Antoine Bambade, Fabian Schramm, Sarah El Kazdadi, Stéphane Caron, Adrien Taylor, Justin Carpentier

To cite this version:

HAL Id: hal-04196897
https://inria.hal.science/hal-04196897
Submitted on 5 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
A. Companion report organisation

This companion report provides the proof associated to the PROXQP algorithm introduced in [1]. It is organized as follows. Section -B provides more details about which properties motivate the update rule presented in Section VI-B. Section -D establishes the global convergence of PROXQP (Theorem 1). It illustrates that BCL algorithm [2] finds appropriate proximal steps to match a desired convergence speed for primal feasibility $p^k$. More precisely, BCL decreases $\mu^k$ until the inequality $p^k \leq \epsilon_{bcl}$ holds for all iteration starting from $k$. At this stage of the algorithm, PROXQP amounts to a primal-dual proximal point algorithm, and we rely on [3] Proposition 1.2 to ensure the global convergence of PROXQP. We study the simplified Algorithm 2, which assumes that at each iteration $k$ of the algorithm $p^k = \mu^k$ compared to Algorithm 1, where $\rho$ is fixed. Updating $\rho$ is not desirable in practice since it would require a numerical factorization at each iteration.

### Section -B: More details about Section VI-B update rule

<table>
<thead>
<tr>
<th>Section</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section -B</td>
<td>More details about Section VI-B update rule</td>
</tr>
<tr>
<td>Section -C</td>
<td>Notations and technical properties</td>
</tr>
<tr>
<td>Section -D</td>
<td>Missing technical elements.</td>
</tr>
<tr>
<td></td>
<td>Proof of Lemma 1</td>
</tr>
<tr>
<td></td>
<td>Proof of Lemma 2</td>
</tr>
<tr>
<td></td>
<td>Proof of Lemma 3</td>
</tr>
<tr>
<td>Section -E</td>
<td>Proof of Theorem 1</td>
</tr>
<tr>
<td>Section -F</td>
<td>Proof of Lemma 1</td>
</tr>
</tbody>
</table>

**TABLE I: Organization of the companion report.**

B. Details about the update rule presented in Section VI-B

The update rule presented in Section VI-B is motivated by the following intuition. Assume the steady state $x_{\text{goal}}$ is still reachable by an unconstrained LQR scheme (see the grey line in Figure 1).

$$
\min_{x_t \in \mathbb{R}^n, u_t \in \mathbb{R}^m} x_T^T Q x_T + \sum_{t=0}^{T-1} x_t^T Q x_t + u_t^T R u_t,
$$

subject to $x_{t+1} = A x_t + B u_t$, $x_0 = x_{\text{init}}$, with a solution sequence $\{x_t^{\text{LQR}}\}_t, \{u_t^{\text{LQR}}\}_t$. As the constraints are linear, strong duality holds, and the primal of problem (1) $p^{\text{LQR}}$ equals its dual $\delta^{\text{LQR}}$ for some dual multipliers $\{y_t^{\text{LQR}}\}_t$.

### Proof of Theorem 1

If we compare now $\|x_t^{\text{LQR}} - x_{\text{goal}}\|_2^2$ with $\|x_1 - x_{\text{goal}}\|_2^2$ assuming that both $x_1^{\text{LQR}} > \bar{x}$ and $\hat{x}_1 > \bar{x}$ we see

$$
\|\hat{x}_1^{\text{LQR}} - x_{\text{goal}}\|_2^2 = \|x_1^{\text{LQR}} - \bar{x}\|_2^2 + \|\bar{x} - x_{\text{goal}}\|_2^2 + 2(\bar{x} - x_{\text{goal}})^T (\bar{x} - x_{\text{goal}}),
$$

and

$$
\|x_1 - x_{\text{goal}}\|_2^2 = \|x_1 - \bar{x}\|_2^2 + \|\bar{x} - x_{\text{goal}}\|_2^2 + 2(\bar{x} - x_{\text{goal}})^T (\bar{x} - x_{\text{goal}}),
$$

Thus we have:

$$
\|x_1^{\text{LQR}} - x_{\text{goal}}\|_2^2 = \|x_1 - \bar{x}\|_2^2 + 2(x_1^{\text{LQR}} - \bar{x})^T (\bar{x} - x_{\text{goal}}).
$$

1 Corresponding author. bambade.antoine@gmail.com
2 Inria - Département d’Informatique de l’École normale supérieure, PSL Research University.
3 École des Ponts, Marne-la-Vallée, France.
We introduce a maximal monotone operator \( I \) where \( \partial \) the Fenchel convex conjugate of a function is defined as \( \text{sup} \{ \phi(x) - y^\top z \} \). Hence, \( T \) \( T \leq \) \( \text{C} \). Further notations and technical properties are necessary for establishing \( \text{P} \) that are necessary for establishing \( \text{Q} \) (see Algorithm 2) is satisfied at infinitely many iterations. It follows from the second point that \( \text{ProxQP} \) corresponds to an appropriate type of approximate primal-dual proximal point method, and hence inherits all of its nice convergence properties. This is formally stated in Section -D.

C. Further notations and technical properties

This section provides a few additional technical ingredients that are necessary for establishing \( \text{ProxQP} \) convergence proof. The Fenchel convex conjugate of a function \( \phi : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{ \infty \} \) is defined as \( \phi^* : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{ \infty \} \) through

\[
\phi^*(y) \overset{\text{def}}{=} \sup_{z \in \mathbb{R}^m} \{ y^\top z - \phi(z) \} = - \inf_{z \in \mathbb{R}^m} \{ \phi(x) - y^\top z \}. \tag{3}
\]

We introduce a maximal monotone operator \( T^{-1}_L(v, t) \) (see for more details [4]), that encodes the set of solutions of the shifted problem (QP) by \( (t, v) \):

\[
T_L(x, z) \overset{\text{def}}{=} \{ (v, t) \mid (v, t) \in \partial L(x, z) \},
\]

\[
T^{-1}_L(v, t) \overset{\text{def}}{=} \arg \min_{x \in \mathbb{R}^n} \{ L(x, z) - x^\top v + z^\top t \},
\]

where \( \partial \) \( L \) refers to

\[
\partial L(x, z) \overset{\text{def}}{=} \left( \nabla f(x) + C^\top z \right) / \left( Cx - \partial I^*_L(z) \right), \tag{4}
\]

where \( I_u \) denotes the indicator function of the constraint set \( Cx \leq u \). The Lagrangian function and \( T_L(x, z) \) encodes the saddle sub-differential for (QP) [5 Section 4.3], \( T_L \) and \( T^{-1}_L \) are generally "point to set" operators. From \( T_L \) and for \( \mu > 0 \) we also introduce the resolvent of \( T_L \) and \( T^{-1}_L \):

- the resolvent of \( T_L \) [4]:

\[
P_\mu \overset{\text{def}}{=} \left( I + \frac{1}{\mu} T_L \right)^{-1}, \tag{5}
\]

- the resolvent of \( T^{-1}_L \) [4]:

\[
Q_\mu \overset{\text{def}}{=} \left( I + \mu T^{-1}_L \right)^{-1} = I - P_\mu. \tag{6}
\]

More precisely, \( T^{-1}_L(0, 0) \) is non-empty, it encodes the set of solutions to (QP). When \( Q \) is feasible, the KKT conditions for (QP) form a polyhedral variational inequality [6 Section 3D]. Hence, \( \exists \alpha > 0, \tau > 0 \) such that \( \forall (u, v) \in \mathbb{R}^{n+m} \):

\[
\|(u, v)\| \leq \tau \implies \text{dist}_{T^{-1}_L(0, 0)}(u, v) \leq \alpha \|(u, v)\|, \forall (x, z) \in T^{-1}_L(0, v)
\]

\( a \) and \( \tau \) are constants describing the geometry of the QP. Rockafellar [7 Thm. 2] has established that it holds as soon as \( T_L(0, 0)^{-1} \) is not empty. That is, \( \forall \mu > 0, \forall (x, z) \in \mathbb{R}^n \times \mathbb{R}^m \):

\[
\mu \|Q_\mu(x, z)\| \leq \tau \implies \text{dist}_{T_L(0, 0)^{-1}}(P_\mu(x, z)) \leq \frac{\alpha \mu}{1 + a^2 \mu^2} \text{dist}_{T_L(0, 0)^{-1}}(x, z). \tag{8}
\]

D. Convergence of ProxQP

The main convergence result for ProxQP is provided by Theorem 1. For proving this theorem (in Section -E), we proceed with the following strategy. This strategy relies on a slightly different formulation of ProxQP provided by Algorithm 2 (it is necessary for the proof to consider the conceptual infinite sequence of iterates generated by the algorithm).

- First, subsection -D1 shows that the "if condition" of ProxQP (see Algorithm 2) is satisfied at infinitely many iterations.

- Second, subsection -D2 is more accurate and shows that the "else condition" of ProxQP (see Algorithm 2) is entered only for a finite number of iterations.

- It follows from the second point that ProxQP corresponds to an appropriate type of approximate primal-dual proximal point method, and hence inherits all of its nice convergence properties. This is formally stated in Section -F.

1) We enter the "if" of Algorithm 2 an infinite number of times

**Lemma 1.** Under Assumption 1 and Assumption 2, \( \forall k \in \mathbb{N}, \exists K \geq 0 \) such that the condition \( \|[Cx^{k+K+1} - u]_+\|_\infty \leq \epsilon^{k+K}_{\text{bcl}} \) of Algorithm 2 is satisfied.

**Proof.** The following proof proceeds by contradiction. That is, we show that there cannot exist a \( k \geq 0 \) such that \( \forall K \geq 0 \) it holds that \( p^{k+K} \overset{\text{def}}{=} \|[Cx^{k+K+1} - u]_+\|_\infty > \epsilon^{k+K}_{\text{bcl}} \) (as defined in (8)). For convenience, and without loss of generality, we assume \( k = 0 \).

By contradiction hypothesis, we thus have that \( p^K > \epsilon^{K}_{\text{bcl}} \) holds \( K \geq 0 \). Therefore, \( \mu^K = \mu^0(\mu^K) \) and \( \epsilon^{K}_{\text{bcl}} = \epsilon^{0}_{\text{bcl}}(\mu^0) \alpha^{\epsilon(\mu^K)}_{\text{bcl}} \) and hence \( \mu^K \rightarrow 0 \) and \( \epsilon^{K}_{\text{bcl}} \rightarrow 0 \) as \( K \rightarrow \infty \), since \( \mu^0 \in (0, 1) \) and \( \alpha_{\text{bcl}} \in (0, 1/2) \). For establishing the desired statement, we show that under the contradiction hypothesis, one has

\[
\epsilon^{K}_{\text{bcl}} < \|[Cx^{K+1} - u]_+\|_\infty \leq A\sqrt{\mu^{K}}, \tag{9}
\]

for all \( K \geq 0 \), thereby reaching a contradiction. Indeed, with \( \alpha_{\text{bcl}} \in (0, 1/2) \), the term \( \sqrt{\mu^K} \) goes faster to zero (as a function of \( K \)) than \( \epsilon^{K}_{\text{bcl}} \). More precisely, the hyperparameter choices within ProxQP ensure that \( \exists K_0 \geq 0 \) such that \( \forall K \geq K_0, \epsilon^{K}_{\text{bcl}} > A\sqrt{\mu^K} \).

For proving our claim, we, therefore, focus on proving (9) in the following lines.

First, at iteration \( K + 1 \), the pair \( (x^{K+1}, z^{K+1}) \) is an approximate solution to (1). In other words, by defining intermediary variables, we have:

\[
u^K \overset{\text{def}}{=} \nabla f(x^{K+1}) + C^\top z^{K+1} + \mu^K(x^{K+1} - x^K),
\]

\[
u^{K} \overset{\text{def}}{=} [C_{x} x^{K+1} - u + \mu^{K}(z^{0} + (\alpha - 1) z^{K+1})_+]_+ - z^{K+1} + 1 \mu^{K},
\]
We are now properly armed to guarantee that where we used the facts that for a feasible point $u$ of the strongly convex function $(\ref{strongly_convex})$, equivalently look at $u$ in the sense of Assumption 2, we get:

$$M_{\rho,\mu,K}(x^{K+1}, z^{K+1}) \leq M_{\rho,\mu,K}(x,0).$$

As $\{x^K\}$ is bounded by Assumption 2, then $\exists M > 0$ independent of $K$ such that

$$M_{\rho,\mu,K}(x^{K+1}, z^{K+1}) \leq f(x) + M \mu K,$$

(10)

where we used the facts that for a feasible point $x$, $|Cx - u + z^0\mu + K|$ (as $Cx - u \leq 0$), and that $\|u^K, v^K\|_\infty \leq \epsilon^K$ implies that $u^K/\mu \leq \epsilon$ and that $\|v^K\|_\infty/\mu \leq \epsilon$.

It follows from the definition of $M_{\rho,\mu,K}$ and nonnegativity of the $||\cdot||_2$ terms that:

- we can bound

$$M_{\rho,\mu,K}(x^{K+1}, z^{K+1}) \geq f(x^{K+1}) + \frac{1}{2\sqrt{\mu}} \|Cx^{K+1} - u + \mu K(z^0 + (\alpha - 1)z^{K+1})\|_2^2$$

which, using (10), can be used to reach

$$\|Cx^{K+1} - u + \mu K(z^0 + (\alpha - 1)z^{K+1})\|_2 \leq \sqrt{2\mu K} f(x) - f(x^{K+1}) + M \mu K).$$

(11)

From (11), since $\{x^K\}_K$ is bounded and $f$ is a quadratic, we deduce that there exists some $M_1 > 0$ independent of $K$ such that

$$\|Cx^{K+1} - u + \mu K(z^0 + (\alpha - 1)z^{K+1})\|_2 \leq \sqrt{\mu K} M_1.$$

(12)

- We can bound

$$M_{\rho,\mu,K}(x^{K+1}, z^{K+1}) \geq f(x^{K+1}) + \frac{1}{\sqrt{\mu}} \|z + \frac{1}{\sqrt{\mu}} v^K\|_2^2$$

which, using (10), can be used to reach

$$\sqrt{\mu} \|z^{K+1}\|_2 \leq \sqrt{\frac{2}{\sqrt{\mu}}} \sqrt{f(x) - f(x^{K+1}) + M \mu K} + \sqrt{\frac{2}{\mu}} \|v^K\|_2.$$}

It shows that $\sqrt{\mu} \|z^{K+1}\|_2$ is bounded since $\{x^K\}$ is bounded and $f$ is a quadratic function.

We are now properly armed to guarantee that $\forall K > 0$:

$$\|Cx^{K+1} - u\|_\infty \leq A \sqrt{\mu K},$$

(13)

for some $A > 0$. Indeed, it holds component-wise

$$|Cx^{K+1} - u|_+ - \mu K(z^0 + (\alpha - 1)z^{K+1})|_+ \leq |Cx^{K+1} - u + \mu K(z^0 + (\alpha - 1)z^{K+1})|_+,$$

(14)

where $|.|$ is the absolute value taken component-wise. We can deduce from (14) that

$$\|Cx^{K+1} - u\|_2 \leq \|Cx^{K+1} - u + \mu K(z^0 + (\alpha - 1)z^{K+1})\|_2$$

$$\leq \sqrt{\mu K} \|z^0 + (\alpha - 1)z^{K+1}\|_2$$

(15)

Hence, from (15), using (12) and the fact that $\{\sqrt{\mu K} (z^0 + (\alpha - 1)z^{K+1})\}_K$ is bounded by some $M_2 > 0$ (since $\{\sqrt{\mu K} z^{K+1}\}_K$ is bounded and $\{\sqrt{\mu K}\}_K$ converges to zero) we get

$$\|Cx^{K+1} - u\|_2 \leq (M_1 + M_2) \sqrt{\mu K}.$$

To reach the desired claim, we use (13) together with the contradiction hypothesis that $\epsilon_{bsl} < \|Cx^{K+1} - u\|_\infty$ which allows reaching the contradiction that $\forall K > 0$:

$$\epsilon_{bsl} \leq A \sqrt{\mu K}.$$

(12)

2) We enter the "else" of Algorithm 2 a finite number of times

As previewed in Section 4, this section establishes that the "else" condition of Algorithm 2 is only entered a finite number of times. The proof can be summarized as follows.

- If the "else" condition is never satisfied, then the result directly holds.
- Otherwise, since Lemma 1 ensures the "if" condition is entered an infinite number of times, we consider the following setting. Without loss of generality, we can assume that the "else" condition is entered at iteration $k$ for the $K$th time and that it is followed by a "if" condition. It implies that

$$\mu = \mu_0 (\mu_f)_K, \epsilon_{bsl} = \epsilon_{bsl}^0 (\mu_f)^{\alpha_{bsl}}.$$

We denote by $N > 2$ the number of consecutive iterations from $k$ for which we enter the "if" condition, meaning that

$$\mu^{k+N} = \mu_0 (\mu_f)_K, \epsilon_{bsl}^{k+N-1} = \epsilon_{bsl}^0 (\mu_f)^{\alpha_{bsl}+(N-1)\beta_{bsl}}.$$

In this situation, one can establish that for QPs, there exists constants $M > 0$ and $b \geq 1$ (describing the geometry of the problem at hand) such that the primal feasibility can be upper-bounded for $N \geq 2$ by

$$\|Cx^{k+N} - u\|_\infty \leq M (b \mu)^{N-1},$$

(16)

For entering the "else" condition in this scenario at the $(k+N)$th step, it is therefore necessary to be in the situation where

$$\epsilon_{bsl}^{k+N-1} < \|Cx^{k+N} - u\|_\infty \leq M (b \mu)^{N-1},$$

Substituting the expressions for $\epsilon_{bsl}^{k+N-1}$ and $\mu^k$, we arrive to the following necessary condition for entering the "else" at iteration $(k+N)$:

$$\epsilon_{bsl}^0 (\mu_f)^{\alpha_{bsl}+(N-1)\beta_{bsl}} \leq M (b \mu_0 (\mu_f)_K)^{N-1}.$$
On the contrary, it is therefore sufficient to satisfy
\[
M(b y^0(\mu_f)^K)^{N-1} \leq \epsilon_{\text{tol}}^2(\mu_f^0(\mu_f)^K)^{\alpha \text{tol} + (N-1)\beta \text{tol}}
\]
for not entering the “else”. Given that \(\alpha \text{tol} \in (0, 1/2)\), \(\beta \text{tol} \in (0, 1)\) with \(\beta \text{tol} + \alpha \text{tol} < 1\), and \(\mu_f \in (0, 1)\), it can be established that this condition is met for all \(N \geq 2\) as soon as
\[
\max\left(\frac{\log(b y^0(\mu_f)^{1-\alpha \text{tol}})}{\epsilon_{\text{tol}}^2}, \frac{\log(b y^0(\mu_f)^{1-\beta \text{tol}})}{\epsilon_{\text{tol}}^2}\right), K_{\tau} \leq K,
\]
for some \(K_{\tau} \in \mathbb{N}\) (for the pair \((\alpha, \tau)\) from \([5]\), which describes the geometry of the QP at hand). It shows that the “else” condition is entered only a finite number of times, thereby arriving to the desired conclusion.

Our main target is thus to establish that \((17)\) holds, as the desired conclusion nicely follows from this fact. For reaching this target, Lemma 2 provides an intermediary result which ensures that for \(K\) sufficiently large (i.e., larger than \(K_{\tau}\) mentioned in \((17)\)), the technical property \((8)\) holds for all the iterates of PROXQP.

**Lemma 2.** Under Assumption 1, Assumption 2, and if \(\mu_k \to 0\) (i.e., we enter an infinite number of times the “else” condition), for any \(\tau > 0\) \(\exists K_{\tau} \in \mathbb{N}\) such that for all \(\forall k \geq K_{\tau}\), \(\mu_k\|Q_{\mu_k}(x_k, z_k)\|_2 \leq \tau\).

**Proof.** If \(\mu_k \to 0\), then by definition of \(Q_{\mu_k}\)
\[
\|Q_{\mu_k}(x_k, z_k)\|_2 = \| (1 - \alpha \mu_k) P_{\mu_k}(x_k, z_k) \|_2 \\
\leq \left\| (x_{\text{prox}} + (1 - \alpha \mu_k)(x_{\text{prox}} - k + 1) - P_{\mu_k}(x_k, z_k) \right\|_2.
\]
(18)

As \(\mu_k \to 0\), combining \((18)\) with \((19)\) allows obtaining the upper bound
\[
\mu_k\|Q_{\mu_k}(x_k, z_k)\|_2 \leq \mu_k \left\| (x_{\text{prox}} + (1 - \alpha \mu_k)(x_{\text{prox}} - k + 1) - P_{\mu_k}(x_k, z_k) \right\|_2 \\
\leq \mu_k \left\| (x_{\text{prox}} - k + 1) - P_{\mu_k}(x_k, z_k) \right\|_2 + \mu_k \epsilon_k.
\]
(20)

Since \(\{ (x_k, z_k) \}\) is bounded (see Assumption 1), and since \(\mu_k \to 0\) (exponentially in the number of times we enter the “else” condition), we conclude from \((20)\) that \(\forall \tau > 0\), there \(\exists K_{\tau} \in \mathbb{N}\), such that \(\forall k \geq K_{\tau}\), \(\mu_k\|Q_{\mu_k}(x_k, z_k)\|_2 \leq \tau\).

Finally, Lemma 3 establishes our main targets. More precisely, we make use of Lemma 2 to establish first the inequality \((16)\) and then \((17)\).

**Lemma 3.** Under Assumption 1 and Assumption 2, \(\exists N_{\text{max}} \in \mathbb{N}\) such that for all \(k \geq N_{\text{max}}\) the condition \(\mu^{k+1} \leq \epsilon_{\text{tol}}\) of Algorithm 2 is satisfied (i.e., we enter the “if” condition).
\[ T_{\varepsilon}^{-1}(0,0), \text{ which can be further upper bounded via a triangle inequality:} \]
\[ \| x^{k+N} - z_{k+N} \|_2 \leq \| x^{k+N} - P_{\mu_k}(x^{k+N-1}, z^{k+N-1}) \|_2 \]
\[ + \| P_{\mu_k}(x^{k+N-1}, z^{k+N-1}) - P_{\mu_k}(x^{k+N-1}, z^{k+N-1}) \|_2. \]

Hence we also have:
\[ \| x^{k+N} - P_{\mu_k}(x^{k+N-1}, z^{k+N-1}) \|_2 \leq \| x^{k+N} - z^{k+N} \|_2 \]
\[ + \| P_{\mu_k}(x^{k+N-1}, z^{k+N-1}) - P_{\mu_k}(x^{k+N-1}, z^{k+N-1}) \|_2. \]

Now, using \( k \geq K_\varepsilon \) from Lemma 2 (for a specific \((a, \tau)\) such that \( \| w \|_{2,\text{f}} \) holds for the QP at hand), then using the property (8) provides
\[ \| P_{\mu_k}(x^{k+N-1}, z^{k+N-1}) - P_{\mu_k}(x^{k+N-1}, z^{k+N-1}) \|_2 \]
\[ \leq \frac{a \mu_k}{\sqrt{1 + a^2(\mu_k)^2}} \text{dist}_{T_{\varepsilon}(0,0)-1}(x^{k+N-1}, z^{k+N-1}) \]
\[ \leq \mu_k \left\| \left( x^{k+N-1} - z^{k+N-1} \right) \right\|_2. \]

Hence incorporating (26) in (25) leads to
\[ \| x^{k+N} - z^{k+N} \|_2 \leq \frac{a \mu_k}{\mu_k} \left\| \left( x^{k+N-1} - z^{k+N-1} \right) \right\|_2. \]

where we also used the fact that \( (x^{k+N}, z^{k+N}) \) satisfies (11) at accuracy \( k^{N-1} \) similarly to (19). Iteratively expanding the upper bound \( (27) \) \( N-1 \) times leads to
\[ \| x^{k+N} - z^{k+N} \|_2 \leq \mu_k \left\| \left( x^{k+N-1} - z^{k+N-1} \right) \right\|_2 \]
\[ \leq w \left( \sum_{l=0}^{N-1} \frac{k^{N-1-l}}{\mu_k} + (a \mu_k)^N \left\| \left( x^{k+N} - z^{k+N} \right) \right\|_2. \]

Because \( k^{N-1-l} = c^0(\mu_k)^{N-l} \), we deduce from (28) that
\[ \| x^{k+N} - z^{k+N} \|_2 \leq w(\mu_k)^{N-1} \left\| \left( x^{k+N} - z^{k+N} \right) \right\|_2 \]
\[ \leq w(\mu_k)^{N-1} \sum_{l=0}^{N-1} (a)^l + B(\mu_k)^N, \]

for some \( B > 0 \), since \( \{ (x^k, z^k) \} \) is bounded by Assumption 2). Finally, incorporating (29) into (22) provides the bound
\[ \| x^{k+N} - u \|_\infty \leq e \| C \|_2 \left( w(\mu_k)^{N-1} \sum_{l=0}^{N-1} (a)^l + B(\mu_k)^N \right), \]

for some constant \( e > 0 \) (from the norm equivalences between \( \| . \|_\infty \) and \( \| . \|_2 \) in \( \mathbb{R}^m \)). Depending on the value of \( a \), we distinguish three different cases to show the inequality (16):

- if \( a \in (0,1) \), then, since \( \{ \mu_k \} \) is bounded by 1 (since \( \mu^0 \in (0,1) \) and \( \mu_f \in (0,1) \)), we can upper bound the geometric sum in (30) with
\[ \| Cx^{k+N} - u \|_\infty \leq e \| C \|_2 \left( w \frac{1}{1-a} + B \right)(\mu_k)^{N-1}, \]
and hence (16) holds with \( b = 1 \).

- if \( a > 1 \), then we can upper bound the geometric sum with (using again that \( \{ \mu_k \} \) is bounded by 1)
\[ \| Cx^{k+N} - u \|_\infty \leq e \| C \|_2 \left( w \frac{1}{1-a} + B \right)(\mu_k)^{N-1}, \]

and hence (16) holds with \( b = a \).

- if \( a = 1 \), then we can upper bound the geometric sum with (using again that \( \{ \mu_k \} \) is bounded by 1)
\[ \| Cx^{k+N} - u \|_\infty \leq e \| C \|_2 \left( wN + B \right)(\mu_k)^{N-1}, \]

and hence (16) holds with \( b = 2 \).

Thus in any case, there exists some \( M > 0 \) and \( b \geq 1 \), such that (16) holds. To conclude, in order to enter at least \( N \geq 2 \) consecutive times in the “if” condition it is sufficient to satisfy
\[ M(b\mu^0(\mu_k)^K)^{N-1} \leq c_{\text{bcl}}(\mu^0(\mu_f)^K)^{\alpha_{\text{bcl}}+(N-1)\beta_{\text{bcl}}}, \]
which is equivalent to
\[ \log(b(\mu_0)^{1-\beta_{\text{bcl}}}) + (N-2) \log(b(\mu_0)^{1-\beta_{\text{bcl}}}) \]
\[ \leq -\log(\mu_f)(1-\alpha_{\text{bcl}} - \beta_{\text{bcl}})K \]
\[ -\log(\mu_f)(1-\beta_{\text{bcl}})(N-2)K. \]

Since \( \mu_f \in (0,1) \), \( \alpha_{\text{bcl}} \in (0,1/2) \), \( \beta_{\text{bcl}} \in (0,1) \) with \( \alpha_{\text{bcl}} + \beta_{\text{bcl}} < 1 \), then (31) holds for any \( N \geq 2 \) as soon as
\[ \max[\{ \log(b(\mu_0)^{1-\beta_{\text{bcl}}}) \}, \{ \log(b(\mu_0)^{1-\beta_{\text{bcl}}}) \}], K \leq K, \]
which establishes (17). In other words, we enter the “else condition” only a finite number of times (which is bounded above by the value of the previous max).

E. Proof of Theorem 1

Theorem 1 (Convergence of ProXQP). Under Assumption 1 (existence of a solution to (QP)), Assumption 2 (bounded iterates), and with all parameters set as required in the inputs of Algorithm 2, the iterates \( \{ (x^k, z^k) \} \) of ProXQP converge to a solution \( (\bar{x}, \bar{z}) \) of (QP).

Proof. The proof consists in showing that ProXQP asymptotically corresponds to a proximal point algorithm (PPA). Then, we leverage the classical global convergence guarantees for PPA, which is automatically inherited by ProXQP. To show
this link between PROXQP and PPA, (i) we first recall PPA global convergence properties, and then (ii) show that in a finite number of iterations PROXQP is a fixed-step PPA.

When (QP) is feasible, the KKT conditions \(0 \in T_\mathcal{E}(x, z)\) forms a polyhedral variational inequality \([6]\) Section 3D. In such settings, if we consider a sequence \(\{\mu^k\}_k\), bounded below by some \(\mu^\infty > 0\), and \(\{\eta^k\}_k\) some summable sequence, then the inexact proximal point iteration

\[
(x^{k+1}, z^{k+1}) \approx_{\eta^k} P_{\mu^k}(x^k, z^k),
\]

with \(\approx_{\eta^k}\) corresponding to

\[
\| (x^{k+1}, z^{k+1}) - P_{\mu^k}(x^k, z^k) \|_2 \leq \eta^k,
\]

is guaranteed to converge globally to some \((x^*, z^*) \in T_\mathcal{E}^{-1}(0, 0)\) \([3, Proposition 1.2]\).

Depending on the value of \(k_{\text{max}}\), we distinguish two different cases to show that PROXQP is asymptotically equivalent to PPA. As we will see, it amounts to show PROXQP always enter the “if” condition in a finite number of iterations.

- If \(k_{\text{max}} < +\infty\), then after \(k_{\text{max}}\) number of iterations, PROXQP iterates always enter the “if” condition.
- Otherwise, under Assumption 1 and Assumption 2, Lemma \([9]\) ensures that there exists some \(N_{\text{max}} \in \mathbb{N}\), after which PROXQP enter only the “if” condition.

Hence, for \(k \geq N \overset{\text{def}}{=} \min(k_{\text{max}}, N_{\text{max}})\), PROXQP always accepts the iterates \((x^{k+1}, z^{k+1})\) from (10). Furthermore, for \(l \geq 0\)

\[
\mu^{N+l} = \mu^N,
\]

\[
\epsilon^{N+l} = \epsilon^k(\mu^N)^l.
\]

Hence the iterates of Algorithm 2 amount to a fixed step PPA, namely for \(k \geq N\)

\[
(x^{k+1}, z^{k+1}) \approx_{\epsilon^k} P_{\mu^N}(x^k, z^k),
\]

for some constant \(\epsilon > 0\) which can be provided following the same strategy as detailed in \([19]\). As \(\{\epsilon^k\}_k\) is summable for \(k \geq N\), it concludes our claim. \(\square\)

### F. Proof of Lemma 1

**Lemma 4.** Under Assumption 1 (existence of solution to (QP)) and Assumption 2 (bounded iterates), and with all parameters set as required in the inputs of Algorithm 2, \(\exists K \in \mathbb{N}\), and \(\mu \in \mathbb{R}\) such that \(\forall k \geq K, \mu^k = \mu\). Hence, \(\forall k \geq 0, p^{K+k} \leq \epsilon_{\text{cl}}(\mu^K)^{\beta_{\text{cl}}k}\).

**Proof.** By Lemma \([3]\) \(\exists \mu \in \mathbb{R}, \exists K \in \mathbb{N}\) such that for all \(k \geq K\), it holds that \(p^k \leq \epsilon_{\text{cl}}\) and \(\mu^k = \mu\) is constant. \(\square\)

### References


