Complexity of a quadratic penalty accelerated inexact proximal point method

W. Kong\(^1\)  J.G. Melo\(^2\)  R.D.C. Monteiro\(^1\)

\(^1\)School of Industrial and Systems Engineering  
Georgia Institute of Technology

\(^2\)Institute of Mathematics and Statistics  
Federal University of Goias

ICERM 2019 - April 30th, Providence
1. The Main Problem

2. The Penalty Approach

3. AIPP Method For Solving the Penalty Subproblem(s)
   - Special Structure of Penalty Subproblem
   - Previous Works
   - AIPP = Inexact Proximal Point + Acceleration
   - AIPP Method and its Complexity

4. Complexity of the Penalty AIPP

5. Computational Results

6. Additional Results and Concluding Remarks
The main problem:

\[(P) \quad \phi^* := \min \{ \phi(z) := f(z) + h(z) : Az = b, \ z \in \mathbb{R}^n \}\]

where

- \( A : \mathbb{R}^n \rightarrow \mathbb{R}^l \) is linear and \( b \in \mathbb{R}^l \)
- \( h : \mathbb{R}^n \rightarrow (-\infty, \infty] \) closed proper convex with bounded domain;
- \( f \) is differentiable (not necessarily convex) on \( \text{dom} \ h \) and, for some \( L_f > 0 \),

\[ \| \nabla f(z) - \nabla f(z') \| \leq L_f \| z - z' \|, \quad \forall z, z' \in \text{dom} \ h \]
The main problem (continued):

\[(P) \quad \phi^* := \min \{ \phi(z) := f(z) + h(z) : Az = b, \ z \in \mathbb{R}^n \}\]

Our goal: Given \((\bar{\rho}, \bar{\eta}) > 0\), find a \((\bar{\rho}, \bar{\eta})\)-approximate solution of \((P)\), i.e., a triple \((\bar{z}, \bar{\omega}; \bar{v})\) such that

\[
\bar{v} \in \nabla f(\bar{z}) + \partial h(\bar{z}) + A^* \bar{\omega}, \quad \|\bar{v}\| \leq \bar{\rho}, \quad ||A\bar{z} - b|| \leq \bar{\eta}
\]

It will be achieved via a penalty approach.
For $c > 0$, consider

$$(P_c) \quad \phi^*_c := \min_{z} \phi_c(z) := f_c(z) + h(z)$$

where

$$f_c(z) := f(z) + \frac{c}{2} \|Az - b\|^2$$

**Quadratic Penalty Approach:**

0. choose initial $c > 0$

1. obtain a $\bar{\rho}$-approximate solution $(\bar{z}; \bar{v})$ of $(P_c)$, i.e., satisfying

$$\bar{v} \in \nabla f_c(\bar{z}) + \partial h(\bar{z}), \quad \|\bar{v}\| \leq \bar{\rho}$$

2. if $\|A\bar{z} - b\| \leq \bar{\eta}$ then stop and output $\bar{z}$; otherwise, set $c \leftarrow 2c$ and go to step 1
For \( c > 0 \), consider

\[
(P_c) \quad \phi_c^* := \min_z \phi_c(z) := f_c(z) + h(z)
\]

where

\[
f_c(z) := f(z) + \frac{c}{2} \|Az - b\|^2
\]

**Quadratic Penalty Approach:**

0. choose initial \( c > 0 \)

1. obtain a \( \bar{\rho} \)-approximate solution \((\bar{z}; \bar{v})\) of \((P_c)\), i.e., satisfying

\[
\bar{v} \in \nabla f_c(\bar{z}) + \partial h(\bar{z}), \quad \|\bar{v}\| \leq \bar{\rho}
\]

2. if \( \|A\bar{z} - b\| \leq \bar{\eta} \) then stop and output \( \bar{z} \); otherwise, set \( c \leftarrow 2c \) and go to step 1
Theorem

Let \((\bar{\rho}, \bar{\eta}) > 0\) be given. Assume that \((\bar{z}; \bar{v})\) is a \(\bar{\rho}\)-approximate solution of \((P_c)\) and define

\[
\bar{w} := c(A\bar{z} - b), \quad R := 2\Delta^*_\phi + 2\bar{\rho}D_h + L_f D_h^2
\]

where

\[
D_h := \sup\{\|z - z'\| : z, z' \in \text{dom } h\},
\]
\[
\Delta^*_\phi := \phi^* - \phi_*, \quad \phi_* := \inf_z\{(f + h)(z) : z \in \mathbb{R}^n\}
\]

Then, \((\bar{z}, \bar{w}; \bar{v})\) is \((\bar{\rho}, \bar{\eta})\)-approximate solution of \((P)\) whenever

\[
c \geq \frac{R}{\bar{\eta}^2}
\]
1. The Main Problem

2. The Penalty Approach

3. AIPP Method For Solving the Penalty Subproblem(s)
   - Special Structure of Penalty Subproblem
   - Previous Works
   - AIPP = Inexact Proximal Point + Acceleration
   - AIPP Method and its Complexity

4. Complexity of the Penalty AIPP

5. Computational Results

6. Additional Results and Concluding Remarks
Recall that the objective function of \((P_c)\) is \(\phi_c = f_c + h\) where
\[
f_c(z) := f(z) + c\|Az - b\|^2 / 2
\]
For every \(z, z' \in \text{dom} \, h\),
\[
-m \leq \frac{f_c(z') - [f_c(z) + \langle \nabla f_c(z), z' - z \rangle]}{\|z' - z\|^2 / 2} \leq M_c
\]
where
\[
m := L_f, \quad M_c := L_f + c\|A\|^2
\]
The complexity of the composite gradient method for solving \((P_c)\) is
\[
\mathcal{O} \left( \frac{M_c}{\rho^2} \frac{mD^2_h}{h^2} \right)
\]
which is high for large \(c\), or when \(M_c >> m\).
Recall that the objective function of \((P_c)\) is \(\phi_c = f_c + h\) where

\[
f_c(z) := f(z) + c\|Az - b\|^2 / 2
\]

For every \(z, z' \in \text{dom } h\),

\[
-m \leq \frac{f_c(z') - \left[f_c(z) + \langle \nabla f_c(z), z' - z \rangle \right]}{\|z' - z\|^2 / 2} \leq M_c
\]

where

\[
m := L_f, \quad M_c := L_f + c\|A\|^2
\]

The complexity of the composite gradient method for solving \((P_c)\) is

\[
O \left( M_c \frac{mD_h^2}{\bar{\rho}^2} \right)
\]

which is high for large \(c\), or when \(M_c >> m\).
1. The Main Problem

2. The Penalty Approach

3. AIPP Method For Solving the Penalty Subproblem(s)
   - Special Structure of Penalty Subproblem
   - Previous Works
     - AIPP = Inexact Proximal Point + Acceleration
     - AIPP Method and its Complexity

4. Complexity of the Penalty AIPP

5. Computational Results

6. Additional Results and Concluding Remarks

**Complexity:**

\[ O \left( \frac{M_c m D_h^2}{\bar{\rho}^2} + \left( \frac{M_c d_0}{\bar{\rho}} \right)^{2/3} \right) \]

The dominant term (i.e., the blue one) is \( O(M_c) \).


obtained a \( O(\sqrt{M_c} \log M_c) \) complexity bound under the assumption that \( h = 0 \).

Our AIPP approach removes the \( \log M_c \) from the above bound and the assumption that \( h = 0 \).
1. The Main Problem

2. The Penalty Approach

3. **AIPP Method For Solving the Penalty Subproblem(s)**
   - Special Structure of Penalty Subproblem
   - Previous Works
   - **AIPP = Inexact Proximal Point + Acceleration**
   - AIPP Method and its Complexity

4. Complexity of the Penalty AIPP

5. Computational Results

6. Additional Results and Concluding Remarks

**AIPP = Inexact Proximal Point + Acceleration**
AIPP for solving \((P_c)\) is based on an IPP scheme whose \(k\)-th iteration is as follows. Given \(z_{k-1}\), it chooses \(\lambda_k > 0\) and approximately solves the ‘prox’ subproblem

\[
(P^k_c) \quad \text{min} \left\{ \lambda_k (f_c + h)(z) + \frac{1}{2} \| z - z_{k-1} \|^2 \right\}
\]

i.e., for some \(\sigma \in (0, 1)\), it computes a point \(z_k\) and a residual pair \((v_k, \epsilon_k) \in \mathbb{R}^n \times \mathbb{R}_+\) such that

\[
v_k \in \partial \epsilon_k \left( \lambda_k (f_c + h) + \frac{1}{2} \| \cdot - z_{k-1} \|^2 \right)(z_k)
\]

\[\| v_k \|^2 + 2\epsilon_k \leq \sigma \| z_{k-1} - z_k + v_k \|^2\]
AIPP = Inexact Proximal Point + Acceleration

AIPP for solving \((P_c)\) is based on an IPP scheme whose \(k\)-th iteration is as follows. Given \(z_{k-1}\), it chooses \(\lambda_k > 0\) and approximately solves the ‘prox’ subproblem

\[
(P^k_c) \quad \min \left\{ \lambda_k (f_c + h)(z) + \frac{1}{2} \| z - z_{k-1} \|^2 \right\}
\]

i.e., for some \(\sigma \in (0, 1)\), it computes a point \(z_k\) and a residual pair \((v_k, \varepsilon_k) \in \mathbb{R}^n \times \mathbb{R}_+\) such that

\[
v_k \in \partial \varepsilon_k \left( \lambda_k (f_c + h) + \frac{1}{2} \| \cdot - z_{k-1} \|^2 \right)(z_k)
\]

\[
\| v_k \|^2 + 2\varepsilon_k \leq \sigma \| z_{k-1} - z_k + v_k \|^2
\]
**AIPP method:** It is an accelerated instance of the above IPP scheme in which for all $k$:

- $\lambda_k = 1/(2m)$, and hence $(P^k_c)$ is a strongly convex problem
- $z_k$ and $(v_k, \varepsilon_k)$ are computed by an accelerated composite gradient (ACG) method applied to $(P^k_c)$ in at most $\mathcal{O}\left(\sqrt{\frac{M_c m}{m}}\right)$ iterations

**Obs:** Each ACG iteration requires one or two evaluations of the resolvent of $h$, i.e., exact solution of

$$\min\{a^T z + h(z) + \theta\|z\|^2\}$$
1. The Main Problem

2. The Penalty Approach

3. AIPP Method For Solving the Penalty Subproblem(s)
   - Special Structure of Penalty Subproblem
   - Previous Works
   - AIPP = Inexact Proximal Point + Acceleration
   - AIPP Method and its Complexity

4. Complexity of the Penalty AIPP

5. Computational Results

6. Additional Results and Concluding Remarks
(0) (beginning of phase I) Let \( c > 0, z_0 \in \text{dom } h, \sigma \in (0, 1) \) and \( \bar{\rho} > 0 \) be given, and set \( \lambda = 1/(2m) \) and \( k = 1 \)

(1) call an ACG variant started from \( z_{k-1} \) to approximately solve \((P^k_c)\), i.e., to obtain \( z_k \) and \((v_k, \epsilon_k)\) such that

\[
v_k \in \partial \epsilon_k \left( \lambda(f_c + h) + \frac{1}{2} \| \cdot - z_{k-1} \|^2 \right)(z_k)
\]

\[
\| v_k \|^2 + 2\epsilon_k \leq \sigma \| z_{k-1} - z_k + v_k \|^2
\]

(2) if \( \| z_{k-1} - z_k + v_k \| > \lambda \bar{\rho}/10 \), then \( k \leftarrow k + 1 \) and go to (1); otherwise, go to (3) (end of phase I)

(3) (phase II) restart the last call to the ACG variant in step 1 to find \( \tilde{z} \) and \((\tilde{v}, \tilde{\epsilon})\) satisfying

\[
\| z_{k-1} - \tilde{z} + \tilde{v} \| \leq \frac{\lambda \bar{\rho}}{2}, \quad \tilde{\epsilon} \leq \lambda \frac{\bar{\rho}^2}{32(M_c + 2m)}
\]

and then refine \((\tilde{z}; \tilde{v}, \tilde{\epsilon})\) to obtain a \( \bar{\rho} \)-approximate solution \((\tilde{z}; \tilde{v})\) for \((P_c)\).
(0) (beginning of phase I) Let $c > 0$, $z_0 \in \text{dom } h$, $\sigma \in (0, 1)$ and $\bar{\rho} > 0$ be given, and set $\lambda = 1/(2m)$ and $k = 1$

(1) call an ACG variant started from $z_{k-1}$ to approximately solve $(P^k_c)$, i.e., to obtain $z_k$ and $(v_k, \varepsilon_k)$ such that

$$v_k \in \partial \varepsilon_k \left( \lambda (f_c + h) + \frac{1}{2} \| \cdot - z_{k-1} \|^2 \right) (z_k)$$

$$\| v_k \|^2 + 2\varepsilon_k \leq \sigma \| z_{k-1} - z_k + v_k \|^2$$

(2) if $\| z_{k-1} - z_k + v_k \| > \lambda \bar{\rho}/10$, then $k \leftarrow k + 1$ and go to (1); otherwise, go to (3) (end of phase I)

(3) (phase II) restart the last call to the ACG variant in step 1 to find $\tilde{z}$ and $(\tilde{v}, \tilde{\varepsilon})$ satisfying

$$\| z_{k-1} - \tilde{z} + \tilde{v} \| \leq \frac{\lambda \bar{\rho}}{2}, \quad \tilde{\varepsilon} \leq \lambda \frac{\bar{\rho}^2}{32(M_c + 2m)}$$

and then refine $(\tilde{z}; \tilde{v}, \tilde{\varepsilon})$ to obtain a $\bar{\rho}$-approximate solution $(\tilde{z}; \tilde{v})$ for $(P_c)$. 
The total number of ACG iterations is

$$
O \left( \frac{\sqrt{M_c m}}{\bar{\rho}^2} \min \{ \Delta_0^*(c), mD_h^2 \} + \sqrt{\frac{M_c}{m}} \log \left( \max \left\{ 1, \frac{M_c}{m \sqrt{m}} \right\} \right) \right)
$$

where $D_h$ is the diameter of $\text{dom} \ h$ and $\Delta_0^*(c) = \phi_c(z_0) - \phi_c^*$

Hence, the complexity of the AIPP method is

$$
O \left( \frac{\sqrt{M_c m} mD_h^2}{\bar{\rho}^2} \right)
$$

while that of the CG or Ghadimi-Lan’s AG is

$$
O \left( M_c \frac{mD_h^2}{\bar{\rho}^2} \right)
$$
Theorem

The total number of ACG iterations is

\[ O \left( \frac{\sqrt{M_c m}}{\varrho^2} \min \{ \Delta_0^*(c), mD_h^2 \} + \sqrt{\frac{M_c}{m}} \log \left( \max \left\{ 1, \frac{M_c}{m\sqrt{m}} \right\} \right) \right) \]

where \( D_h \) is the diameter of \( \text{dom} \ h \) and \( \Delta_0^*(c) = \phi_c(z_0) - \phi_c^* \)

Hence, the complexity of the AIPP method is

\[ O \left( \frac{\sqrt{M_c m}}{\varrho^2} \frac{mD_h^2}{\varrho^2} \right) \]

while that of the CG or Ghadimi-Lan’s AG is

\[ O \left( M_c \frac{mD_h^2}{\varrho^2} \right) \]
Complexity of the quadratic penalty AIPP: Recall that a sufficient condition for attaining \( \|A\tilde{z} - b\| \leq \tilde{\eta} \) is that \( c \geq R/\tilde{\eta}^2 \) where

\[
R := 2\Delta^*_\phi + 2\bar{\rho}D_h + L_f D_h^2
\]

Theorem

The quadratic penalty AIPP method performs a total of at most

\[
\mathcal{O}\left(\frac{\sqrt{R\|A\|L_f^2D_h^2}}{\bar{\rho}^2\tilde{\eta}} + \frac{L_f^2D_h^2}{\bar{\rho}^2}\right)
\]

ACG iterations to find a \((\bar{\rho}, \tilde{\eta})\)-approximate solution of \((P)\).

Hence, the complexity of the penalty AIPP is \(\mathcal{O}\left(1/(\bar{\rho}^2\tilde{\eta})\right)\).
Complexity of the quadratic penalty AIPP: Recall that a sufficient condition for attaining \( \|Az - b\| \leq \bar{\eta} \) is that \( c \geq R / (\bar{\eta})^2 \) where

\[ R := 2\Delta^*_\phi + 2\bar{\rho}D_h + L_f D_h^2 \]

**Theorem**

The quadratic penalty AIPP method performs a total of at most

\[ O \left( \frac{\sqrt{R\|A\|L_f^3/2} D_h^2}{\bar{\rho}^2 \bar{\eta}} + \frac{L_f^2 D_h^2}{\bar{\rho}^2} \right) \]

ACG iterations to find a \((\bar{\rho}, \bar{\eta})\)-approximate solution of \((P)\).

Hence, the complexity of the penalty AIPP is \( O \left( 1 / (\bar{\rho}^2 \bar{\eta}) \right) \)
Computational Results

- AIPP was benchmarked against Ghadimi-Lan’s AG method.
- The nonconvex optimization problem tested was

$$\min_{z \in S^n_+} \left\{ f(z) := -\frac{\zeta}{2} \|DB(z)\|^2 + \frac{\tau}{2} \|A(z) - b\|^2 : z \in P_n \right\}$$

where $P_n$ is the unit spectraplex, i.e.,

$$P_n := \{ z \in S^n_+ : \text{tr}(z) = 1 \}$$

$A : S^n \to \mathbb{R}^n$, $B : S^n \to \mathbb{R}^l$ are linear operators, $D$ is a positive diagonal matrix, $b \in \mathbb{R}^n$

- Values in $A$, $B$ and $b$ were sampled from the $U[0, 1]$ distribution at sparsity level $d$ and values for $D$ were sampled from $U[0, 1000]$ distribution.
## Results for composite unconstrained problems

\((l = 50, n = 200, d = 0.025, \bar{\rho} = 10^{-7})\)

<table>
<thead>
<tr>
<th>Size</th>
<th>Iteration Count</th>
<th>Runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AG</td>
<td>AIPP</td>
</tr>
<tr>
<td>1000000</td>
<td>1</td>
<td>3.84E+01</td>
</tr>
<tr>
<td>100000</td>
<td>1</td>
<td>3.82E+00</td>
</tr>
<tr>
<td>10000</td>
<td>1</td>
<td>3.67E-01</td>
</tr>
<tr>
<td>1000</td>
<td>1</td>
<td>2.05E-02</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>-1.74E-02</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>-3.65E-02</td>
</tr>
</tbody>
</table>
The Main Problem
Penalty Problem and Approach
AIPP Method For Solving the Penalty Subproblem(s)
Complexity of the Penalty

Results for composite unconstrained problems
\((l = 50, n = 1000, d = 0.001, \bar{\rho} = 10^{-7})\)

<table>
<thead>
<tr>
<th>Size</th>
<th>(\bar{f})</th>
<th>Iteration Count</th>
<th>Runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>AIPP</td>
</tr>
<tr>
<td>(M)</td>
<td>(m)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000000</td>
<td>1</td>
<td>2.98E+03</td>
<td>2351</td>
</tr>
<tr>
<td>100000</td>
<td>1</td>
<td>2.98E+02</td>
<td>2351</td>
</tr>
<tr>
<td>10000</td>
<td>1</td>
<td>2.97E+01</td>
<td>2347</td>
</tr>
<tr>
<td>1000</td>
<td>1</td>
<td>2.91E+00</td>
<td>2312</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>2.28E-01</td>
<td>1969</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>-6.80E-02</td>
<td>603</td>
</tr>
</tbody>
</table>
QP-AIPP was benchmarked against a penalty version of G-L’s AG method.

The linearly constrained nonconvex optimization problem tested was

\[
\min_{z \in S^n_+} \left\{ f(z) = -\frac{\xi}{2} \|DB(z)\|^2 : z \in P_n, \ A(z) = b \right\}
\]

where \( A : S^n \to \mathbb{R}^n \), \( B : S^n \to \mathbb{R}^l \) and \( D \) were generated as before.

\( b \) was chosen so as to make \( l/n \) feasible.
Results for composite linearly constrained problems

\((l = 50, n = 20, \, d = 1, \, \bar{\rho} = 10^{-3}, \, \bar{\eta} = 10^{-6})\)

<table>
<thead>
<tr>
<th>(L_f)</th>
<th>(\bar{F})</th>
<th>Iteration Count</th>
<th>Runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>AIPP</td>
</tr>
<tr>
<td>1000000</td>
<td>-1.49E+03</td>
<td>110415</td>
<td>17673</td>
</tr>
<tr>
<td>100000</td>
<td>-1.49E+02</td>
<td>110414</td>
<td>17673</td>
</tr>
<tr>
<td>10000</td>
<td>-1.49E+01</td>
<td>110386</td>
<td>17673</td>
</tr>
<tr>
<td>1000</td>
<td>-1.49E+00</td>
<td>110135</td>
<td>17673</td>
</tr>
<tr>
<td>100</td>
<td>-1.49E-01</td>
<td>107942</td>
<td>17393</td>
</tr>
<tr>
<td>10</td>
<td>-1.49E-02</td>
<td>96776</td>
<td>16499</td>
</tr>
</tbody>
</table>
### Results for composite linearly constrained problems

\( (l = 50, n = 100, d = 0.0015, \bar{\rho} = 10^{-3}, \bar{\eta} = 10^{-6}) \)

<table>
<thead>
<tr>
<th>( L_f )</th>
<th>( \bar{f} )</th>
<th>Iteration Count</th>
<th>Runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>AIPP</td>
</tr>
<tr>
<td>1000000</td>
<td>-5.22E+04</td>
<td>33330</td>
<td>6426</td>
</tr>
<tr>
<td>100000</td>
<td>-5.22E+03</td>
<td>33290</td>
<td>5405</td>
</tr>
<tr>
<td>10000</td>
<td>-5.22E+02</td>
<td>32897</td>
<td>3897</td>
</tr>
<tr>
<td>1000</td>
<td>-5.22E+01</td>
<td>29611</td>
<td>8321</td>
</tr>
<tr>
<td>100</td>
<td>-5.22E+00</td>
<td>17289</td>
<td>7042</td>
</tr>
<tr>
<td>10</td>
<td>-5.22E-01</td>
<td>5917</td>
<td>4644</td>
</tr>
</tbody>
</table>
Implementation Remarks

- Even though Phase II is theoretically needed, it was never needed for solving the instances in our test.

- $\lambda_k$ has been chosen aggressively in all instances, i.e., $\lambda_k > 1/m$. 
Additional results

\[ p_* := \min_x \{ f(x) + h(x) : Ax = b \} \]

where now

\[ f(x) = \max_{y \in Y} \Phi(x, y) \]

Assume that \( Y \) is a closed convex set whose diameter

\[ D_y := \sup_{y, y' \in Y} \| y - y' \| \]

is finite
It is also assumed that

- $\Phi(x, \cdot)$ is concave on $Y$ for every $x \in X$;
- $\Phi(\cdot, y)$ is continuously differentiable on $\text{dom } h$ for every $y \in Y$;
- there exist scalars $(L_x, L_y) \in \mathbb{R}^2_{++}$, and $m \in (0, L_x]$ such that

$$
\Phi(x', y) - \left[ \Phi(x, y) + \left< \nabla_x \Phi(x, y), x' - x \right> \right] \geq - \frac{m}{2} \|x - x'\|_X^2
$$

$$
\| \nabla_x \Phi(x, y) - \nabla_x \Phi(x', y') \|_X \leq L_x \|x - x'\|_X + L_y \|y - y'\|_Y
$$

for every $x, x' \in \text{dom } h$ and $y, y' \in Y$. 
$f$ can now be nonsmooth and nonconvex but it can easily be approximated by a smooth nonconvex function, namely,

$$f_\xi(x) := \max_{y \in Y} \left\{ \Phi_\xi(x, y) := \Phi(x, y) - \frac{1}{2\xi} \| y - y_0 \|^2_Y : y \in Y \right\}$$

where $y_0 \in Y$ and $\xi > 0$

Similar to the one used by Nesterov in his smooth approximation acceleration scheme!
Applying the penalty AIPP method to

$$\min_x \{ f_\xi(x) + h(x) : Ax = b \}$$

for some well-chosen $\xi$, yields a quintuple $(\bar{u}, \bar{v}, \bar{x}, \bar{y}, \bar{w})$ satisfying

$$\left( \begin{array}{c} \bar{u} \\ \bar{v} \end{array} \right) \in \left( \begin{array}{c} \nabla_x \Phi(\bar{x}, \bar{y}) + A^* \bar{w} \\ 0 \end{array} \right) + \left( \begin{array}{c} \partial h(\bar{x}) \\ [-\Phi(\bar{x}, \cdot)](\bar{y}) \end{array} \right)$$

$$\|\bar{u}\|_{\lambda^*} \leq \rho_x, \quad \|\bar{v}\|_{\gamma^*} \leq \rho_y, \quad \|A\bar{x} - b\|_U \leq \eta.$$

in a total number of ACG iterations bounded by

$$\mathcal{O} \left( m^{3/2} D_h^2 \left[ \frac{L_x^{1/2}}{\rho_x^2} + \frac{L_y D_y^{1/2}}{\rho_y^{1/2} \rho_x^2} + \frac{m^{1/2} \|A\| D_h}{\eta \rho_x^2} \right] \right)$$

The complexity is still $\mathcal{O}(1/\eta^3)$ under the assumption that $\rho_x = \rho_y = \eta.$
Concluding Remarks

- We have presented the quadratic penalty AIPP method for "solving" a linearly constrained composite smooth nonconvex program and have shown that its associated bound is

\[ O \left( \frac{1}{\bar{\rho}^2 \bar{\eta}} \right) \]

If instead either the PG or AG method were used to solve subproblems \( (P_c) \), the bound would be \( O \left( 1/ [\bar{\rho}^2 \bar{\eta}^2] \right) \).

- We have also argued that the above complexity ‘remains the same’ in the context of linearly constrained composite nonsmooth nonconvex min-max programs.
THE END

Thanks!
The Main Problem
Penalty Problem and Approach
AIPP Method For Solving the Penalty Subproblem(s)
Complexity of the Penalty AIPP
Additional Results and Concluding Remarks

Example
On first slide.

Example
On second slide.
Example
On first slide.

Example
On second slide.
Theorem

On first slide.

Corollary

On second slide.
The Main Problem

Penalty Problem and Approach

AIPP Method For Solving the Penalty Subproblem(s)

Complexity of the Penalty

Additional Results and Concluding Remarks

Theorem

On first slide.

Corollary

On second slide.
Theorem

In left column.

Corollary

In right column.

New line
Theorem

In left column.

Corollary

In right column.
New line
You can control text size using special keywords

Text Text Text Text Text Text Text Text Text Text

You can also specify the text size directly
This sentence has 0.5 centimeters of space between lines.
This sentence is 1x the size of normal sentences
This sentence is 2x the size of normal sentences
- You can control spacing between bullet points with the `vspace*` command
- This bullet point will have addition vertical spacing after it
- This bullet point will have less vertical spacing after it
- This is the last item
The first main message of your talk in one or two lines.
The second main message of your talk in one or two lines.
Perhaps a third message, but not more than that.

Outlook
- What we have not done yet.
- Even more stuff.
A. Author.  
*Handbook of Everything.*  

S. Someone.  
On this and that.  
*Journal on This and That.* 2(1):50–100, 2000.