On projecting in primal-dual splitting methods for solving constrained convex optimization

L. M. Briceño-Arias
Universidad Técnica Federico Santa María

Flatiron Institute, NYC, USA
March 21st, 2019
1 Problem and motivation

2 Algorithm, convergence, and numerical simulations

3 Application

4 Open questions
Convex non-differentiable optimization problem

Problem (P)

\[
\min_{x \in \mathbb{R}^N} f(x) + g(x) + h(Lx).
\]

- \( f : \mathbb{R}^N \to \mathbb{R} \) is differentiable and \( \nabla f \) is \( \beta^{-1} \)-Lipschitz.
- \( g \in \Gamma_0(\mathbb{R}^N) \) and \( h \in \Gamma_0(\mathbb{R}^M) \).
- \( L \) is a \( M \times N \) real matrix.
- The set of solutions is nonempty.
- \( \text{ri}(\text{dom } h) \cap L(\text{ri}(\text{dom } g)) \neq \emptyset \).

Important case: \( h = \nu_b \) for some \( b \in \mathbb{R}^M \)

\[
\min_{Lx = b} f(x) + g(x).
\]
Motivation: Mean Field Games

- Mean Field Games were introduced by Lasry-Lions (2006) and models a game in which the number of players (indistinguishable) tend to infinity.

- The limiting behavior collapses into two coupled PDE’s, which are the optimality condition of the optimization problem

\[
\inf_{(m,w)} \int_{T^d} \left[ b_2(m(x), w(x)) + F(x, m(x)) \right] dx,
\]

subject to
\[
\begin{cases}
-\nu \Delta m + \text{div}(w) = 0 \text{ in } T^d, \\
\int_{T^d} m(x) dx = 1.
\end{cases}
\]

where \( m \) represents the probability distribution on the actions of players and \( w \) involves the gradient of the cost function of a “small player”. \( F \) models the influence of other players on the cost function of each one.
Motivation: Mean Field Games

- The function

\[ b_2(m, w) := \begin{cases} \frac{|w|^2}{2m}, & \text{if } m > 0, \\ 0, & \text{if } (m, w) = (0, 0), \\ +\infty, & \text{otherwise} \end{cases} \]

is convex but not differentiable.

- It is possible to compute the proximity operator of

\[ \varphi: (m, w) \mapsto \int_{T_d} \left[ b_2(m(x), w(x)) + F(x, m(x)) \right] dx. \]

It involves a cubic real equation for each variable (after discretization).

- The constraints are affine linear.

- After discretization, the matrix involved in the constraints is usually bad conditioned.
Classic approach: ADMM

Problem (P)

\[
\min_{x \in \mathbb{R}^N} f(x) + g(x) + h(Lx).
\]

An equivalent formulation is

\[
\min_{Lx = y} f(x) + g(x) + h(y)
\]

from which we define the Augmented Lagrangian \((\gamma > 0)\):

\[
\mathcal{L}_\gamma(x, y, u) = f(x) + g(x) + h(y) + u \cdot (Lx - y) + \frac{\gamma}{2} \|Lx - y\|^2.
\]

- Under qualification conditions \(x\) solves \((P)\) iff \((x, Lx, u)\) is a saddle point of \(\mathcal{L}_\gamma\).
- From an alternating minimization-maximization method we obtain the classical Alternating Direction method of Multipliers (ADMM) (Gabay-Mercier 80’s):
Classic approach: ADMM

\[ x^{k+1} = \arg\min_x \mathcal{L}_{\gamma}(x, y^k, u^k) \]
\[ y^{k+1} = \arg\min_y \mathcal{L}_{\gamma}(x^{k+1}, y, u^k) \]
\[ u^{k+1} = u^k + \gamma(Lx^{k+1} - y^{k+1}). \]
Classic approach: ADMM

\[ x^{k+1} = \arg\min_x \left\{ f(x) + g(x) + u^k \cdot Lx + \frac{\gamma}{2} \|Lx - y^k\|^2 \right\} \]

\[ y^{k+1} = \arg\min_y \left\{ h(y) - u^k \cdot y + \frac{\gamma}{2} \|Lx^{k+1} - y\|^2 \right\} \]

\[ u^{k+1} = u^k + \gamma(Lx^{k+1} - y^{k+1}). \]
Classic approach: ADMM

\[ x^{k+1} = \text{argmin}_x \left\{ f(x) + g(x) + u^k \cdot Lx + \frac{\gamma}{2} \|Lx - y^k\|^2 \right\} \]

\[ y^{k+1} = \text{argmin}_y \left\{ h(y) - u^k \cdot y + \frac{\gamma}{2} \|Lx^{k+1} - y\|^2 \right\} \]

\[ u^{k+1} = u^k + \gamma (Lx^{k+1} - y^{k+1}) . \]

- In the case when \( h = \iota_{\{b\}} \) for some \( b \in \mathbb{R}^M \), we have

\[ x^{k+1} = \text{argmin}_x \left\{ f(x) + g(x) + u^k \cdot Lx + \frac{\gamma}{2} \|Lx - b\|^2 \right\} \]

\[ u^{k+1} = u^k + \gamma (Lx^{k+1} - b) . \]
Drawbacks ADMM

- The primal iterates \((x^k)_{k \in \mathbb{N}}\) do not satisfy the constraints \((Lx^k \neq b)\).
- Moreover, the first step it is not easy in general (involves \(L\) and \(f + g\)).
- It can be solved efficiently only in specific instances: \(f + g\) quadratic, \(L^\top L = \alpha \text{Id.}\).
- **Idea:** Try to split the influence of \(L, f\) and \(g\) in the first step.
Drawbacks ADMM

- The primal iterates \((x^k)_{k \in \mathbb{N}}\) do not satisfy the constraints \((Lx^k \neq b)\).
- Moreover, the first step it is not easy in general (involves \(L\) and \(f + g\)).
- It can be solved efficiently only in specific instances: \(f + g\) quadratic, \(L^\top L = \alpha \text{Id}\).
- **Idea:** Try to split the influence of \(L, f\) and \(g\) in the first step.

Given \(x \in \mathbb{R}^N\), \(\text{prox}_f x\) is the unique solution to

\[
\minimize_{y \in \mathbb{R}^N} f(y) + \frac{1}{2} \|x - y\|^2.
\]

- Several functions \(f\) have an explicit or efficiently computable \(\text{prox}_f\). **Examples:** \(\| \cdot \|_1\), \(\iota_C\) (\(\text{prox}_{\iota_C} = \text{P}_C\)), \(d_C\), etc...
Other approaches

Problem (P)

\[ \min_{x \in \mathbb{R}^N} f(x) + g(x) + h(Lx). \]

Combettes-Pesquet (2012)

Let \( 0 < \gamma < (\|L\| + \beta)^{-1} \), \( x^0 \in \mathbb{R}^N \) and \( u^0 \in \mathbb{R}^M \) and iterate

\[
\begin{align*}
  p^k_1 &= \text{prox}_{\gamma g}(x^k - \gamma(\nabla f(x^k) + L^T u^k)) \\
  p^k_2 &= \text{prox}_{\gamma h^*}(u^k + \gamma Lx^k) \\
  x^{k+1} &= p^k_1 - \gamma(L^T p^k_2 + \nabla f(p^k_1) - L^T u^k - \nabla f(x^k)) \\
  u^{k+1} &= p^k_2 + \gamma(Lp^k_1 - Lx^k).
\end{align*}
\]

In the case \( f = 0 \), is the method proposed in BA-Combettes (2011).
Other approaches

Problem (P) case $h = \lambda \{ b \}$

$$\min_{Lx=b} f(x) + g(x).$$

Combettes-Pesquet (2012)

Let $0 < \gamma < (\|L\| + \beta)^{-1}$, $x^0 \in \mathbb{R}^N$ and $u^0 \in \mathbb{R}^M$ and iterate

$$
\begin{align*}
    p_1^k &= \text{prox}_{\gamma g} (x^k - \gamma(\nabla f(x^k) + L^\top u^k)) \\
    p_2^k &= u^k + \gamma(Lx^k - b) \\
    x^{k+1} &= p_1^k - \gamma(L^\top p_2^k + \nabla f(p_1^k) - L^\top u^k - \nabla f(x^k)) \\
    u^{k+1} &= p_2^k + \gamma(Lp_1^k - Lx^k).
\end{align*}
$$

- Also the influences of $L, f$ and $g$ have been split, but primal iterates do not satisfy the constraints.
- The method does not exploit cocoercivity of $\nabla f$. 
Problem and motivation

Convergence and Numerics

Application

Open questions

Other approaches

Problem (P)

\[
\min_{x \in \mathbb{R}^N} f(x) + g(x) + h(Lx).
\]

Condat-Vũ (2013)

\[
x^0 = \bar{x}^0 \in \mathbb{R}^N \text{ and } u^0 \in \mathbb{R}^M, \tau, \gamma > 0 \text{ such that } \tau \gamma \|L\|^2 < 1 - \frac{\tau}{2\beta}
\]

\[
\begin{align*}
    u^{k+1} &= \text{prox}_{\gamma h^*}(u^k + \gamma L\bar{x}^k) \\
x^{k+1} &= \text{prox}_{\tau g}(x^k - \tau(\nabla f(x^k) + L^T u^{k+1})) \\
    \bar{x}^{k+1} &= 2x^{k+1} - x^k.
\end{align*}
\]

- The case \( f = 0 \), the method was first proposed by Chambolle-Pock (2011).
Other approaches

Problem (P) case \( h = \iota \{ b \} \)

\[
\min_{Lx = b} f(x) + g(x).
\]

Condat-Vũ (2013)

\( x^0 = \bar{x}^0 \in \mathbb{R}^N \) and \( u^0 \in \mathbb{R}^M \), \( \tau, \gamma > 0 \) such that \( \tau \gamma \|L\|^2 < 1 - \frac{\tau}{2\beta} \)

\[
\begin{align*}
    u^{k+1} &= u^k + \gamma (L\bar{x}^{k+1} - b) \\
    x^{k+1} &= \text{prox}_{\tau g}(x^k - \tau (\nabla f(x^k) + L^\top u^{k+1})) \\
    \bar{x}^{k+1} &= 2x^{k+1} - x^k.
\end{align*}
\]

- \( \text{prox}_{\gamma h^*} = \text{Id} - \gamma \text{prox}_{h/\gamma} \circ (\text{Id}/\gamma) = \text{Id} - \gamma b. \)
- Now the influences of \( L, f \) and \( g \) have been split, but the primal iterates do not satisfy the constraints. Same with other Aug. Lagrangian approach proposed by Chen & Teboulle (1994).
If $P_{L^{-1}b}$ is computable...

- Suppose that it is possible to compute $P_{L^{-1}b} = \text{prox}_{h\circ L}$.
- If $LL^\top$ is invertible, $P_{L^{-1}b}x = x - L^\top(LL^\top)^{-1}(Lx - b)$.
- In this case we can avoid splitting $L$ from $g$ and use several methods for solving optimization problems without involving linear operators.
- In all these methods, primal iterates satisfy the constraints.
- But, in several cases $LL^\top$ is not invertible or it is bad conditioned. In those cases...
Goal of this talk

- Provide an algorithm which can ensure that primal iterates satisfy some of the constraints by adding projections onto the related affine space (a priori knowledge).
- Provide convergence results and some numerical experiences.
- Application to Mean Field Games.
Goal of this talk

- Provide an algorithm which can ensure that primal iterates satisfy some of the constraints by adding projections onto the related affine space (a priori knowledge).
- Provide convergence results and some numerical experiences.
- Application to Mean Field Games.
- Open questions.
1 Problem and motivation

2 Algorithm, convergence, and numerical simulations

3 Application

4 Open questions
Projected primal-dual splitting\textsuperscript{1}

Let $C$ be a nonempty closed convex set such that $P_C$ is easy to compute.

**Problem (P)**

$$\min_{x \in \mathbb{R}^N} f(x) + g(x) + h(Lx).$$

**Condat-Vũ (2013)**

$x^0 = \bar{x}^0 \in \mathbb{R}^N$ and $u^0 \in \mathbb{R}^M$, $\tau, \gamma > 0$ such that $\tau \gamma \|L\|^2 < 1 - \frac{\tau^2}{2\beta}$

\[
\begin{align*}
    u^{k+1} &= \text{prox}_{\gamma h^*}(u^k + \gamma L\bar{x}^k) \\
    p^{k+1} &= \text{prox}_{\tau g}(x^k - \tau(\nabla f(x^k) + L^T u^{k+1})) \\
    x^{k+1} &= p^{k+1} \\
    \bar{x}^{k+1} &= x^{k+1} + p^{k+1} - x^k.
\end{align*}
\]

Then, $x^k \rightarrow x$ solution to $(P)$

\textsuperscript{1}Joint work with Sergio López Rivera, USM
Projected primal-dual splitting

Let $C$ be a nonempty closed convex set such that $P_C$ is easy to compute.

Problem (P) with a priori information $C$

$$\begin{align*}
(Q) \quad \text{find } \hat{x} \in C \cap \arg\min_{x \in \mathbb{R}^N} f(x) + g(x) + h(Lx) \neq \emptyset.
\end{align*}$$

Projected primal-dual splitting

$$\begin{align*}
x^0 &= \bar{x}^0 \in \mathbb{R}^N \text{ and } u^0 \in \mathbb{R}^M, \tau, \gamma > 0 \text{ such that } \tau \gamma \|L\|^2 < 1 - \frac{\tau}{2\beta}
\end{align*}$$

$$\begin{align*}
u^{k+1} &= \text{prox}_{\gamma h^*}(u^k + \gamma L\bar{x}^k) \\
p^{k+1} &= \text{prox}_{\tau g}(x^k - \tau (\nabla f(x^k) + L^\top u^{k+1})) \\
x^{k+1} &= P_C p^{k+1} \\
\bar{x}^{k+1} &= x^{k+1} + p^{k+1} - x^k.
\end{align*}$$

Then, $(x^k)_{k \in \mathbb{N}} \subset C$ and $x^k \to \hat{x}$ solution to $(Q)$.

Joint work with Sergio López Rivera, USM
Projected primal-dual splitting: case $h = \nu\{b\}$

- In the case when $h = \nu\{b\}$, suppose that $M = r + s$, $L: x \mapsto (Rx, Sx)$, where $R$ and $S$ are $r \times N$ and $s \times N$ real matrices, and $b = (c, d)$, where $c \in \mathbb{R}^r$ and $d = \mathbb{R}^s$.
- Suppose that it is possible to project onto $C := R^{-1}c \subset L^{-1}b$.

**Problem (P)**

$$\min_{Lx=b} f(x) + g(x).$$

**Condat-Vũ (2013)**

$x^0 = \bar{x}^0 \in \mathbb{R}^N$ and $u^0 \in \mathbb{R}^M$, $\tau, \gamma > 0$ such that $\tau\gamma\|L\|^2 < 1 - \frac{\tau}{2\beta}$

\[
\begin{align*}
u^{k+1} &= u^k + \gamma(L\bar{x}^k - b) \\
x^{k+1} &= \text{prox}_{\tau g}(x^k - \tau(\nabla f(x^k) + L^\top u^{k+1})) \\
\bar{x}^{k+1} &= 2x^{k+1} - x^k.
\end{align*}
\]

Then, $x^k \to x$ solution to $(P)$
Projected primal-dual splitting: case $h = \nu\{b\}$

- In the case when $h = \nu\{b\}$, suppose that $M = r + s$, $L: x \mapsto (Rx, Sx)$, where $R$ and $S$ are $r \times N$ and $s \times N$ real matrices, and $b = (c, d)$, where $c \in \mathbb{R}^r$ and $d = \mathbb{R}^s$.
- Suppose that it is possible to project onto $C := R^{-1}c \subset L^{-1}b$.

Problem (P) with $C = R^{-1}c$

(Q) find $\hat{x} \in R^{-1}c \cap \arg\min_{Rx=c, Sx=d} f(x) + g(x)$.

Projected primal-dual splitting

$x^0 = \bar{x}^0 \in \mathbb{R}^N$ and $u^0 \in \mathbb{R}^M$, $\tau, \gamma > 0$ such that $\tau\gamma\|L\|^2 < 1 - \frac{\tau}{2\beta}$

\[
\begin{align*}
  u^{k+1} &= u^k + \gamma(L\bar{x}^k - b) \\
  p^{k+1} &= \text{prox}_{\tau g}(x^k - \tau(\nabla f(x^k) + L^\top u^{k+1})) \\
  x^{k+1} &= p^{k+1} - R^\top (RR^\top)^{-1}(Rp^{k+1} - c) \\
  \bar{x}^{k+1} &= x^{k+1} + p^{k+1} - x^k.
\end{align*}
\]

Then, $(x^k)_{k \in \mathbb{N}} \subset C$ and $x^k \to \hat{x}$ solution to (Q).
**Problem and motivation**

**Convergence and Numerics**

**Application**

**Open questions**

---

# Numerical experiences

\( \min_{\|x\|_1} \)

\( R_{x=c} \)

\( S_{x=d} \)

- Is the case \( f = 0, h = \nu \{(c,d)\} \) and \( g = \| \cdot \|_1 \).
- Dimensions: \( N = 1000, s = 100 \) and \( r \in \{1, 10, 30\} \).
- We consider 20 random realizations of matrices \( R, S \) and vectors \( c \) and \( d \) and we measure the relative error

\[
R_k = \sqrt{\frac{\|u^{k+1}-u^k\|^2 + \|x^{k+1}-x^k\|^2}{\|u^k\|^2 + \|x^k\|^2}}.
\]

<table>
<thead>
<tr>
<th>( r = 1, s = 100 )</th>
<th>( e = 10^{-4} )</th>
<th>( e = 5 \cdot 10^{-5} )</th>
<th>( e = 10^{-5} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>iter</td>
<td>time (s)</td>
<td>iter</td>
</tr>
<tr>
<td>PCP</td>
<td>9265</td>
<td>22.28</td>
<td>14570</td>
</tr>
<tr>
<td>CP</td>
<td>9732</td>
<td>23.04</td>
<td>15718</td>
</tr>
<tr>
<td>%improv.</td>
<td>4.8</td>
<td>3.3</td>
<td>7.3</td>
</tr>
</tbody>
</table>

**Table:** Average time and number of iterations when \( m = 1 \) for obtaining

\( R_k < e \).
**Numerical experiences**

\[
\min_{R x = c, S x = d} \|x\|_1
\]

- Is the case \( f = 0, h = \nu \{ (c, d) \} \) and \( g = \| \cdot \|_1 \).
- Dimensions: \( N = 1000, s = 100 \) and \( r \in \{ 1, 10, 30 \} \).
- We consider 20 random realizations of matrices \( R, S \) and vectors \( c \) and \( d \) and we measure the relative error

\[
R_k = \sqrt{\frac{\|u^{k+1} - u^k\|_2^2 + \|x^{k+1} - x^k\|_2^2}{\|u^k\|_2^2 + \|x^k\|_2^2}}.
\]

<table>
<thead>
<tr>
<th>( r = 10, s = 100 )</th>
<th>( e = 10^{-4} )</th>
<th>( e = 5 \cdot 10^{-5} )</th>
<th>( e = 10^{-5} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>iter</td>
<td>time (s)</td>
<td>iter</td>
<td>time (s)</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>PCP</td>
<td>6865</td>
<td>18.65</td>
<td>10229</td>
</tr>
<tr>
<td>CP</td>
<td>9280</td>
<td>23.72</td>
<td>16033</td>
</tr>
<tr>
<td>%improv.</td>
<td>26.0</td>
<td>21.4</td>
<td>36.2</td>
</tr>
</tbody>
</table>

**Table:** Average time and number of iterations when \( m = 10 \) for obtaining \( R_k < e \).
Numerical experiences

\[
\min_{Rx=c, Sx=d} \|x\|_1
\]

- Is the case \( f = 0, h = \nu\{(c,d)\} \) and \( g = \| \cdot \|_1 \).
- Dimensions: \( N = 1000, s = 100 \) and \( r \in \{1, 10, 30\} \).
- We consider 20 random realizations of matrices \( R, S \) and vectors \( c \) and \( d \) and we measure the relative error

\[
R_k = \sqrt{\frac{\|u^{k+1}-u^k\|^2 + \|x^{k+1}-x^k\|^2}{\|u^k\|^2 + \|x^k\|^2}}.
\]

<table>
<thead>
<tr>
<th>( r = 30, s = 100 )</th>
<th>( e = 10^{-4} )</th>
<th>( e = 5 \cdot 10^{-5} )</th>
<th>( e = 10^{-5} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iter</td>
<td>Time (s)</td>
<td>Iter</td>
<td>Time (s)</td>
</tr>
<tr>
<td>PCP</td>
<td>5146</td>
<td>7.68</td>
<td>7143</td>
</tr>
<tr>
<td>CP</td>
<td>9941</td>
<td>12.93</td>
<td>16438</td>
</tr>
<tr>
<td>%improv.</td>
<td>48.2</td>
<td>40.6</td>
<td>56.5</td>
</tr>
</tbody>
</table>

**Table:** Average time and number of iterations when \( m = 30 \) for obtaining
\( R_k < e \).
1. Problem and motivation

2. Algorithm, convergence, and numerical simulations

3. Application

4. Open questions
Application to Mean Field Games

\[ \inf_{(m, w)} \int_{T^d} \left[ b_2(m(x), w(x)) + F(x, m(x)) \right] \, dx, \]

subject to \[
\begin{align*}
-\nu \Delta m + \text{div}(w) &= 0 \quad \text{in } T^d, \\
\int_{T^d} m(x) \, dx &= 1.
\end{align*}
\]

After discretizing

\[ \inf_{(m, w)} \sum_{i, j} \left[ b_2(m_{i,j}, w_{i,j}) + F(x_{i,j}, m_{i,j}) \right] \, dx, \]

subject to \[
\begin{align*}
-\nu \Delta_h m + \text{div}_h w &= 0 \\
h^2 \sum_{i,j} m_{i,j} &= 1.
\end{align*}
\]

Application to Mean Field Games

Denoting \( \varphi = b_2 + F \), \( A = -\nu \Delta_h \), \( B = \text{div}_h \), the classic CP splitting reads:

\[
\begin{align*}
\begin{pmatrix} u^{k+1} \\ \lambda^{k+1} \end{pmatrix} &= \begin{pmatrix} u^k + \gamma(A\bar{m}^k + B\bar{w}^k) \\ \lambda^k + \gamma(h^2 \sum_{i,j} \bar{m}_{i,j}^k - 1) \end{pmatrix} \\
\begin{pmatrix} m^{k+1} \\ w^{k+1} \end{pmatrix} &= \text{prox}_{\tau \varphi} \left( m^k - \tau(A^\top u^{k+1} + h^2 \lambda^{k+1} \mathbf{1}) \right) \\
\begin{pmatrix} \bar{m}^{k+1} \\ \bar{w}^{k+1} \end{pmatrix} &= \begin{pmatrix} 2m^{k+1} - m^k \\ 2w^{k+1} - w^k \end{pmatrix}
\end{align*}
\]

but primal iterates do not satisfy any of the constraints, which extremely affect the speed of the method.

\[^3\text{BA-Kalise-Silva Alvarez (2018).}\]
Application to Mean Field Games

Imposing, the constraint $\int m = 1$ the projected CP reads

$$
\begin{align*}
(u^{k+1}, \lambda^{k+1}) &= (u^k + \gamma(A\bar{m}^k + B\bar{w}^k), \\
 & \quad \lambda^k + \gamma(h^2 \sum_{i,j} \bar{m}_{i,j}^k - 1))
\end{align*}
$$

$$
\begin{align*}
(n^{k+1}, v^{k+1}) &= \text{prox}_{\tau \varphi} \left( m^k - \tau(A^\top u^{k+1} + h^2 \lambda^{k+1} 1) \right) \\
 & \quad w^k - \tau B^\top u^{k+1}
\end{align*}
$$

$$
\begin{align*}
(m^{k+1}, w^{k+1}) &= (1 + (n^{k+1} - 1 \sum_{i,j=1}^{N_h} n_{i,j}^{k+1}) v^{k+1}) \\
 & \quad m^k + n^{k+1} - m^k
\end{align*}
$$

$$
\begin{align*}
(\bar{m}^{k+1}, \bar{w}^{k+1}) &= (m^{k+1} + n^{k+1} - m^k) \\
 & \quad w^{k+1} + v^{k+1} - w^k
\end{align*}
$$

which is much faster, and any of the primal iterates satisfy the imposed constraint. By including splitting, we do not need any matrix inversion.

1 Problem and motivation

2 Algorithm, convergence, and numerical simulations

3 Application

4 Open questions
Redundancy in constrained primal-dual algorithm

Problem (P) with \( C = R^{-1}c \)

\[
(Q) \quad \text{find} \quad \hat{x} \in R^{-1}c \cap \arg \min_{Rx=c} f(x) + g(x).
\]

Projected primal-dual splitting

\( x^0 = \bar{x}^0 \in \mathbb{R}^N \) and \( u^0 \in \mathbb{R}^M \), \( \tau, \gamma > 0 \) such that \( \tau \gamma \|L\|^2 < 1 - \frac{\tau}{2\beta} \)

\[
\begin{align*}
  u^{k+1} &= u^k + \gamma (L\bar{x}^k - b) \\
  p^{k+1} &= \text{prox}_{\tau g}(x^k - \tau (\nabla f(x^k) + L^\top u^{k+1})) \\
  x^{k+1} &= p^{k+1} - R^\top (RR^\top)^{-1} (Rp^{k+1} - c) \\
  \bar{x}^{k+1} &= x^{k+1} + p^{k+1} - x^k.
\end{align*}
\]

Then, \( (x^k)_{k \in \mathbb{N}} \subseteq C \) and \( x^k \to \hat{x} \) solution to \( (Q) \).

- Open question: Redundancy is better than splitting?
Randomized Kaczmarz ($f = g = 0$)

- If we focus on the system $Lx = b$ ($f = g = 0$), in the results discussed above, we select deterministically some rows onto which we project.

- Kaczmarz (1937): alternating projection onto the $M$ lines of $Lx = b$ for solving the linear system (under uniqueness of the solution).

- Strohmer-Vershynin (2009): Randomized Kaczmarz (proportional to the norm of $\ell_i$). Linear convergence depending on the condition number of $L$. 
Randomized Kaczmarz for constrained primal-dual problems

- In the case when $f$ or $g$ are not trivial, we are currently studying if Kaczmarz (or other deterministic ways) or randomized Kaczmarz help in the efficiency of primal-dual methods. Case $h = \nu\{b\}$.

- In the general case, we may assume that $C = \cap_{i \in I} C_i$ and use deterministic or randomized approaches.

- **Open questions**: Blocks. Can we exploit structure of $L$? Ex.: block diagonal, other possibilities/ideas?

- And if $f: (x_1, \ldots, x_n) \mapsto \sum_{i=1}^n f_i(x_i)$? Can we exploit separability at the same time?
References