] where it is desirable to use a different perturbation vector for each set of local constraints in order to reduce the communication burden. In this technical note we will analyze the convergence properties of primal-dual interior points with non-uniform complementarity perturbation vector for quadratic programming. Following closely the analysis of [6], we will show that in this generalized setting, identical worst-case complexity results can be obtained, provided that certain centrality conditions are enforced. These conditions are automatically satisfied in the uniform setting, and provide guidelines for ensuring safe operation for practical implementations of coordinated or prediction correction type primal-dual interior-point methods.

In the following we will first introduce the standard primal-dual interior point method and then provide the complexity result for the non-uniform setting.

II. Primal-Dual Interior Point Method

This section provides a summary of standard primal-dual interior point framework. See [8], [9] for a more in depth treatment and [5] for an up to date survey.

The first order optimality conditions for problem (1), or the Karush-Khun-Tucker(KKT) system, can be written as

$$r(x, s, \lambda) = 0. \tag{2}$$

Where $r = (r_{\text{dual}}^T, r_{\text{comp}}^T, r_{\text{prim}}^T)^T$ and the following definitions are used:

$$\begin{aligned} r_{\text{dual}}(x, \lambda) &= Hx + q + A^T \lambda \ , \\ r_{\text{prim}}(x, s) &= Ax + b + s \ , \\ r_{\text{comp}}(\lambda, s) &= \Lambda S \mathbf{1} \ , \end{aligned} \tag{3}$$

where $\Lambda = \text{diag}(\lambda)$, $S = \text{diag}(s)$ and the operator diag(.) constructs a diagonal matrix from the input vector.

*This technical note includes some of the explanatory material from [1] and provides the proof of Theorem 1 of [1].

This work has received support from the European Research Council under the European Unions Seventh Framework Programme (FP/2007-2013), ERC Grant Agreement 307608.

The authors are with Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland.
altug.bitlislioglu - colin.jones @ epfl.ch

Primal-dual interior point (PDIP) methods operate by applying Newton's method on a smoothed version of the nonlinear set of equations (2). The smoothing is obtained by perturbing the complementary slackness condition:

$$r_{\text{cent}}(\beta, \lambda, s) = \Lambda S \mathbf{1} - \beta, \quad \beta > 0 \ , \tag{4}$$

and using the modified function $\hat{r}(\beta) = (r_{\text{dual}}^T, r_{\text{cent}}^T(\beta), r_{\text{prim}}^T)^T$. In order to retain the optimizer of the original KKT system, the perturbation parameter is decreased progressively. As $\beta \to 0$, the primal-dual pairs converge to the optimizer of (1) following the so called 'central path'. In general the perturbation vector β is obtained by multiplying a scalar with a vector of ones. However we will let it attain different values for each constraint. The Newton direction $\Delta z = (\Delta x, \Delta y, \Delta s, \Delta\lambda, \Delta\mu, \Delta\gamma)$, for the perturbed KKT system can be found by solving the following linear system of equations:

$$\begin{bmatrix} H & 0 & A^T \\ 0 & \Lambda & S \\ A & \mathbf{I}_s & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta s \\ \Delta \lambda \end{bmatrix} = -\hat{r}(\beta) \ . \tag{5}$$

'Long step' methods use the maximum step size that preserves strict positivity of the pair λ, s such that $(s + t\Delta s, \lambda + t\Delta\lambda) > 0$ holds, whereas 'short step' variants further restrict the step size to preserve 'centrality' and/or to ensure sufficient decrease in the residuals. The primal dual interior point method will terminate when the duality gap measure

$$\eta = \lambda^T s$$

and the primal dual residuals are reduced below desired thresholds. η becomes equal to the actual duality gap, when the iterates are primal and dual feasible, $r_{\text{dual}} = 0$, $r_{\text{primal}} = 0$ [2]. For a uniform perturbation vector β, when the perturbed KKT system is solved, that is we obtain a point on the central path, we have that

$$\hat{\eta} := \frac{\eta}{M} = \frac{\lambda^T s}{M}, \ \ \hat{\eta}\mathbf{1} = \beta \ ,$$

where $\lambda, s \in \mathbb{R}^M$, M is the total number of inequality constraints and the average complementarity value $\hat{\eta}$ provides a measure for the duality gap. In practice it is not necessary to wait for convergence to the central path, before reducing the perturbation parameter. In the standard 'feasible path following' method, for which convergence and complexity results are relatively easy to provide [8], the perturbation vector at each step is chosen as

$$\beta = \sigma\hat{\eta}\mathbf{1}, \quad \sigma \in (0, 1) \ .$$

To ensure that proximity to the central path is preserved, the iterates can be constrained to stay within certain neighborhoods [3], [8].

III. Convergence with non-uniform perturbation

For brevity, we write the Netwon system in compact form

$$\mathcal{M}\Delta z = -\begin{bmatrix} r_{\text{primal,dual}} \\ \Lambda S\mathbf{1} - \beta \end{bmatrix} \ .$$

Consider the modified Newton step with a non-uniform perturbation vector on the complementarity condition:

$$\mathcal{M}\Delta z = -\begin{bmatrix} r_{\text{prim,dual}} \\ \Lambda S\mathbf{1} - \Lambda S\xi \end{bmatrix} \ , \tag{6}$$

where the vector ξ specifies how much change in each complementarity pair is desired, with $\xi \in \mathbb{R}^M$ and $\xi > 0$. For a given $\sigma > 0$, ξ can be selected as $\xi_j = \sigma\frac{\hat{\eta}}{\lambda_j s_j}$ to recover the standard uniform perturbation vector with σ. Note that any corrector applied together with the uniform-perturbation can be considered as a non-uniform perturbation vector.

In this section, we will state the convergence conditions for quadratic programs (QP) in the feasible case, where the primal-dual residuals are equal to 0, to make the analysis simpler, as is common in the interior point literature [6], [10]. The results can be extended to the non-feasible case and for LPs and QPs, a feasible point can be obtained with reformulations of the problem [5], [6].

We restrict the iterations to lie inside the 'symmetric neighborhood' [6],

$$\mathcal{N}_s = \{(x, s, \lambda) \in \mathcal{F}^0 : \ \frac{1}{\psi}\hat{\eta} \geq \lambda_j s_j \geq \psi\hat{\eta}, \quad \forall j \in [1, M]\} \tag{7}$$

for some $1 \gg \psi > 0$, where the strictly feasible set $\mathcal{F}^0$ is defined as

$$\mathcal{F}^0 = \{(x, s, \lambda) \ : \ (r_{\text{prim}}, r_{\text{dual}}) = 0, \ (s, \lambda) > 0\}.$$

Furthermore we restrict the reduction vector ξ to satisfy the following criteria,

$$\exists c\ : 0 < c < 1,\ \frac{\sum_j \xi_j \lambda_j s_j}{M} = c\hat{\eta} \tag{8}$$

$$\frac{1}{\phi}\frac{c\hat{\eta}}{\lambda_j s_j} \geq \xi_j \geq \phi \frac{c\hat{\eta}}{\lambda_j s_j} \tag{9}$$

for some $1 \gg \phi > \psi$. The constraints (9), (8) enforce that the next point we would like to arrive with the Newton step, does not have to be on the central path as in the uniform case, but it should be at least strictly within the $\mathcal{N}_s$ and aim to reduce the overall duality gap.

By applying these conditions, we will ensure that the linearization error in the Newton step is bounded, and there is always some finite step-size that can reduce the duality-gap and keep the system within $\mathcal{N}_s$. The following theorem states the convergence and complexity result.

Theorem 1: If the problem (1) is a convex QP, given $\epsilon > 0$, suppose that a starting point $(x_0, s_0, \lambda_0) \in \mathcal{N}_s$ satisfies

$$\eta_0 \leq \frac{1}{\epsilon^\kappa} \tag{10}$$

for some positive constant κ. Let $\{\eta_k\}$ be a sequence generated by the iteration scheme, which takes steps using the Newton relation (6) with admissible parameters ϕ, ψ and reduction vectors ξ_k that satisfy (8), (9), with some $\alpha, \gamma \in (0,1), \gamma < 1-\alpha$ and $c_k \leq \gamma$. Then, there exists an admissible step-size $t_k \sim \mathcal{O}(\frac{1}{M})$ and an index $K \sim \mathcal{O}(M \log(\frac{1}{\epsilon}))$, such that

$$\eta_k \leq \epsilon,\ \ \forall k \geq K.$$

The analysis and the proof will be similar to that of [6] by showing that starting from a point within $\mathcal{N}_s$, it is possible to keep the linearization error small and obtain an admissible step-size of complexity $\mathcal{O}(\frac{1}{M})$ that reduces the duality gap sufficiently and keep the iterates within $\mathcal{N}_s$. Allowing non-uniformity necessitates additional constraints (9), (8) on the perturbation vector and results in different bounding terms compared to [6], however the complexity result is identical.

We start the analysis by deriving bounds on the linearization error.

Lemma 1: For $(x, s, \lambda) \in \mathcal{N}_s$, the Newton direction resulting from (6) satisfies the following relations;

$$0 \leq \Delta\lambda^T \Delta s \leq \theta \lambda^T s \tag{11}$$

$$-\theta\lambda^T s \leq \Delta\lambda_j \Delta s_j \leq \sqrt{2}\theta\lambda_j s_j \tag{12}$$

where θ is defined as

$$\theta := 2^{-3/2}\max\ (\frac{c}{\phi\psi} - 1, 1 - \phi\psi c). \tag{13}$$

Proof: We start by proving (11). Under the feasibility assumption, primal-dual residuals are set to 0 and from (5), we have the following relations

$$\Delta s = -A\Delta x,$$
$$A^T \Delta\lambda = -H\Delta x,$$

and therefore

$$\Delta s^T \Delta\lambda = -\Delta x^T A^T \Delta\lambda$$
$$= \Delta x^T H \Delta x\ .$$

Since the cost function is assumed to be convex, we have $H \succeq 0$, which validates the left-hand side inequality in (11).

For the right-hand side of (11) we need a technical result ([5], Lemma 3.3.). Given two vectors $u, v \in \mathbb{R}^n$, if $u^T v \geq 0$ then,

$$\|UV\mathbf{1}\| \leq 2^{-3/2}\|u+v\|^2 \tag{14}$$

where $U := \text{diag}(u),\ V := \text{diag}(v)$. With relation (14) at hand, we start with the non-uniformly perturbed centrality equation

$$S\Delta\lambda + \Lambda\Delta s = \Lambda S(\xi - \mathbf{1}). \tag{15}$$

Multiply both sides with $(\Lambda S)^{-1/2}$;

$$\Lambda^{-1/2}S^{1/2}\Delta\lambda + \Lambda^{1/2}S^{-1/2}\Delta s = (\Lambda S)^{1/2}(\xi - \mathbf{1}). \tag{16}$$

Since $\Delta\Lambda, \Lambda, S, \Delta S$ are all diagonal matrices we have that

$$\|\Delta\Lambda\Delta S\mathbf{1}\| = \|\Lambda^{-1/2}S^{1/2}\Delta\lambda\Lambda^{1/2}S^{-1/2}\Delta s\|$$

Next we use (14) with $u := \Lambda^{-1/2}S^{1/2}\Delta\lambda$ and $v := \Lambda^{1/2}S^{-1/2}\Delta s$. Note that the condition $u^T v \geq 0$ is satisfied since $\Delta\lambda^T\Delta s \geq 0$.

$$\begin{aligned}
\|\Delta\Lambda\Delta S\mathbf{1}\| &= \|\Lambda^{-1/2}S^{1/2}\Delta\lambda\Lambda^{1/2}S^{-1/2}\Delta s\| \\
&\leq 2^{-3/2}\|\Lambda^{-1/2}S^{1/2}\Delta\lambda + \Lambda^{1/2}S^{-1/2}\Delta s\|^2 \\
&= 2^{-3/2}\|(\Lambda S)^{1/2}(\xi - \mathbf{1})\|^2 \\
&= 2^{-3/2}\sum_j(\lambda_j s_j)(\xi_j - 1)^2 \\
&\leq 2^{-3/2}\|\xi - \mathbf{1}\|_\infty^2\lambda^T s \qquad \text{due to}(\lambda_j s_j > 0,\ \forall j)
\end{aligned}$$

Relying on the constraints (9), (8) imposed on the non-uniform perturbation vector ξ, we can obtain a parametric expression for $\|\xi - \mathbf{1}\|_\infty$

$$\|\xi - \mathbf{1}\|_\infty^2 = \max\ (\frac{c}{\phi\psi} - 1, 1 - \phi\psi c)$$

which is equal to $\theta/(2^{-3/2})$, by definition (13). Therefore, we can write the bound on $\|\Delta\Lambda\Delta S\mathbf{1}\|$ as

$$\|\Delta\Lambda\Delta S\mathbf{1}\| \leq \theta\lambda^T s$$

Finally using Cauchy-Schwarz inequality we arrive at the upper bound on $\Delta\lambda^T\Delta s$;

$$\Delta\lambda^T\Delta s = (\Delta\Lambda\Delta S\mathbf{1})^T\mathbf{1} \leq \|\Delta\Lambda\Delta S\mathbf{1}\|\|\mathbf{1}\| = \theta\lambda^T s$$

which proves the right-hand side of (11).

Next, we prove the relation (12) that provide bounds on $\Delta\lambda_j\Delta s_j$. We write equation (16) element-wise and square both sides

$$\begin{aligned}
&\lambda_j^{-1/2}s_j^{1/2}\Delta\lambda_j + \lambda_j^{1/2}s_j^{-1/2}\Delta s_j = \lambda_j^{1/2}s_j^{1/2}(\xi_j - 1) \\
\Rightarrow\quad &\lambda_j^{-1}s_j\Delta\lambda_j^2 + \lambda_j s_j^{-1}\Delta s_j^2 + 2\Delta\lambda_j\Delta s_j = (\xi_j - 1)^2\lambda_j s_j
\end{aligned}$$

Since $\lambda, s > 0$, we have that

$$\begin{aligned}
&2\Delta\lambda_j\Delta s_j \leq (\xi_j - 1)^2\lambda_j s_j \\
\Rightarrow &\Delta\lambda_j\Delta s_j \leq 0.5\|(\xi_j - \mathbf{1})\|_\infty^2\lambda_j s_j
\end{aligned}$$

Using the relation $|\Delta\lambda_j\Delta s_j| \leq \|\Delta\Lambda\Delta S\mathbf{1}\|_\infty \leq \|\Delta\Lambda\Delta S\mathbf{1}\|$, and the definition of θ, we arrive at

$$-\theta\lambda^T s \leq \Delta\lambda_j\Delta s_j \leq \sqrt{2}\theta\lambda_j s_j$$

■

In the following, we will show that a finite step-size can be obtained such that the neighborhood constraints are preserved and a sufficient decrease in the duality gap is achieved.

For a given direction, we describe the values after a step with a step-size $t \in (0, 1]$ as

$$\lambda_j(t) = \lambda_j + t\Delta\lambda_j, \quad s_j(t) = s_j + t\Delta s_j\ .$$

Lemma 2: For a given point $(x, s, \lambda) \in \mathcal{N}_s$, and the Newton direction resulting from (6), if the step-size $t \in (0, 1]$ satisfies the following conditions;

$$t \leq \frac{c(\phi - \psi)}{(\psi + M)\theta} \tag{17}$$

$$t \leq \frac{c(\phi - \psi)}{\sqrt{2}\theta\phi} \tag{18}$$

then the new point attained by taking a step in the Newton direction with step-size t stays in the neighborhood $\mathcal{N}_s$, that is

$$(x(t), s(t), \lambda(t)) \in \mathcal{N}_s. \tag{19}$$

Proof: Symmetric neighborhood $\mathcal{N}_s$ implies lower and upper bound constraints on complementarity pairs and feasibility of the iterates. We start with the lower bound constraint. Using Lemma 1, the centrality equation (15) and neighborhood constraints on the current point $\lambda_j s_j \geq \psi\hat{\eta}$ and the perturbation vector $\xi_j \geq \phi \frac{c\hat{\eta}}{\lambda_j s_j}$ we obtain an initial lower bound on $\lambda_j(t)s_j(t)$

$$\begin{aligned}\lambda_j(t)s_j(t) &= (\lambda_j + t\Delta\lambda_j)(s_j + t\Delta s_j)\\ &= \lambda_j s_j(1-t) + t\xi_j\lambda_j s_j + t^2\Delta\lambda_j\Delta s_j\\ &\geq \psi\hat{\eta}(1-t) + t\phi c\hat{\eta} - t^2\theta M\hat{\eta}\end{aligned}$$

Next, we derive an upper bound on $\hat{\eta}(t)$, using the relation (8), centrality equation (15) and Lemma 1

$$\begin{aligned}\hat{\eta}(t) &= \lambda(t)^T s(t)/M\\ &= \left(\lambda^T s + t(c-1)\lambda^T s + t^2\Delta\lambda^T\Delta s\right)/M\\ &\leq \left(\lambda^T s + t(c-1)\lambda^T s + t^2\theta\lambda^T s\right)/M\\ &= \hat{\eta}\left((1-t) + tc + t^2\theta\right)\end{aligned} \tag{20}$$

Using the upper bound on $\hat{\eta}(t)$ and lower bound on $\lambda_j(t)s_j(t)$ we can write a sufficient condition for $\lambda_j(t)s_j(t) \geq \psi\hat{\eta}(t)$,

$$\begin{gathered}\psi\hat{\eta}(1-t) + t\phi c\hat{\eta} - t^2\theta M\hat{\eta} \geq \psi\hat{\eta}\,(1-t) + t\psi c\hat{\eta} + \psi t^2\theta\hat{\eta}\\ (\phi-\psi)c\hat{\eta} \geq t\left((\psi+M)\theta\right)\hat{\eta}\end{gathered}$$

which is identical to condition (17).

Next, we tackle the upper bound constraint on $\lambda_j(t)s_j(t) \leq \frac{1}{\psi}\hat{\eta}(t)$ which is equivalent to

$$\begin{aligned}&\lambda_j s_j(1-t) + t\xi_j\lambda_j s_j + t^2\Delta\lambda_j\Delta s_j \leq\\ &\qquad\qquad \frac{1}{\psi}\left(\hat{\eta}(1-t) + tc\hat{\eta} + t^2\Delta\lambda^T\Delta s/M\right)\end{aligned}$$

Seeing that $\lambda_j s_j \leq \frac{1}{\psi}\hat{\eta}$, and $\xi_j \leq \frac{1}{\phi}\frac{c\hat{\eta}}{\lambda_i s_i}$ a sufficient condition is

$$\begin{aligned}&\frac{1}{\psi}\hat{\eta}(1-t) + t\frac{1}{\phi}c\hat{\eta} + t^2\Delta\lambda_j\Delta s_j \leq\\ &\qquad\qquad \frac{1}{\psi}\hat{\eta}(1-t) + t\frac{1}{\psi}c\hat{\eta} + t^2\frac{1}{\psi}\Delta\lambda^T\Delta s/M\end{aligned}$$

$$\begin{gathered}t\frac{1}{\phi}c\hat{\eta} + t^2\Delta\lambda_j\Delta s_j \leq t\frac{1}{\psi}c\hat{\eta} + t^2\frac{1}{\psi}\Delta\lambda^T\Delta s/M\\ t^2\left(\Delta\lambda_j\Delta s_j - \frac{1}{\psi}\Delta\lambda^T\Delta s/n\right) \leq tc\hat{\eta}\left(\frac{1}{\psi} - \frac{1}{\phi}\right).\end{gathered}$$

Since $\Delta\lambda_j\Delta s_j \leq \sqrt{2}\theta\lambda_j s_j \leq \sqrt{2}\theta\frac{1}{\psi}\hat{\eta}$ and $\Delta\lambda^T\Delta s \geq 0$ another sufficient condition is

$$\begin{aligned}t\sqrt{2}\theta\frac{1}{\psi}\hat{\eta} &\leq c\hat{\eta}\left(\frac{1}{\psi} - \frac{1}{\phi}\right)\\ t &\leq \frac{\psi c(\frac{1}{\psi} - \frac{1}{\phi})}{\sqrt{2}\theta} = \frac{c(\phi-\psi)}{\sqrt{2}\theta\phi}\,.\end{aligned}$$

which is equivalent to (18).

For a LP and QP, the KKT system is composed of linear equations, except for the complementarity equation. Therefore the Newton step with $r_{\text{dual}} = 0, r_{\text{primal}} = 0$, will preserve the feasibility for any step-size, from which we conclude that a step-size $t \in (0,1]$ satisfying (17), (18) will keep the iterates within $\mathcal{N}_s$. ■

Lemma 3: For a given point $(x,s,\lambda) \in \mathcal{N}_s$, parameter $\alpha \in (0,1)$ and the Newton direction resulting from (6), if the step-size t, and the mean reduction parameter c, satisfy the conditions;

$$1 - c > \alpha, \quad 0 < t \leq \frac{1-c-\alpha}{\theta} \tag{21}$$

then the new duality gap η(t) satisfies the reduction requirement

$$\eta(t) \leq (1-\alpha t)\,\eta\,. \tag{22}$$

Proof:

From (20), we have that

$$\eta(t) \leq (1-\gamma(t))\,\eta$$

with $\gamma(t) = t(1-c) - t^2\theta$. Therefore, a sufficient condition for satisfying (22) is

$$\gamma(t) = t(1-c) - t^2\theta \geq \alpha t$$
$$\Rightarrow t^2\theta \leq t(1-c-\alpha) \Rightarrow \; 0 < t \leq \frac{1-c-\alpha}{\theta}$$

which is equivalent to (21).■

In order to obtain the convergence and complexity result, first we will show that there exists a $t \sim \mathcal{O}(\frac{1}{M})$ that satisfies (18), (17) and (21). Then, application of Thorem 3.2 from [10] will provide the desired result.

The set of admissible step-sizes are given by;

$$\{t \mid 0 < t \leq \min\left(\frac{1-c-\alpha}{\theta}, \frac{c(\phi-\psi)}{(\psi+M)\theta}, \frac{c(\phi-\psi)}{\sqrt{2}\theta\phi}\right)\} \tag{23}$$

where $\theta := 2^{-3/2}\text{max}\ (\frac{c}{\phi\psi} - 1, 1 - \phi\psi c)$.

Choose $\alpha = 0.1$, $\psi = 0.1$ and $\phi = 0.2$ and fix reduction parameter $c = 0.8$ for every step. For these values θ is given by $2^{-3/2}(\frac{0.8}{0.2*0.1} - 1) \simeq 13.8$ It can be shown that the step-size selection $\hat{t} = \frac{1}{200M}$ will be always admissible, which provides $t \sim \mathcal{O}(\frac{1}{M})$.

Proof of Theorem 1: We have already shown that with judicious selection of parameters, it is possible to find an admissible $t \sim \mathcal{O}(\frac{1}{M})$. Plugging this into the reduction requirement (22), we get

$$\eta_{k+1} \leq (1 - \frac{\delta}{M})\eta_k, \quad k = 1, 2, \ldots, \tag{24}$$

for some constant $\delta > 0$. From here, we can directly apply Thorem 3.2 from [10] and state the result. We provide the proof for completeness. First we take the $\log$ of both sides in (24)

$$\log \eta_{k+1} \leq \log(1 - \frac{\delta}{M}) + \log \eta_k \ .$$

Writing the right hand side recursively and using (10) we get

$$\log \eta_{k+1} \leq k \log(1 - \frac{\delta}{M}) + \log \eta_0 \leq k \log(1 - \frac{\delta}{M}) + \kappa \log(\frac{1}{\epsilon}) \ .$$

The log function has the property that

$$\log(1 + x) \leq x, \quad \forall x > -1 \ .$$

Since $\alpha, t < 1$, this leads to

$$\log \eta_{k+1} \leq k(-\frac{\delta}{M}) + \kappa \log(\frac{1}{\epsilon}) \ .$$

For the convergence criterion $\eta_k < \epsilon$ to be satisfied, it is sufficient to have

$$k(-\frac{\delta}{M}) + \kappa \log(\frac{1}{\epsilon}) \leq \log \epsilon \ ,$$

which is satisfied for all k, such that

$$k \geq K = (1 + \kappa)\frac{M}{\delta} \log \frac{1}{\epsilon} \ .$$

■

References

[1] A. Bitlislioglu and C. N. Jones. On coordinated primal-dual interior-point methods for multi-agent optimization. In *56th IEEE Conference on Decision and Control (to appear), Melbourne, Australia*, December 2017.

[2] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004.

[3] M. Colombo and J. Gondzio. Further development of multiple centrality correctors for interior point methods. *Computational Optimization and Applications*, 41(3):277–305, December 2008.

[4] J. Gondzio. Multiple centrality corrections in a primal-dual method for linear programming. *Computational Optimization and Applications*, 6(2):137–156, September 1996.

[5] J. Gondzio. Interior point methods 25 years later. *European Journal of Operational Research*, 218(3):587–601, May 2012.

[6] J. Gondzio. Convergence Analysis of an Inexact Feasible Interior Point Method for Convex Quadratic Programming. *SIAM Journal on Optimization*, 23(3):1510–1527, January 2013.

[7] S. Mehrotra. On the Implementation of a Primal-Dual Interior Point Method. *SIAM Journal on Optimization*, 2(4):575–601, November 1992.

[8] J. Nocedal and J. Wright, S. *Numerical Optimization*. Springer, New York, 2000.

[9] F. A. Potra and S. J. Wright. Interior-point methods. *Journal of Computational and Applied Mathematics*, 124(12):281–302, December 2000.

[10] S. J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, 1997.