Weak and Strong Superiorization: Between Feasiblity-seeking and Minimization

Yair Censor

University of Haifa

The Fourth Annual Loma Linda
Imaging and IMRT/IMPT Algorithm
Workshop 2018

August 6-8, 2018, LLU, Loma Linda, CA

TRANSFORM METHODS VS FULL DISCRETIZATION OF INVERSE PROBLEMS

$g(x)$ observation

$f(x)$ the casual factors that produced $g(x)$

Froward problem

$f(x)$ is known ▶ find $g(x)$

$$g(x) = \mathcal{R}(f)(x)$$

Inverse problem

$g(x)$ is known ▶ find $f(x)$

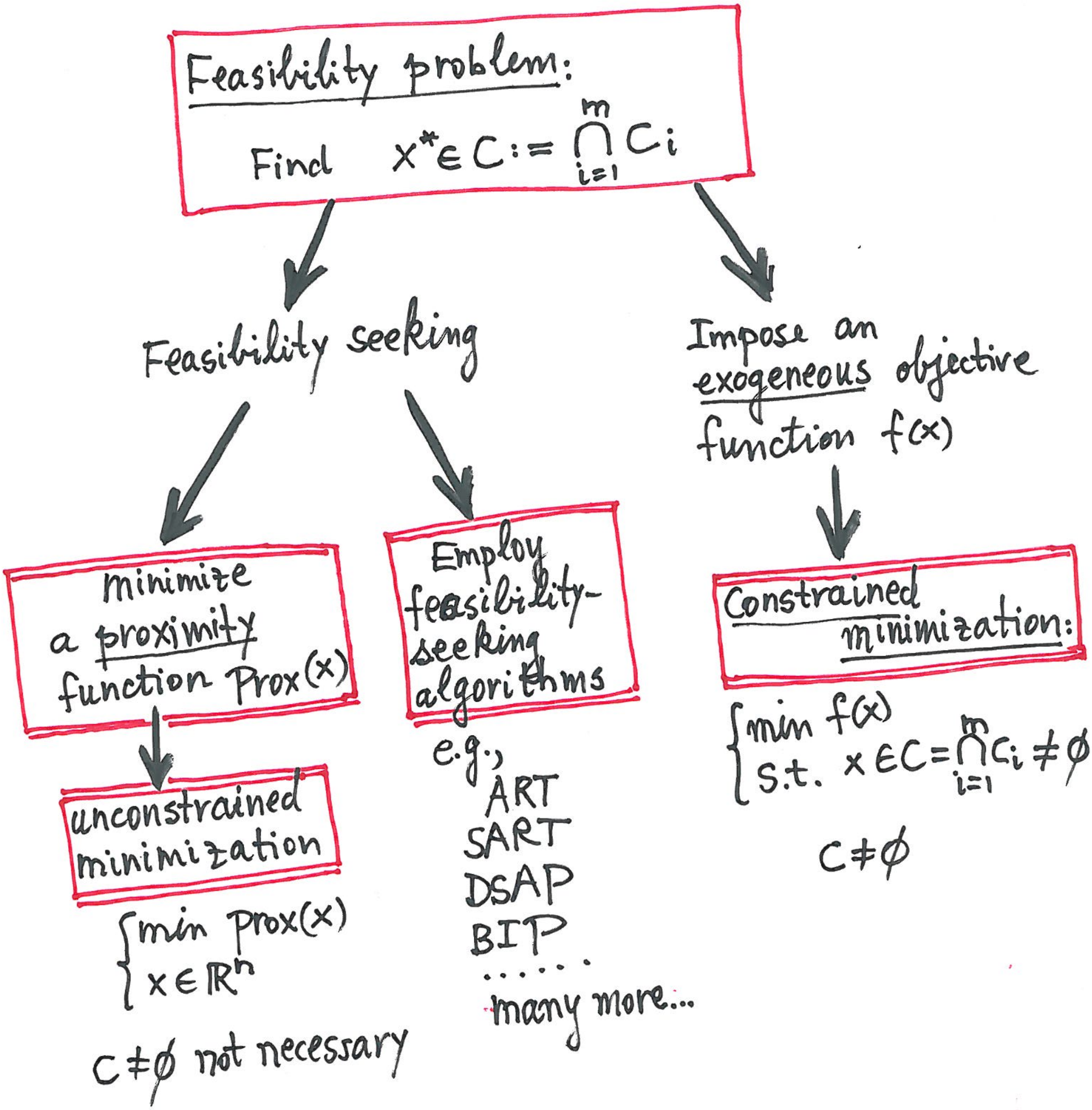
$$f(x) = \mathcal{R}^{-1}(g)(x)$$

↙ $\mathcal{R}$ and $\mathcal{R}^{-1}$ known and tractable → **Transform methods**

↘ otherwise → Full discretization → **Iterative feasibility-seeking of constraints**

The physical model gives rise to constraints such as

$$x \in C_i \quad i = 1, 2, \ldots, m, \qquad C_i \subseteq \mathbb{R}^n$$



Proximity function minimization leads to a simultaneous projection method

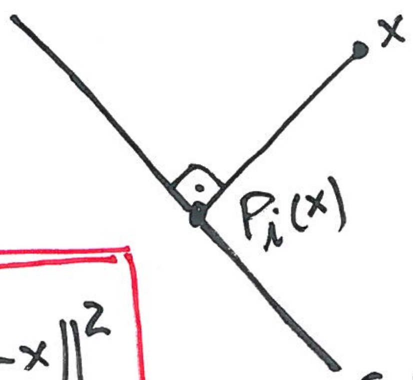
$$Ax = b \iff x \in C := \bigcap_{i=1}^{m} C_i$$

$$C_i = \{ x \in \mathbb{R}^n \mid \langle a^i, x \rangle = b_i \} \qquad i = 1, 2, \dots, m$$

$$\begin{cases} \min \frac{1}{2} \| Ax - b \|^2 \\ x \in \mathbb{R}^n \end{cases}$$

$$\Updownarrow$$

$$\begin{cases} \min \frac{1}{2} \sum_{i=1}^{m} \omega_i \| P_i(x) - x \|^2 \\ x \in \mathbb{R}^n \end{cases}$$



C_i is a hyperplane

Applying gradient descent method to this unconstrained minimization yields precisely the Cimmino simultaneous projection algorithm

$$\begin{cases} x^0 \in \mathbb{R}^n \\ x^{k+1} = x^k + \lambda_k \left(\sum_{i=1}^{m} P_i(x^k) - x^k \right) \end{cases}$$

Question: Can we do unconstrained minimization of another proximity function to obtain other projection methods? Such as ART?

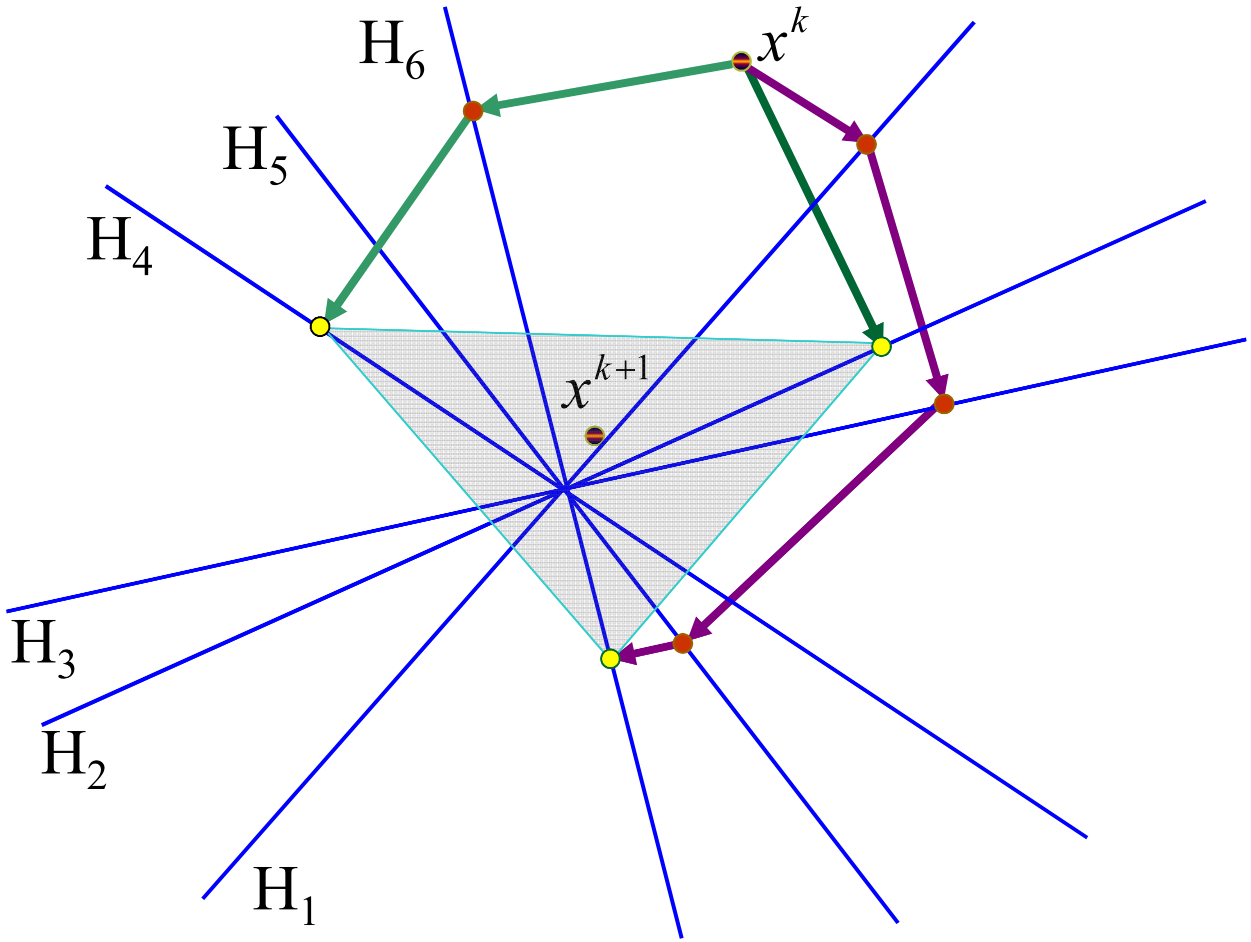
Answer: No.

Baillon, Combettes and Cominetti (2012)

Conclusion: Unconstrained proximity function minimization is equivalent to feasibility-seeking in principle but it cannot recover all algorithms for the feasibility problem that can be developed.

String averaging in general

Strings: $I_1 = (1,3,5,6)$ $I_2 = (2)$ $I_3 = (6,4)$



The string averaging algorithmic structure

For $t = 1, 2, \ldots, M,$ let $I_t = (i_1^t, i_2^t, \ldots, i_{m(t)}^t),$

be an ordered subset of $\{1, 2, \ldots, m\}$

$$x^0 \in S,$$

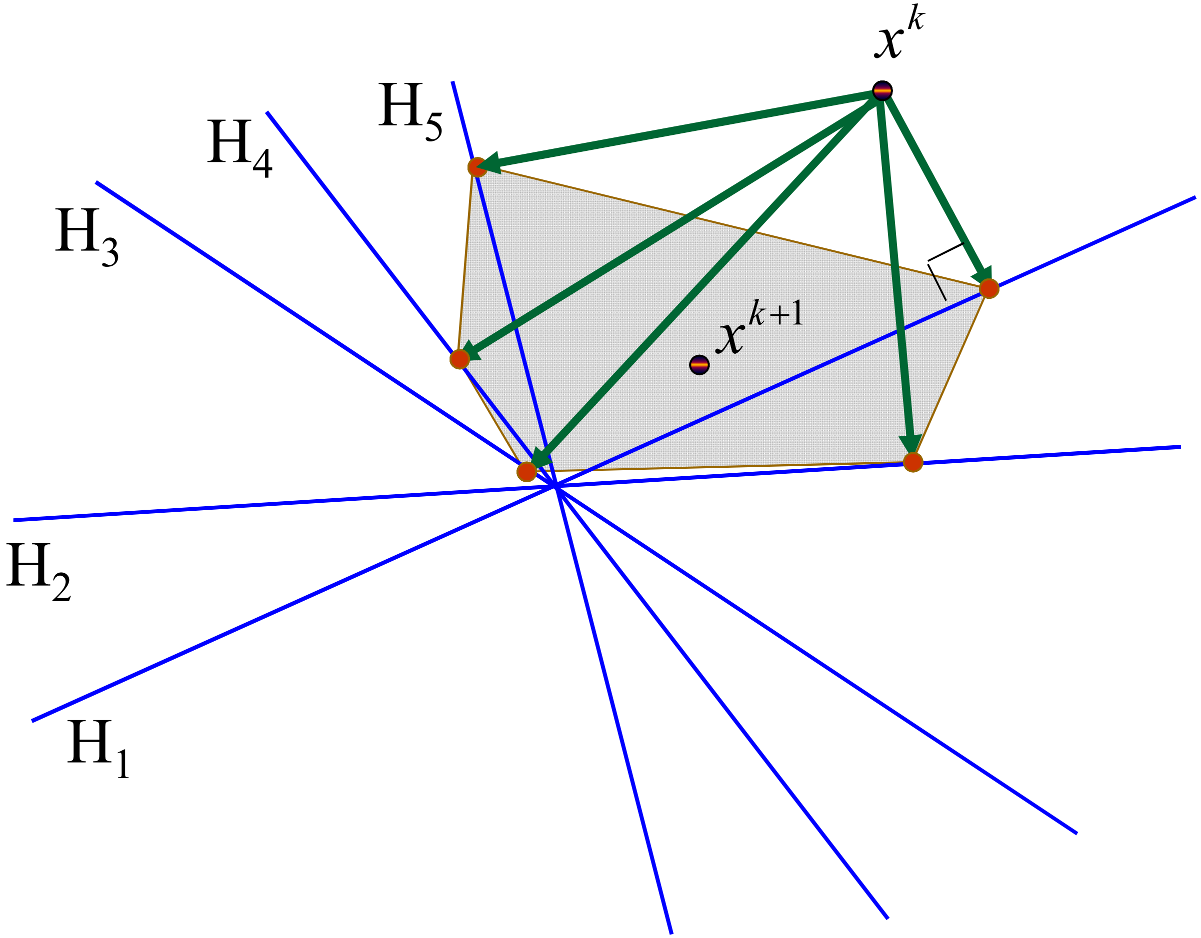
$$T_t x^k = R_{i_{m(t)}^t} \ldots R_{i_2^t} R_{i_1^t} x^k,$$

$$x^{k+1} = R(T_1 x^k, T_2 x^k, \ldots, T_M x^k).$$

Censor, Elfving and Herman, 2001

For example, if all the sets are hyperplanes…

$$\lambda_k = 1 \qquad x^{k+1} = x^k + \lambda_k \sum_{i=1}^{m} w_i (P_i(x^k) - x^k),$$



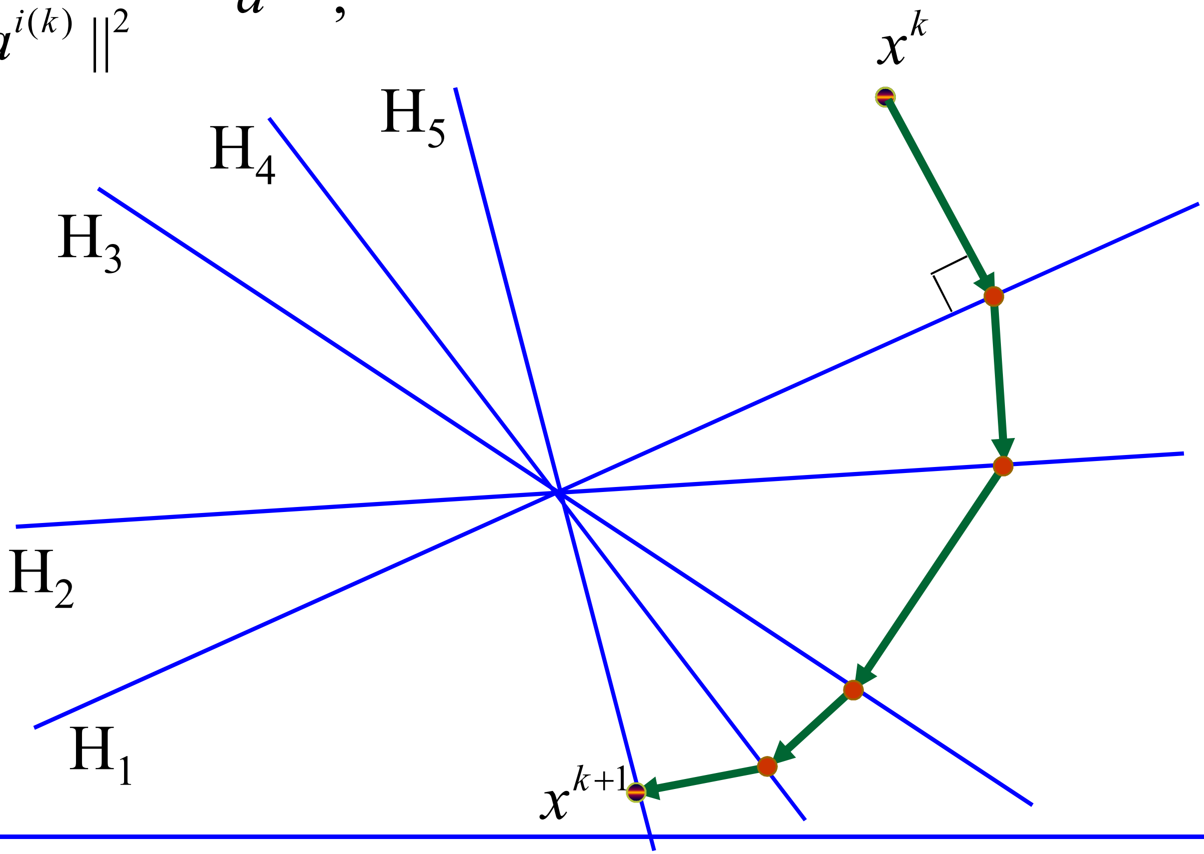
Cimmino 1937

Sequential Successive Projections (POCS, ART, Kaczmarz, Row-Action)

$$x^{k+1} = x^k + \lambda_k (P_{C_{i(k)}}(x^k) - x^k),\ \lambda_k = 1$$

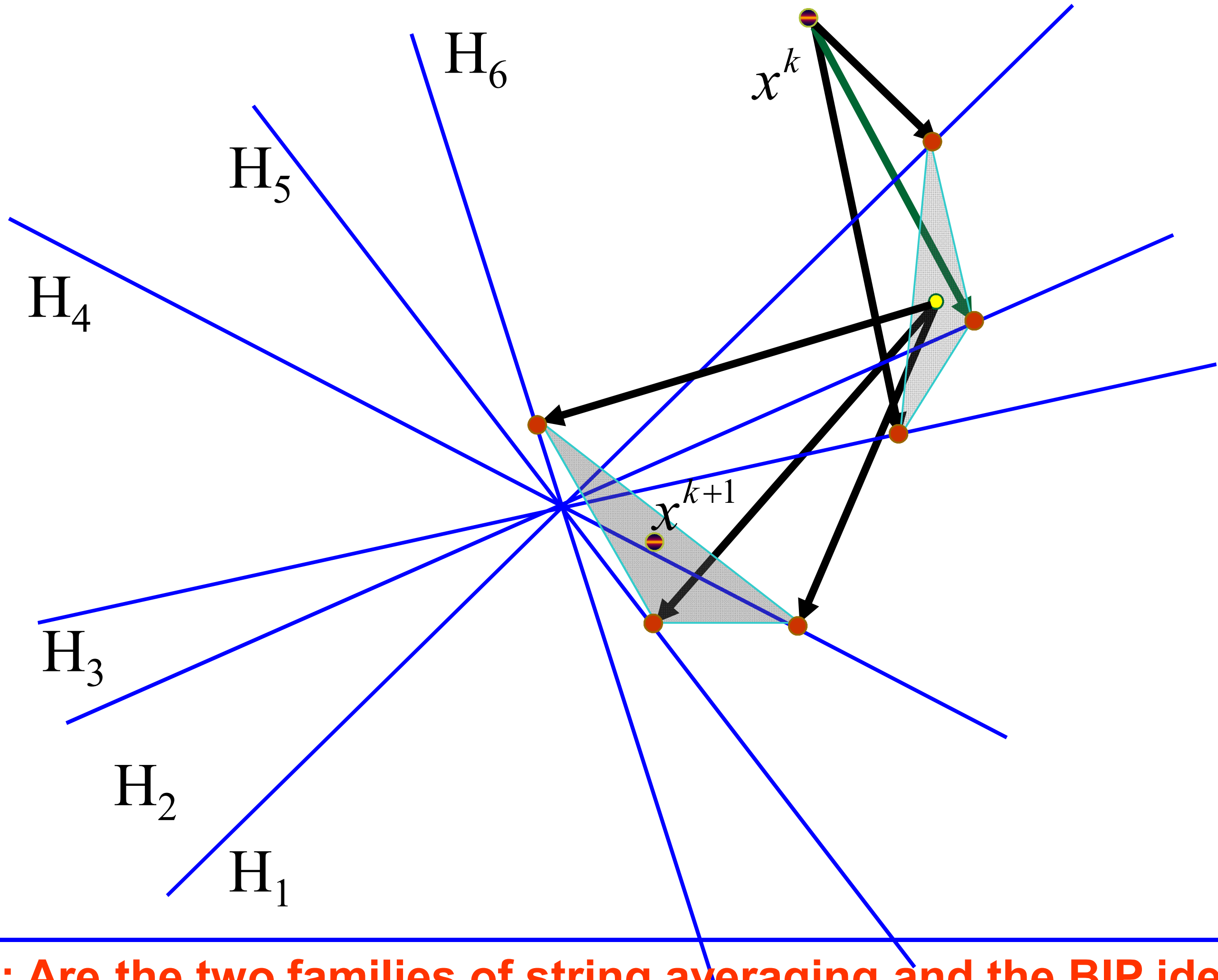
$$\langle a^i, x \rangle - b_i = 0, \quad i = 1, 2, \ldots, m,$$

$$x^{k+1} = x^k + \lambda_k \frac{b_{i(k)} - \langle a^{i(k)}, x^k \rangle}{\| a^{i(k)} \|^2} a^{i(k)},$$



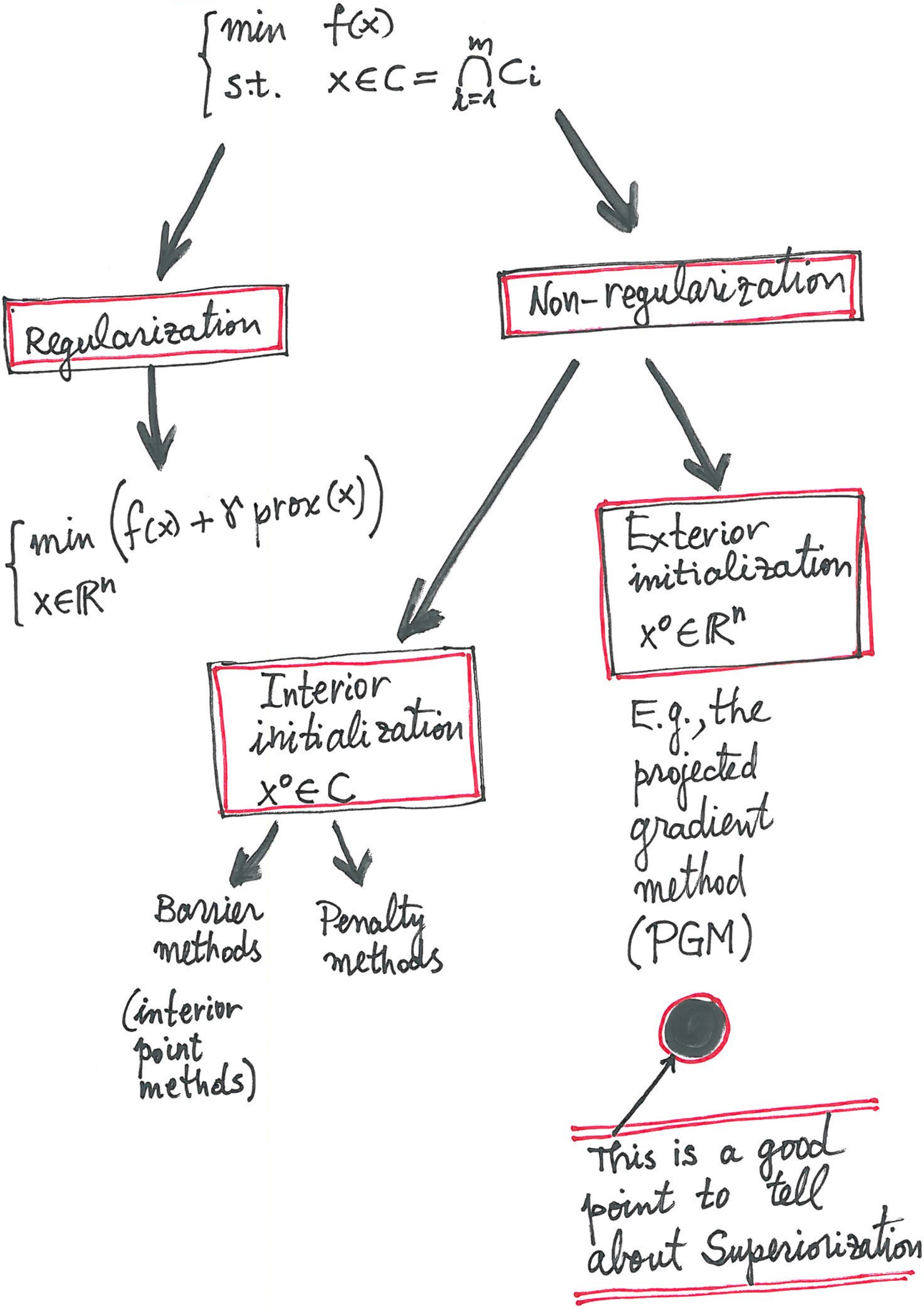
Block iterative projections (BIP)

Blocks: $B_1 = (1, 2, 3) \quad B_2 = (4, 5, 6)$



Question: Are the two families of string averaging and the BIP identical?

Constrained minimization



Projected Subgradient Minimization (PSM) method

Constrained minimization: $\min\{\phi(x) \mid x \in C\}$.

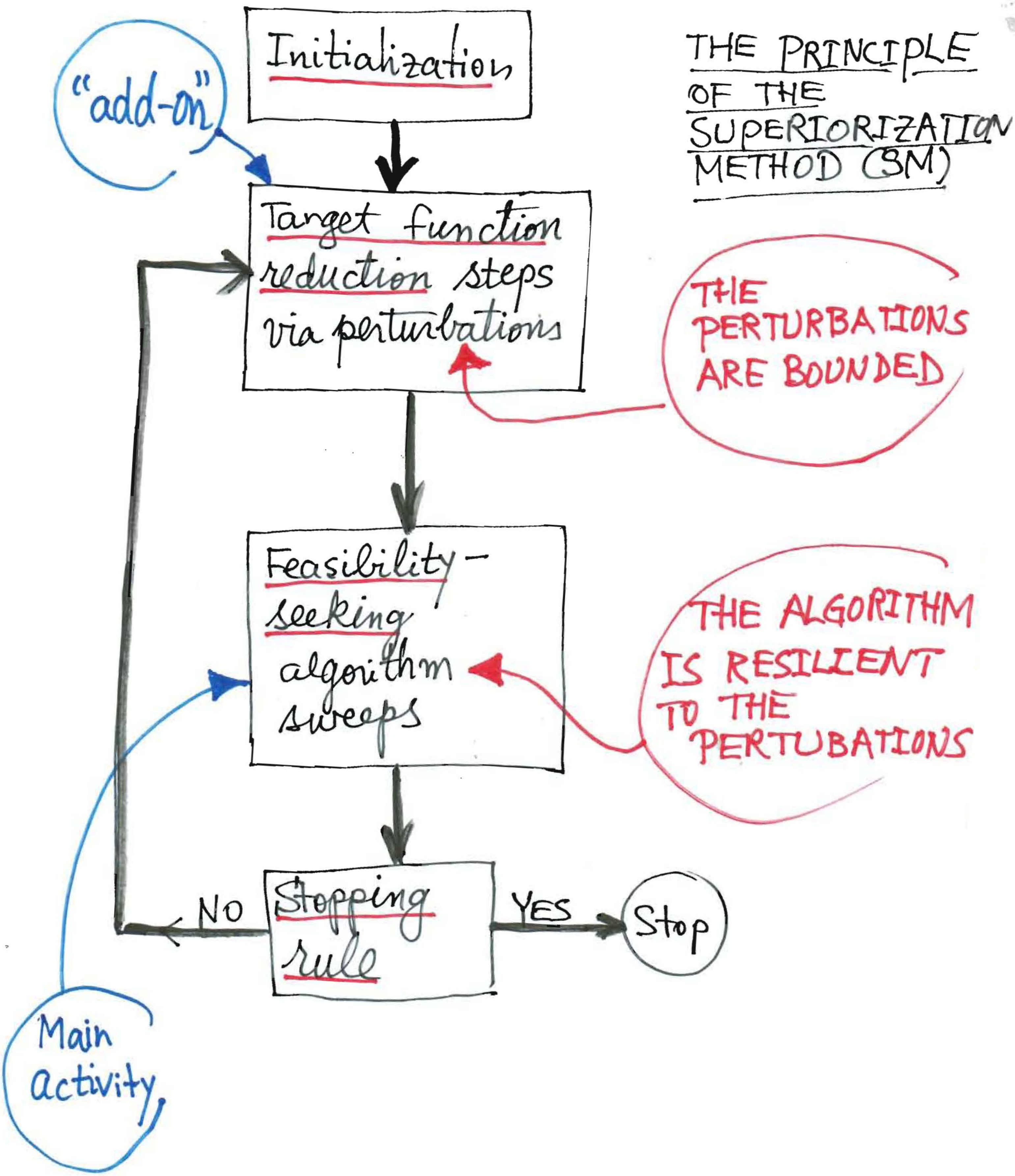
• C is nonempty closed convex set and ϕ is a convex function with domain that contains C.

$$\boxed{x^{k+1} = P_C\left(x^k - t_k\phi'(x^k)\right)}$$

• step-sizes $t_k > 0$, $\phi'(x^k) \in \partial\phi(x^k)$, and P_C is the projection onto C.

• **Underlying philosophy of PSM**: perform unconstrained objective function descent steps via $z^k := x^k - t_k\phi'(x^k)$ and repeatedly **regain feasibility** by doing a projection $P_C\left(z^k\right)$ onto C.

• **Major difficulty**: If C is not "simple to project onto" then the projection requires an independent inner-loop calculation to minimize the distance from the point z^k to the set C, which hampers overall effectiveness.



THE PRINCIPLE OF THE SUPERIORIZATION METHOD (SM)
Initialization
"add-on"
Target function reduction steps via perturbations
THE PERTURBATIONS ARE BOUNDED
Feasibility-seeking algorithm sweeps
THE ALGORITHM IS RESILIENT TO THE PERTUBATIONS
Stopping rule
NO
YES
Stop
Main activity

FEATURES OF THE SUPERIORIZATION METHOD (SM)

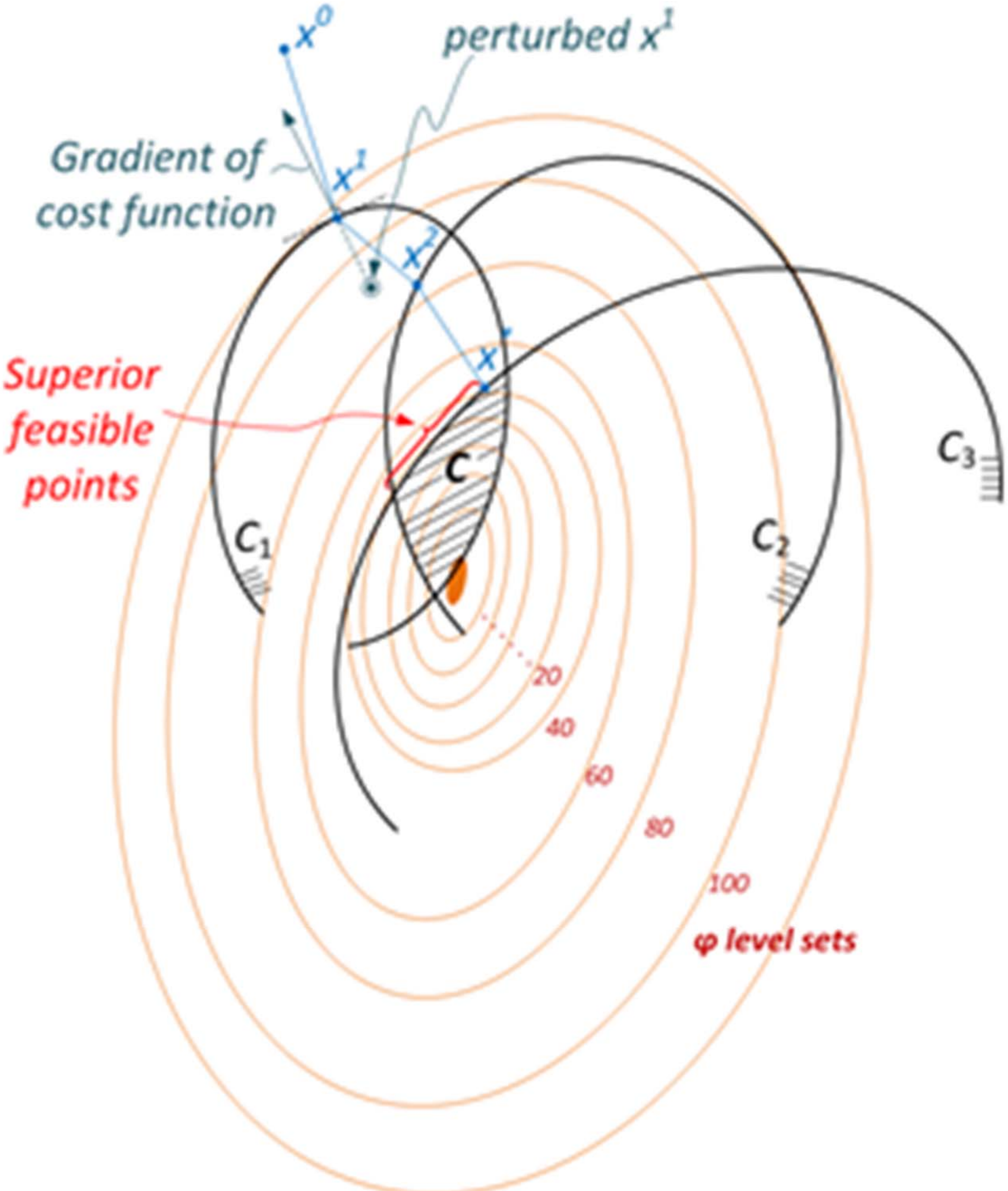
- Based on feasibility-seeking plus a target function reduction "add-on"
  - limited faith in the exogeneous target function
  - aims at function reduction not exact constrained minimization
- External initialization
- Allows to use efficient iterative feasibility-seeking algorithms
- Open mathematical problem:
  Missing certificate: Validation of global accumulation of all local function reductions
- 79 items to date since 2009 on:

http://math.haifa.ac.il/yair/bib-superiorization-censor.html

- Special Issue of Inverse Problems Vol. 33, April 2017.

Superiorization Diagram

- C is the feasible set defined by the intersection of many convex sets C_i
- ϕ is a target function to reduce (here not to minimize)
- Systematically perturb (add perturbation term) intermediate iterates from iterative projections in the direction of the negative gradient of ϕ
- Leads to feasible solution that is "superior" to one found without perturbations



5 Figures

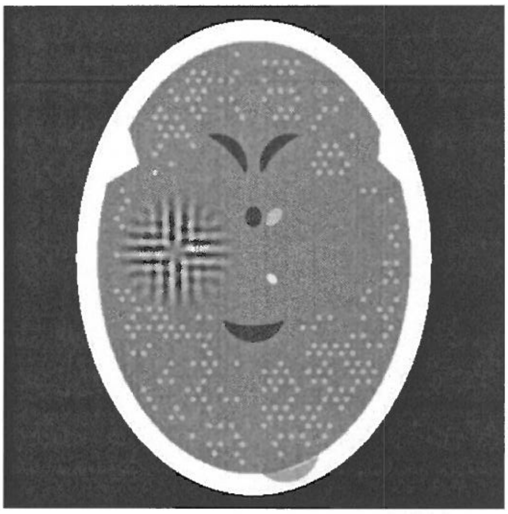


Figure 1: The head phantom. Its tomographic data was obtained for 60 views. It has TV=984.

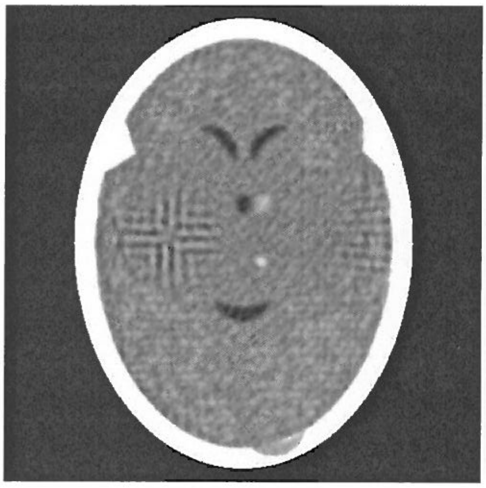


Figure 2: The image reconstructed by the projected subgradient method (PSM) has TV=919 and was obtained after 5257 seconds.

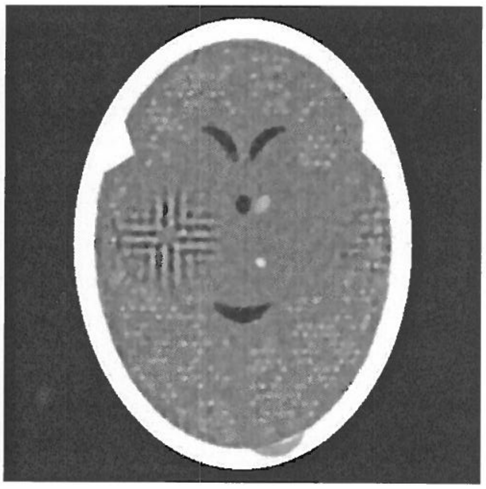


Figure 3: The image reconstructed by the superiorization method has TV=873 and was obtained after 318 seconds.

We explain now what we see in these figures. All computational work was done on a single machine, an Intel i5-3570K 3.4Ghz with 16GB RAM using the SNARK09 software package [27]; the phantom, the data, the reconstructions and displays were all generated within this same framework. In particular, this implies that differences in the reported reconstruction times are not due to the

Thank you

תודה רבה←

Pronounced: "Toda Raba"

The University of Haifa atop Mount Carmel (480 m), embedded in the Carmel Forest National Park