Algorithms Based on Unions of Nonexpansive Maps

Matthew K. Tam
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

joint work with Minh N. Dao (University of Newcastle)

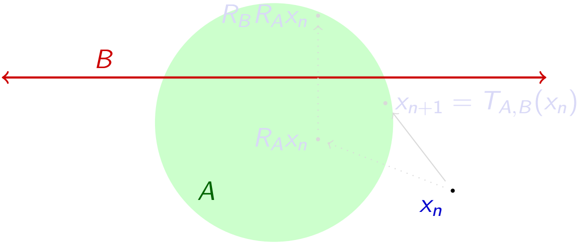
ISMP in Bordeaux, July 1–6, 2018

Background: The Douglas–Rachford Algorithm

Let $A, B \subseteq \mathbb{R}^d$. Given $x_0 \in \mathbb{R}^d$, the Douglas–Rachford algorithm can be compactly described as the fixed point iteration given by $T_{A,B}$, that is,

$$x_{n+1} \in T_{A,B}(x_n) := \left(\frac{\mathrm{Id} + R_B R_A}{2}\right)(x_n) \qquad \forall n \in \mathbb{N}$$

where $R_A := 2P_A - \mathrm{Id}$ and $P_A(x) =$ nearest point(s) to x in A.



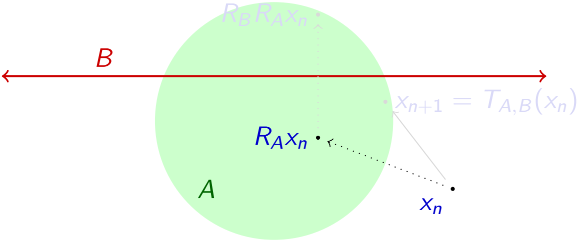
$$A = \{x \in \mathbb{R}^d : \|x\| \leq 1\}, \quad B = \{x \in \mathbb{R}^d : \langle a, x \rangle = b\}.$$

Background: The Douglas–Rachford Algorithm

Let $A, B \subseteq \mathbb{R}^d$. Given $x_0 \in \mathbb{R}^d$, the Douglas–Rachford algorithm can be compactly described as the fixed point iteration given by $T_{A,B}$, that is,

$$x_{n+1} \in T_{A,B}(x_n) := \left(\frac{\mathrm{Id} + R_B R_A}{2}\right)(x_n) \qquad \forall n \in \mathbb{N}$$

where $R_A := 2P_A - \mathrm{Id}$ and $P_A(x) =$ nearest point(s) to x in A.



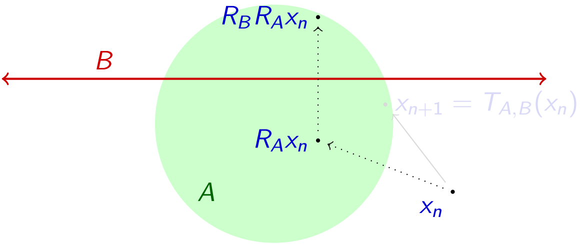
$$A = \{x \in \mathbb{R}^d : \|x\| \leq 1\}, \quad B = \{x \in \mathbb{R}^d : \langle a, x \rangle = b\}.$$

Background: The Douglas–Rachford Algorithm

Let $A, B \subseteq \mathbb{R}^d$. Given $x_0 \in \mathbb{R}^d$, the Douglas–Rachford algorithm can be compactly described as the fixed point iteration given by $T_{A,B}$, that is,

$$x_{n+1} \in T_{A,B}(x_n) := \left(\frac{\mathrm{Id} + R_B R_A}{2}\right)(x_n) \qquad \forall n \in \mathbb{N}$$

where $R_A := 2P_A - \mathrm{Id}$ and $P_A(x) =$ nearest point(s) to x in A.



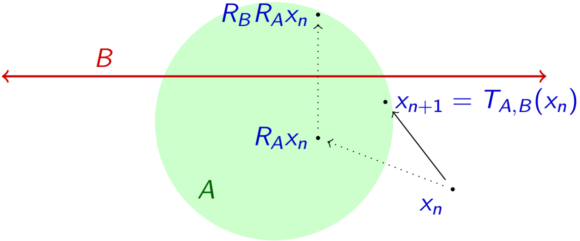
$$A = \{x \in \mathbb{R}^d : \|x\| \leq 1\}, \quad B = \{x \in \mathbb{R}^d : \langle a, x\rangle = b\}.$$

Background: The Douglas–Rachford Algorithm

Let $A, B \subseteq \mathbb{R}^d$. Given $x_0 \in \mathbb{R}^d$, the Douglas–Rachford algorithm can be compactly described as the fixed point iteration given by $T_{A,B}$, that is,

$$x_{n+1} \in T_{A,B}(x_n) := \left(\frac{\mathrm{Id} + R_B R_A}{2}\right)(x_n) \qquad \forall n \in \mathbb{N}$$

where $R_A := 2P_A - \mathrm{Id}$ and $P_A(x) =$ nearest point(s) to x in A.



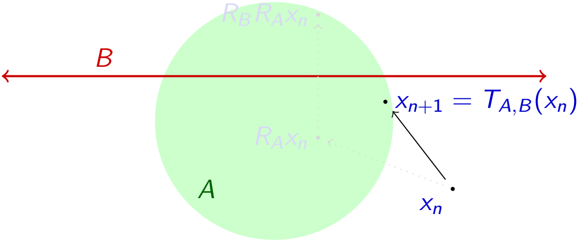
$$A = \{x \in \mathbb{R}^d : \|x\| \leq 1\}, \quad B = \{x \in \mathbb{R}^d : \langle a, x \rangle = b\}.$$

Background: The Douglas–Rachford Algorithm

Let $A, B \subseteq \mathbb{R}^d$. Given $x_0 \in \mathbb{R}^d$, the Douglas–Rachford algorithm can be compactly described as the fixed point iteration given by $T_{A,B}$, that is,

$$x_{n+1} \in T_{A,B}(x_n) := \left(\frac{\mathrm{Id} + R_B R_A}{2}\right)(x_n) \qquad \forall n \in \mathbb{N}$$

where $R_A := 2P_A - \mathrm{Id}$ and $P_A(x) =$ nearest point(s) to x in A.



$$A = \{x \in \mathbb{R}^d : \|x\| \leq 1\}, \quad B = \{x \in \mathbb{R}^d : \langle a, x\rangle = b\}.$$

Background: The Douglas–Rachford Algorithm

A non-comprehensive history:

1. Douglas–Rachford ('56): proposed for solving heat flow problems.
2. Lions–Mercier ('79): globally convergent if A and B are convex.
3. Bauschke–Noll ('14): locally convergent around strong fixed points if $A = \cup_{i \in I} A_i$ and $B = \cup_{j \in J} B_i$ are finite unions of closed convex sets.

For a set-valued map $T : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$, there are two types of fixed points.

- Weak fixed point set:

$$\operatorname{Fix} T := \{x \in \mathbb{R}^d : x \in T(x)\}.$$

- Strong fixed point set:

$$\mathbf{Fix}\, T := \{x \in \mathbb{R}^d : \{x\} = T(x)\}.$$

Both notions coincide for single-valued operators.

Background: The Douglas–Rachford Algorithm

A non-comprehensive history:

1. Douglas–Rachford ('56): proposed for solving heat flow problems.
2. Lions–Mercier ('79): globally convergent if A and B are convex.
3. Bauschke–Noll ('14): locally convergent around strong fixed points if $A = \cup_{i \in I} A_i$ and $B = \cup_{j \in J} B_i$ are finite unions of closed convex sets.

For a set-valued map $T : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$, there are two types of fixed points.

- Weak fixed point set:

$$\mathrm{Fix}\, T := \{x \in \mathbb{R}^d : x \in T(x)\}.$$

- Strong fixed point set:

$$\mathbf{Fix}\, T := \{x \in \mathbb{R}^d : \{x\} = T(x)\}.$$

Both notions coincide for single-valued operators.

Union Averaged Operators

An operator $S : \mathbb{R}^d \to \mathbb{R}^d$ is α-averaged if $\alpha \in (0,1)$ and it holds that

$$S = \alpha I + (1-\alpha)R$$

for some nonexpansive operator $R : \mathbb{R}^d \to \mathbb{R}^d$.

Definition (Union Averaged)

We say a (set-valued) operator $T : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is union α-averaged if it can be represented in the form

$$T(x) = \{T_i(x) : i \in \phi(x)\} \quad \forall x \in \mathbb{R}^d, \tag{1}$$

where

- I is a finite index set,
- $\{T_i\}_{i \in I}$ is a collection of α-averaged operators, and
- $\phi : \mathbb{R}^d \rightrightarrows I$ is nonempty-valued and outer semicontinuous (osc):

$$\phi(x) \supseteq \operatorname*{Lim\,sup}_{y \to x} \phi(y) := \{i \in I : \exists (x_n, i_n) \to (x, i) \text{ s.t. } i_n \in \phi(x_n)\}.$$

Union Averaged Operators: A First Example

Let $A = \bigcup_{i \in I} A_I$ where A_i is convex. Then $P_A : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is given by

$$P_A(x) := \{a \in A : \|x - a\| = d(x, A)\} = \{P_{A_i}(x) : i \in \phi(x)\}$$

where $\phi(x) := \{i \in I : d(x, A_i) = \min_{i \in I} d(x, A_j)\}$.

To illustrate the idea, consider the union of two convex sets $A = A_1 \cup A_2$.

On the dashed line, $\phi(x) = \{1, 2\}$ which makes ϕ outer semicontinuous.

Union Averaged Operators: A First Example

Let $A = \bigcup_{i \in I} A_I$ where A_i is convex. Then $P_A : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is given by

$$P_A(x) := \{a \in A : \|x - a\| = d(x, A)\} = \{P_{A_i}(x) : i \in \phi(x)\}$$

where $\phi(x) := \{i \in I : d(x, A_i) = \min_{i \in I} d(x, A_j)\}$.

To illustrate the idea, consider the union of two convex sets $A = A_1 \cup A_2$.

On the dashed line, $\phi(x) = \{1, 2\}$ which makes ϕ outer semicontinuous.

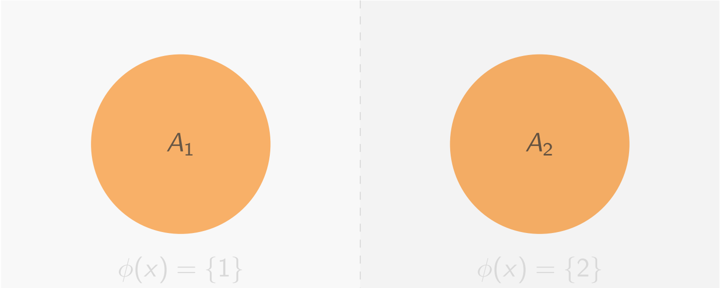
Union Averaged Operators: A First Example

Let $A = \bigcup_{i \in I} A_I$ where A_i is convex. Then $P_A : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is given by

$$P_A(x) := \{a \in A : \|x - a\| = d(x, A)\} = \{P_{A_i}(x) : i \in \phi(x)\}$$

where $\phi(x) := \{i \in I : d(x, A_i) = \min_{i \in I} d(x, A_j)\}$.

To illustrate the idea, consider the union of two convex sets $A = A_1 \cup A_2$.



On the dashed line, $\phi(x) = \{1, 2\}$ which makes ϕ outer semicontinuous.

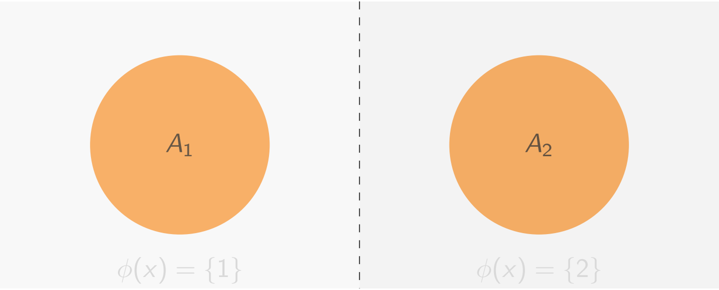
Union Averaged Operators: A First Example

Let $A = \bigcup_{i \in I} A_I$ where A_i is convex. Then $P_A : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is given by

$$P_A(x) := \{a \in A : \|x - a\| = d(x, A)\} = \{P_{A_i}(x) : i \in \phi(x)\}$$

where $\phi(x) := \{i \in I : d(x, A_i) = \min_{i \in I} d(x, A_j)\}$.

To illustrate the idea, consider the union of two convex sets $A = A_1 \cup A_2$.



On the dashed line, $\phi(x) = \{1, 2\}$ which makes ϕ outer semicontinuous.

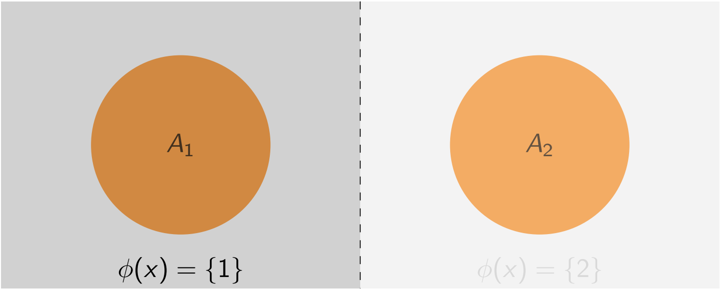
Union Averaged Operators: A First Example

Let $A = \bigcup_{i \in I} A_I$ where A_i is convex. Then $P_A : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is given by

$$P_A(x) := \{a \in A : \|x - a\| = d(x, A)\} = \{P_{A_i}(x) : i \in \phi(x)\}$$

where $\phi(x) := \{i \in I : d(x, A_i) = \min_{i \in I} d(x, A_j)\}$.

To illustrate the idea, consider the union of two convex sets $A = A_1 \cup A_2$.



On the dashed line, $\phi(x) = \{1, 2\}$ which makes ϕ outer semicontinuous.

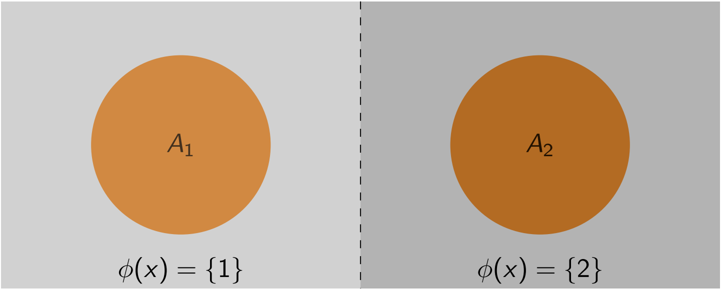
Union Averaged Operators: A First Example

Let $A = \bigcup_{i \in I} A_I$ where A_i is convex. Then $P_A : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is given by

$$P_A(x) := \{a \in A : \|x - a\| = d(x, A)\} = \{P_{A_i}(x) : i \in \phi(x)\}$$

where $\phi(x) := \{i \in I : d(x, A_i) = \min_{i \in I} d(x, A_j)\}$.

To illustrate the idea, consider the union of two convex sets $A = A_1 \cup A_2$.



On the dashed line, $\phi(x) = \{1, 2\}$ which makes ϕ outer semicontinuous.

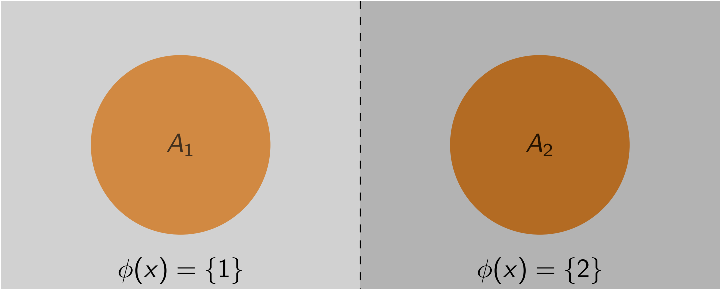
Union Averaged Operators: A First Example

Let $A = \bigcup_{i \in I} A_I$ where A_i is convex. Then $P_A : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is given by

$$P_A(x) := \{a \in A : \|x - a\| = d(x, A)\} = \{P_{A_i}(x) : i \in \phi(x)\}$$

where $\phi(x) := \{i \in I : d(x, A_i) = \min_{i \in I} d(x, A_j)\}$.

To illustrate the idea, consider the union of two convex sets $A = A_1 \cup A_2$.



On the dashed line, $\phi(x) = \{1, 2\}$ which makes ϕ outer semicontinuous.

Union Averaged Operators: Sparsity Constraints

Let $A = \bigcup_{i \in I} A_I$ where A_i is convex. Then $P_A : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is given by

$$P_A(x) := \{a \in A : \|x - a\| = d(x, A)\} = \{P_{A_i}(x) : i \in \phi(x)\}$$

where $\phi(x) := \{i \in I : d(x, A_i) = \min_{i \in I} d(x, A_j)\}$.

Let $s \in \{0, 1 \ldots, d\}$. Consider the sparsity constraint given by

$$A := \{x \in \mathbb{R}^d : \|x\|_0 \leq s\}.$$

By denoting $\mathcal{I} := \{I \in 2^{\{1,2,\ldots,d\}} : |I| = s\}$, we may express the sparsity constraint A as a finite union of subspaces:

$$A = \bigcup_{I \in \mathcal{I}} A_I \quad \text{where} \quad A_I := \{x \in \mathbb{R}^d : x_i \neq 0 \text{ only if } i \in I\}.$$

It follows that P_A is union 1/2-averaged nonexpansive.

Union Averaged Operators: Sparsity Constraints

Let $A = \bigcup_{i \in I} A_I$ where A_i is convex. Then $P_A : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is given by

$$P_A(x) := \{a \in A : \|x - a\| = d(x, A)\} = \{P_{A_i}(x) : i \in \phi(x)\}$$

where $\phi(x) := \{i \in I : d(x, A_i) = \min_{j \in I} d(x, A_j)\}$.

Let $s \in \{0, 1 \ldots, d\}$. Consider the sparsity constraint given by

$$A := \{x \in \mathbb{R}^d : \|x\|_0 \leq s\}.$$

By denoting $\mathcal{I} := \{I \in 2^{\{1,2,\ldots,d\}} : |I| = s\}$, we may express the sparsity constraint A as a finite union of subspaces:

$$A = \bigcup_{I \in \mathcal{I}} A_I \quad \text{where} \quad A_I := \{x \in \mathbb{R}^d : x_i \neq 0 \text{ only if } i \in I\}.$$

It follows that P_A is union $1/2$-averaged nonexpansive.

Building New Union Averaged Operators

Union averaged operators are closed under various operators.

Proposition (Dao–T.)

Let $I := \{1, \dots, m\}$ and, for each $i \in I$, suppose $T_i : X \rightrightarrows X$ is union α_i-averaged nonexpansive. Then the following assertions hold.

(a) $T := \sum_{i \in I} \omega_i T_i$ is union α-averaged nonexpansive with

$$\alpha := \sum_{i \in I} \omega_i \alpha_i$$

whenever $(\omega_i)_{i \in I} \subseteq \mathbb{R}_{++}$ with $\sum_{i \in I} \omega_i = 1$.

(b) $T' := T_m \circ \cdots \circ T_2 \circ T_1$ is union α'-averaged nonexpansive with

$$\alpha' := \left(1 + \left(\sum_{i \in I} \frac{\alpha_i}{1 - \alpha_i}\right)^{-1}\right)^{-1}.$$

Local Convergence

Theorem (T., 2018)

Suppose $T : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is union α-averaged with $x^* \in \textbf{Fix}\, T$. Define

$$r := \sup \{\delta > 0 : \phi(x) \subseteq \phi(x^*) \text{ for all } x \in \mathbb{B}(x^*;\delta)\}. \quad (2)$$

Then $r > 0$ and, for any $\epsilon \in (0, r)$, it holds that

$$\|y - x^*\| \leq \|x - x^*\| \text{ whenever } x \in \mathbb{B}(x^*;\epsilon),\ y \in T(x). \quad (3)$$

Furthermore, if $x_0 \in \mathbb{B}(x^*;\epsilon)$ and $x_{n+1} \in T(x_n)$ for all $n \in \mathbb{N}$, then the fixed point iterates $(x_n)_{n\in\mathbb{N}}$ converge to a point $\overline{x} \in \text{Fix}\, T \cap \mathbb{B}(x^*;\epsilon)$.

Consequences and remarks:

- If $\exists x^* \in \textbf{Fix}\, T$ s.t. $\phi(x^*) = I$, then $r = +\infty \implies$ glob. convergence.
- In general, $(x_n)_{n\in\mathbb{N}}$ need not stay in $\mathbb{B}(x^*;\epsilon)$ if $x^* \in \text{Fix}\, T \setminus \textbf{Fix}\, T$.
- In general, the limit $\overline{x} = \lim_{n\to\infty} x_n$ need not be an element of $\textbf{Fix}\, T$.

Local Convergence

Theorem (T., 2018)

Suppose $T : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is union α-averaged with $x^* \in \mathbf{Fix}\, T$. Define

$$r := \sup\left\{\delta > 0 : \phi(x) \subseteq \phi(x^*) \text{ for all } x \in \mathbb{B}(x^*;\delta)\right\}. \tag{2}$$

Then $r > 0$ and, for any $\epsilon \in (0, r)$, it holds that

$$\|y - x^*\| \leq \|x - x^*\| \quad \text{whenever} \quad x \in \mathbb{B}(x^*;\epsilon),\ y \in T(x). \tag{3}$$

Furthermore, if $x_0 \in \mathbb{B}(x^*;\epsilon)$ and $x_{n+1} \in T(x_n)$ for all $n \in \mathbb{N}$, then the fixed point iterates $(x_n)_{n\in\mathbb{N}}$ converge to a point $\overline{x} \in \text{Fix}\, T \cap \mathbb{B}(x^*;\epsilon)$.

Consequences and remarks:

- If $\exists x^* \in \mathbf{Fix}\, T$ s.t. $\phi(x^*) = I$, then $r = +\infty \implies$ glob. convergence.
- In general, $(x_n)_{n\in\mathbb{N}}$ need not stay in $\mathbb{B}(x^*;\epsilon)$ if $x^* \in \text{Fix}\, T \setminus \mathbf{Fix}\, T$.
- In general, the limit $\overline{x} = \lim_{n\to\infty} x_n$ need not be an element of $\mathbf{Fix}\, T$.

Local Convergence

This following example is due to Bauschke and Noll (2014).

Example (strong fixed point needed)

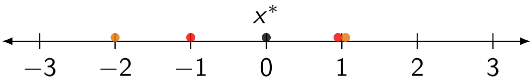
Consider the DR operator $T_{A,B} : \mathbb{R} \rightrightarrows \mathbb{R}$ given by

$$T_{A,B} := \frac{I + R_B R_A}{2}$$

for the sets $A = \{-1, 1\}$ and $B = \{-2, 1\}$. Then

$$x^* := 0 \in \text{Fix } T_{A,B} \setminus \textbf{Fix } T_{A,B}$$

but, for any $x \in (-\epsilon, 0)$ with $\epsilon \approx 0$, $x_+ \in T_{A,B}x$ need be in $\mathbb{B}(x^*, \epsilon)$.



- Fixed point iterates are *stable* around strong fixed points.

Local Convergence

Example (limit points need not be strong fixed points)

Consider $T : \mathbb{R} \rightrightarrows \mathbb{R}$ defined by

$$T(x) := \{T_1(x), T_2(x)\},$$

where $T_1(x) = 0$ and $T_2(x) = x$ (*i.e.*, $\phi(x) = \{1, 2\}$ for all $x \in \mathbb{R}$). Then

$$\mathbf{Fix}\, T = \{0\} \quad \text{and} \quad \mathrm{Fix}\, T = \mathbb{R}.$$

For any $\overline{x} \in \mathbb{R} \setminus \{0\}$, the constant sequence $x_n = x$ for all $n \in \mathbb{N}$ satisfies $x_{n+1} \in T(x_n)$ but (trivially) converges to $\overline{x} \in \mathrm{Fix}\, T \setminus \mathbf{Fix}\, T$.

1. Nonconvex Feasibility Problems

Let $J = \{1, \ldots, m\}$. Consider the nonconvex feasibility problem:

$$\text{find } x \in \bigcap_{j \in J} C_j \quad \text{with} \quad C_j = \bigcup_{i \in I_j} C_{i,j}, \tag{4}$$

where all index sets are finite, and the $C_{i,j}$'s are closed and convex.

- The projector P_{C_j} is union $\frac{1}{2}$-averaged nonexpansive, for all $j \in J$.

While combinatorial feasibility problems (*e.g.*, Sudoku) are of the form (4), local convergence guarantees are not usually interesting because one can simply "round" near solutions.

		5	3					
8							2	
	7			1		5		
4					5	3		
	1			7				6
		3	2				8	
	6		5					9
		4					3	
					9	7		

1	4	5	3	2	7	6	9	8
8	3	9	6	5	4	1	2	7
6	7	2	9	1	8	5	4	3
4	9	6	1	8	5	3	7	2
2	1	8	4	7	3	9	5	6
7	5	3	2	9	6	4	8	1
3	6	7	5	4	2	1	8	9
9	8	4	7	6	1	2	3	5
5	2	1	8	3	9	7	6	4

1. Nonconvex Feasibility Problems

Let $J = \{1, \ldots, m\}$. Consider the nonconvex feasibility problem:

$$\text{find } x \in \bigcap_{j \in J} C_j \quad \text{with} \quad C_j = \bigcup_{i \in I_j} C_{i,j}, \tag{4}$$

where all index sets are finite, and the $C_{i,j}$'s are closed and convex.

- The projector P_{C_j} is union $\frac{1}{2}$-averaged nonexpansive, for all $j \in J$.

While combinatorial feasibility problems (*e.g.*, Sudoku) are of the form (4), local convergence guarantees are not usually interesting because one can simply "round" near solutions.

		5	3					
8							2	
	7			1		5		
4					5	3		
	1			7				6
		3	2				8	
	6		5					9
		4					3	
					9	7		

1	4	5	3	2	7	6	9	8
8	3	9	6	5	4	1	2	7
6	7	2	9	1	8	5	4	3
4	9	6	1	8	5	3	7	2
2	1	8	4	7	3	9	5	6
7	5	3	2	9	6	4	8	1
3	6	7	5	4	2	1	8	9
9	8	4	7	6	1	2	3	5
5	2	1	8	3	9	7	6	4

1. Nonconvex Feasibility Problems

Theorem (Dao-T.)

Let $x^* \in \cap_{j\in J} C_j$ for sets C_j as in (4). Then there exists $r > 0$ such that if $x_0 \in \mathbb{B}(x^*; r)$ and $(x_n)_{n\in\mathbb{N}}$ satisfies

$$\forall n \in \mathbb{N}, \quad x_{n+1} \in T(x_n),$$

then $x_n \to \overline{x}$ and either $\overline{x} \in \cap_{j\in J} C_j$ or $P_{C_1}(\overline{x}) \cap (\cap_{j\in J} C_j) \neq \emptyset$ whenever

1. $T := P_{C_m} \circ \cdots \circ P_{C_1}$ (method of cyclic projections).
2. $T = T_{C_m,C_1} \circ \ldots T_{C_2,C_3} \circ T_{C_1,C_2}$ and C_1 convex (CA-DR algorithm).
3. $T := T_{C_1,C_2}$ (DR algorithm).

- Not immediate that the limit point $\overline{x}$ is related to the intersection!

Other projection methods still converge locally, but it unclear whether their weak fixed points can be used to obtain a point in the intersection.

2. The Proximal Point Algorithm

Definition (min-convex)

A function $g : \mathbb{R}^d \to (-\infty, +\infty]$ is min-convex if there exists a finite set I and a collection of proper, lsc, convex functions $\{f_i\}_{i \in I}$ such that

$$g(x) = \min_{i \in I} g(x), \quad \forall x \in \mathbb{R}^d.$$

- g min-convex & $\gamma > 0 \implies \mathrm{prox}_{\gamma g}$ union $\frac{1}{2}$-averaged nonexpansive.

Theorem (Dao-T.)

Let $\gamma > 0$, suppose g is min-convex and $x^* \in \mathbf{Fix}(\mathrm{prox}_{\gamma g})$. Then there exists an $r > 0$ such that if $x_0 \in \mathbb{B}(x^*; r)$ and $(x_n)_{n \in \mathbb{N}}$ satisfies

$$\forall n \in \mathbb{N}, \quad x_{n+1} \in \mathrm{prox}_{\gamma g}(x_n),$$

then $(x_n)_{n \in \mathbb{N}}$ converges to a local minimum of g.

3. Sparsity Constrained Minimisation

Consider the sparsity constrained minimisation problem

$$\min_{x\in\mathbb{R}^d}\{f(x) : \|x\|_0 \le s\}, \tag{P}$$

where $f : \mathbb{R}^d \to \mathbb{R}$ is convex with L-Lipschitz gradient, $\|x\|_0$ denotes the ℓ_0-functional, and $s \in \{0, 1, \dots, d-1\}$ is an *a priori* sparsity estimate.

The forward-backward algorithm for (P) can be compactly described as

$$x_{n+1} \in T(x_n) := P_A(x_n - \gamma\nabla f(x_n)) \quad \forall n \in \mathbb{N},$$

where the step-size satisfies $\gamma \in (0, 2/L)$ and $A := \{x \in \mathbb{R}^d : \|x\|_0 \le s\}$.

1. Near strong fixed points of T, any sequence (x_n) satisfying (14) is locally convergent to a weak fixed point.
2. Every weak fixed point is a local minimia of the (P).
3. Weak fixed points are strong whenever γ is sufficiently small.

3. Sparsity Constrained Minimisation

Consider the sparsity constrained minimisation problem

$$\min_{x\in\mathbb{R}^d}\{f(x) : \|x\|_0 \le s\}, \qquad \text{(P)}$$

where $f : \mathbb{R}^d \to \mathbb{R}$ is convex with L-Lipschitz gradient, $\|x\|_0$ denotes the ℓ_0-functional, and $s \in \{0, 1, \ldots, d-1\}$ is an *a priori* sparsity estimate.

The forward-backward algorithm for (P) can be compactly described as

$$x_{n+1} \in T(x_n) := P_A\left(x_n - \gamma\nabla f(x_n)\right) \quad \forall n \in \mathbb{N},$$

where the step-size satisfies $\gamma \in (0, 2/L)$ and $A := \{x \in \mathbb{R}^d : \|x\|_0 \le s\}$.

1. Near strong fixed points of T, any sequence (x_n) satisfying (14) is locally convergent to a weak fixed point.
2. Every weak fixed point is a local minimia of the (P).
3. Weak fixed points are strong whenever γ is sufficiently small.

3. Sparsity Constrained Minimisation

In general, this is a special case of the following result. Consider

$$\min_{x\in\mathbb{R}^d}\{f(x)+g(x)\},$$

where the function $f:\mathbb{R}^d\to\mathbb{R}$ is convex with L-Lipschitz gradient and $g:\mathbb{R}^d\to(-\infty,+\infty]$ is min-convex.

(For sparsity constrained min, take g to be the indicator function of A.)

Theorem (T., 2018, Dao-T.)

Suppose that $x^*\in\textbf{Fix}\ T_{\rm FB}$. Then there exists an $r>0$ such that if $x_0\in\mathbb{B}(x^*;r)$ and $(x_n)_{n\in\mathbb{N}}$ satisfies

$$x_{n+1}\in T(x_n):=\operatorname{prox}_{\gamma g}(x_n-\gamma\nabla f(x_n))\quad\forall n\in\mathbb{N},\tag{5}$$

then $x_n\to\overline{x}\in\text{Fix}\ T_{\rm FB}$ which is local minimum of $f+g$.

Concluding Remarks

- We introduced the notion of union averaged operators which generalises the ideas of Bauschke & Noll (2014).
- The corresponding fixed point iterations are locally convergent around strong fixed points.
- For reasonable (proximal) methods, there is a correspondence between fixed points and local minima.
- Applications in nonconvex feasibility problems, the proximal point algorithm, and in sparsity constrained minimisation.
- **Open question:** The "right" framework for infinite dimensional extensions (weak osc of ϕ is too strong in general).

M.K. Tam: **Algorithms based on unions of nonexpansive maps**, *Optimization Letters*, 2018. Preprint: `arXiv:1510.06823`

M.N. Dao and M.K. Tam: **Union averaged operators for proximal algorithms**, *in preparation*.