

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)

$$\underset{x \in \mathbb{R}^N}{\text{minimize}} \quad f(x) + g(x) + h(Lx).$$

- $f: \mathbb{R}^N \rightarrow \mathbb{R}$ is differentiable and ∇f is β^{-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: $h = \iota_{\{b\}}$ for some $b \in \mathbb{R}^M$

$$\underset{Lx=b}{\text{minimize}} \quad f(x) + g(x).$$

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

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

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

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 $\varphi: (m, w) \mapsto \int_{T^d} [b_2(m(x), w(x)) + F(x, m(x))] dx$. 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)

$$\underset{x \in \mathbb{R}^N}{\text{minimize}} \quad f(x) + g(x) + h(Lx).$$

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

$$x^{k+1} = \operatorname{argmin}_x \mathcal{L}_\gamma(x, y^k, u^k)$$

$$y^{k+1} = \operatorname{argmin}_y \mathcal{L}_\gamma(x^{k+1}, y, u^k)$$

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

Classic approach: ADMM

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

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

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

Classic approach: ADMM

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

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

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

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

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

$$u^{k+1} = u^k + \gamma(Lx^{k+1} - b).$$

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

$$\underset{y \in \mathbb{R}^N}{\text{minimize}} \quad f(y) + \frac{1}{2} \|x - y\|^2.$$

- Several functions f have an explicit or efficiently computable prox_f . Examples: $\|\cdot\|_1$, ι_C ($\text{prox}_{\iota_C} = P_C$), 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

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

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

Other approaches

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

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

Combettes-Pesquet (2012)

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

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

- 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 ∇f .

Other approaches

Problem (P)

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

Condat-Vũ (2013)

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

$$\begin{cases} u^{k+1} = \text{prox}_{\gamma h^*}(u^k + \gamma L\bar{x}^k) \\ x^{k+1} = \text{prox}_{\tau g}(x^k - \tau(\nabla f(x^k) + L^\top u^{k+1})) \\ \bar{x}^{k+1} = 2x^{k+1} - x^k. \end{cases}$$

- 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 $\tau\gamma\|L\|^2 < 1 - \frac{\tau}{2\beta}$

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

- $\text{prox}_{\gamma h^*} = \text{Id} - \gamma \text{prox}_{h/\gamma} \circ (\text{Id}/\gamma) = \text{Id} - \gamma b$.
- 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 splitting¹

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 $\tau\gamma\|L\|^2 < 1 - \frac{\tau}{2\beta}$

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

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

¹Joint work with Sergio López Rivera, USM

Projected primal-dual splitting¹

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

Problem (P) with a priori information C

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

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 $\tau\gamma\|L\|^2 < 1 - \frac{\tau}{2\beta}$

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

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

¹Joint work with Sergio López Rivera, USM

Projected primal-dual splitting: case $h = \iota_{\{b\}}$

- In the case when $h = \iota_{\{b\}}$, 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 $d \in \mathbb{R}^s$.
- 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)

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

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

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

Projected primal-dual splitting: case $h = \iota_{\{b\}}$

- In the case when $h = \iota_{\{b\}}$, 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 $d \in \mathbb{R}^s$.
- Suppose that it is possible to project onto $C := R^{-1}c \subset L^{-1}b$.

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

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

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 $\tau\gamma\|L\|^2 < 1 - \frac{\tau}{2\beta}$

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

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

Numerical experiences

$$\min_{\substack{Rx=c \\ Sx=d}} \|x\|_1$$

- Is the case $f = 0$, $h = \iota_{\{(c,d)\}}$ 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

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

$r = 1, s = 100$	$e = 10^{-4}$		$e = 5 \cdot 10^{-5}$		$e = 10^{-5}$	
	iter	time (s)	iter	time (s)	iter	time (s)
PCP	9265	22.28	14570	37.02	46191	116.26
CP	9732	23.04	15718	39.21	50544	125.49
%improv.	4.8	3.3	7.3	5.6	8.6	7.4

Table: Average time and number of iterations when $m = 1$ for obtaining

$$R_k < e.$$

Numerical experiences

$$\min_{\substack{Rx=c \\ Sx=d}} \|x\|_1$$

- Is the case $f = 0$, $h = \iota_{\{(c,d)\}}$ 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

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

$r = 10, s = 100$	$e = 10^{-4}$		$e = 5 \cdot 10^{-5}$		$e = 10^{-5}$	
	iter	time (s)	iter	time (s)	iter	time (s)
PCP	6865	18.65	10229	27.86	22855	65.05
CP	9280	23.72	16033	39.13	49526	129.78
%improv.	26.0	21.4	36.2	28.8	53.9	49.9

Table: Average time and number of iterations when $m = 10$ for obtaining

$$R_k < e.$$

Numerical experiences

$$\min_{\substack{Rx=c \\ Sx=d}} \|x\|_1$$

- Is the case $f = 0$, $h = \iota_{\{(c,d)\}}$ 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

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

$r = 30, s = 100$	$e = 10^{-4}$		$e = 5 \cdot 10^{-5}$		$e = 10^{-5}$	
	iter	time (s)	iter	time (s)	iter	time (s)
PCP	5146	7.68	7143	10.67	13421	19.70
CP	9941	12.93	16438	21.37	50841	64.23
%improv.	48.2	40.6	56.5	50.1	73.6	69.3

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 Games²

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

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

After discretizing

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

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

²BA-Kalise- Silva Alvarez (2018).

Application to Mean Field Games³

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

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

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

³BA-Kalise-Silva Alvarez (2018).

Application to Mean Field Games³

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

$$\begin{aligned} \begin{pmatrix} u^{k+1} \\ \lambda^{k+1} \end{pmatrix} &= \begin{pmatrix} u^k + \gamma(A\bar{m}^k + B\bar{w}^k) \\ \lambda^k + \gamma(h^2 \sum_{i,j} \bar{m}_{i,j}^k - 1) \end{pmatrix} \\ \begin{pmatrix} n^{k+1} \\ v^{k+1} \end{pmatrix} &= \text{prox}_{\tau\varphi} \begin{pmatrix} m^k - \tau(A^\top u^{k+1} + h^2 \lambda^{k+1} \mathbf{1}) \\ w^k - \tau B^\top u^{k+1} \end{pmatrix} \\ \begin{pmatrix} m^{k+1} \\ w^{k+1} \end{pmatrix} &= \begin{pmatrix} \mathbf{1} + (n^{k+1} - \mathbf{1} \sum_{i,j=1}^{N_h} n_{i,j}^{k+1}) \\ v^{k+1} \end{pmatrix} \\ \begin{pmatrix} \bar{m}^{k+1} \\ \bar{w}^{k+1} \end{pmatrix} &= \begin{pmatrix} m^{k+1} + n^{k+1} - m^k \\ w^{k+1} + v^{k+1} - w^k \end{pmatrix} \end{aligned}$$

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

³BA-Kalise-Silva Alvarez (2018).

- 1 Problem and motivation
- 2 Algorithm, convergence, and numerical simulations
- 3 Application
- 4 Open questions

Redundancy in constrained primal-dual algorithm

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

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

Projected primal-dual splitting

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

$$\begin{aligned} u^{k+1} &= u^k + \gamma(L\bar{x}^k - b) \\ p^{k+1} &= \text{prox}_{\tau g}(x^k - \tau(\nabla f(x^k) + L^\top u^{k+1})) \\ x^{k+1} &= \textcolor{red}{p}^{k+1} - \textcolor{red}{R}^\top (\textcolor{red}{R}\textcolor{red}{R}^\top)^{-1} (\textcolor{red}{R}p^{k+1} - c) \\ \bar{x}^{k+1} &= x^{k+1} + p^{k+1} - x^k. \end{aligned}$$

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

- Open question: Redundancy is better than splitting ?

References

- L. M. Briceño-Arias and S. López Rivera, A projected primal-dual splitting for solving constrained monotone inclusions, *J. Optim. Theory Appl.*, vol. 180, no. 3, pp. 907–924, 2019.
- L. M. Briceño-Arias, D. Kalise, and F. J. Silva, Proximal methods for stationary Mean Field Games with local couplings, *SIAM J. Control Optim.* 56, 801–836, 2018.
- L. M. Briceño-Arias and P. L. Combettes, A monotone + skew splitting model for composite monotone inclusions in duality, *SIAM J. Optim.* 21, 1230–1250, 2011.
- A. Chambolle and T. Pock, A first order primal dual algorithm for convex problems with applications to imaging, *J. Math. Imaging Vis.* 40, 120–145, 2011.
- L. Condat, A primal-dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms, *J. Optim. Theory Appl.* 158, 460–479, 2013.
- G. Chen and M. Teboulle, A proximal-based decomposition method for convex minimization problems, *Math. Programming Ser. A* 64, 81–101, 1994.
- J.-M. Lasry and P.-L. Lions. Jeux à champ moyen I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343:619–625, 2006.
- B.C. Vũ, A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Adv. Comput. Math.* 38, 667–681, 2013.