Accelerated Alternating Descent Methods for Dykstra-like problems

Antonin Chambolle, Pauline Tan and **Samuel Vaiter**[1]

[1]CNRS & IMB, Dijon, (Burgundy) France

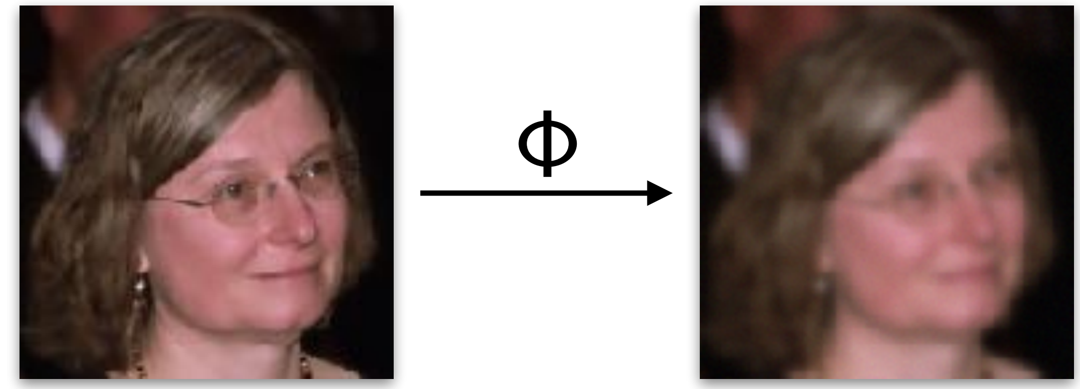
September 8, 2017

@Optimization 2017 — Lisbon

Total Variation in Imaging

Linear inverse problem

$$y = \Phi x_0 + w$$



Total Variation regularization

$$x^\star \in \underset{x \in \mathbb{R}^n}{\text{Argmin}} \; \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda J(x)$$

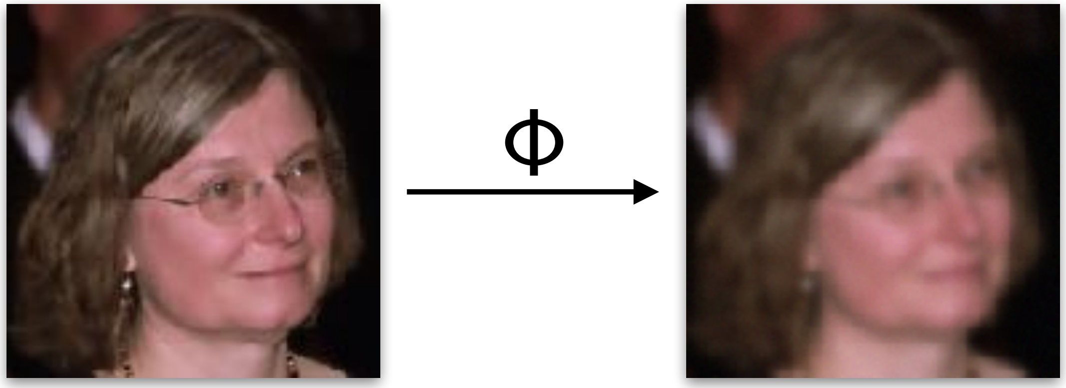
data fidelity

regularity of the edges

Total Variation in Imaging

Linear inverse problem

$$y = \Phi x_0 + w$$



Total Variation regularization

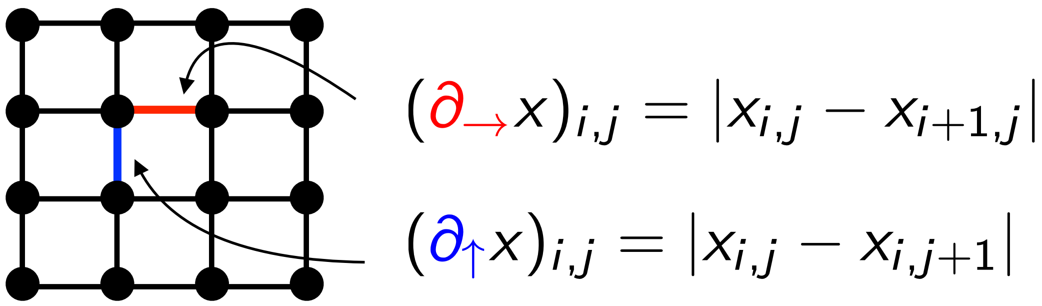
$$x^\star \in \underset{x \in \mathbb{R}^n}{\text{Argmin}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda J(x)$$

isotropic

$$J(x) = \sum_{i,j} \sqrt{(\partial_\rightarrow x)_{i,j}^2 + (\partial_\uparrow x)_{i,j}^2}$$

anisotropic

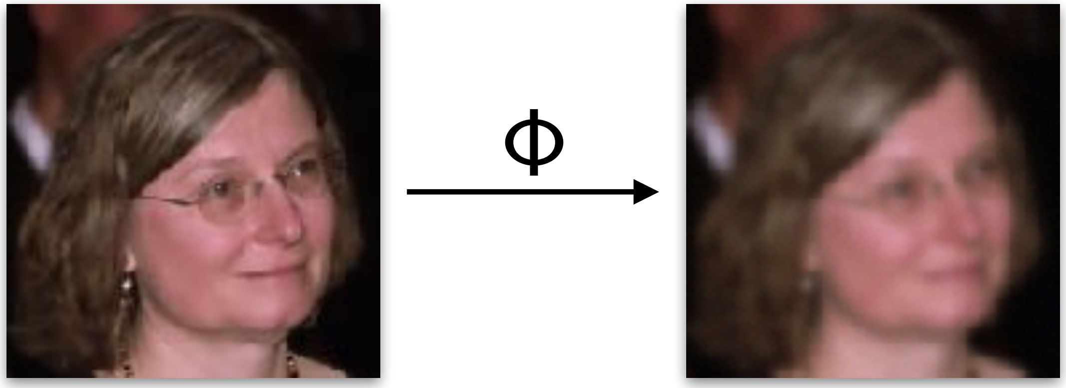
$$J(x) = \sum_{i,j} (\partial_\rightarrow x)_{i,j} + (\partial_\uparrow x)_{i,j}$$



Total Variation in Imaging

Linear inverse problem

$$y = \Phi x_0 + w$$



Total Variation regularization

$$x^\star \in \underset{x \in \mathbb{R}^n}{\operatorname{Argmin}} \frac{1}{2}\|y - \Phi x\|_2^2 + \lambda J(x)$$

This talk : How to compute this optimization problem (and some related problems)

Forward—Backward Method

$$x^\star \in \underset{x \in \mathbb{R}^n}{\text{Argmin}} \; \underbrace{\frac{1}{2}\|y - \Phi x\|_2^2}_{F(x)} + \lambda J(x)$$

Forward-backward

$$x^{(k+1)} = \text{Prox}_{\lambda J}(x^{(k)} - \gamma \nabla F(x^{(k)}))$$

Proximity operator

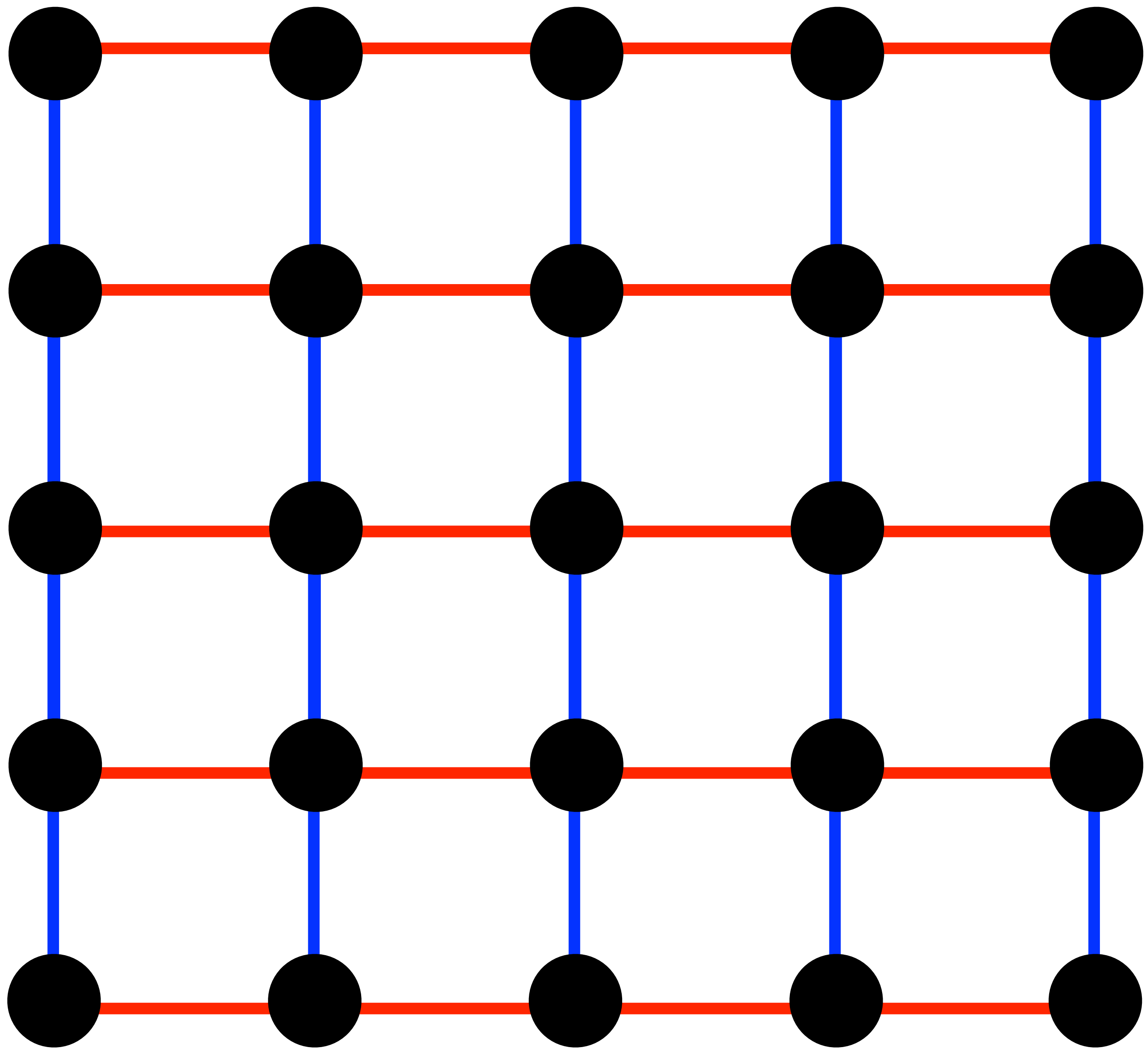
$$\text{Prox}_{\lambda J}(x) = \underset{z \in \mathbb{R}^n}{\text{argmin}} \; \frac{1}{2}\|x - z\|_2^2 + \lambda J(z) \longrightarrow$$ no closed formula for TV

Row/Column splitting [Condat 2013]

Augmented method [Chambolle-Pock 2011]

An Alternative Total Variation

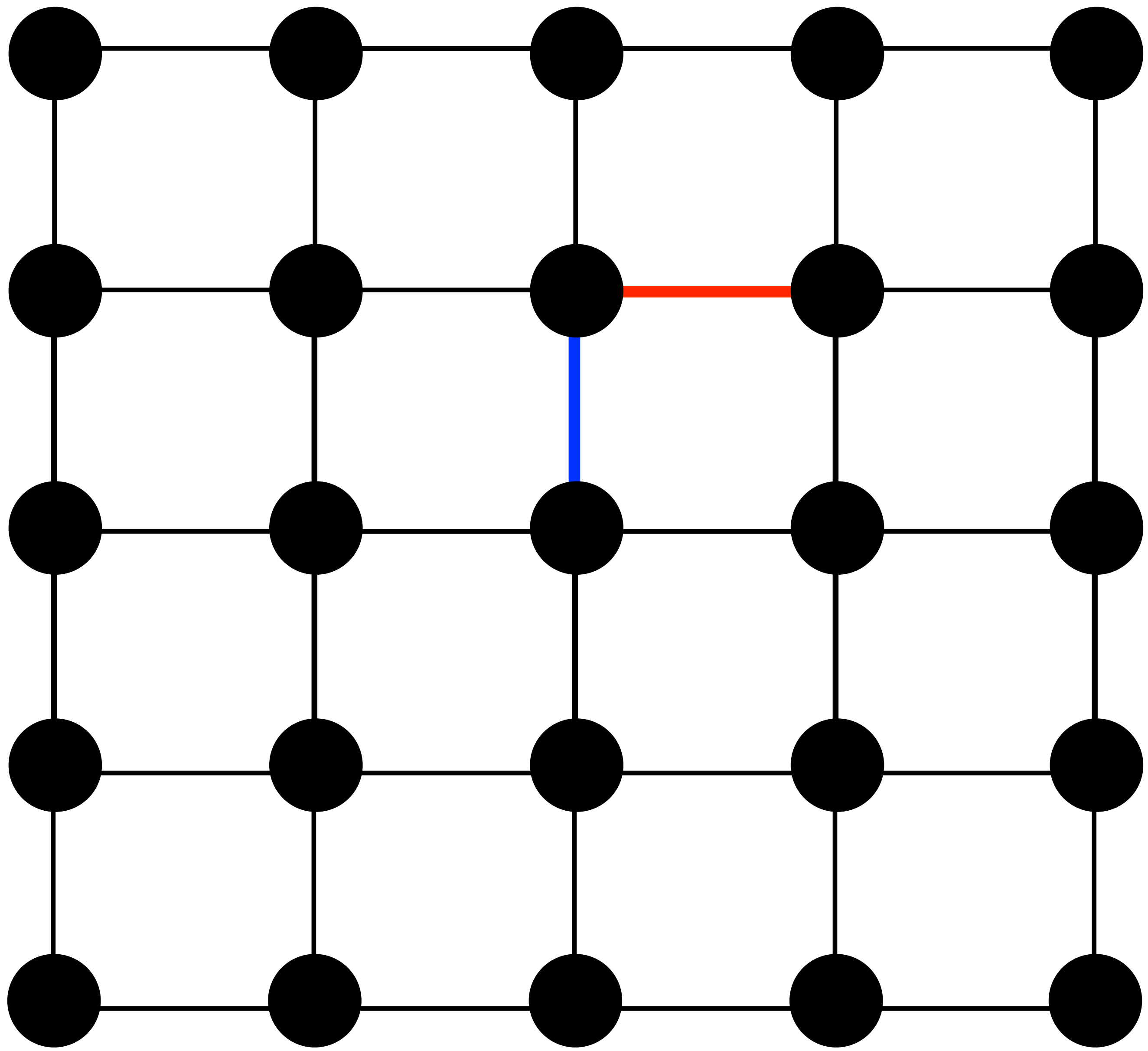
$$J(x) = \sum_{i,j} (\partial_{\rightarrow} x)_{i,j} + (\partial_{\uparrow} x)_{i,j}$$



[Condat 2013]

An Alternative Total Variation

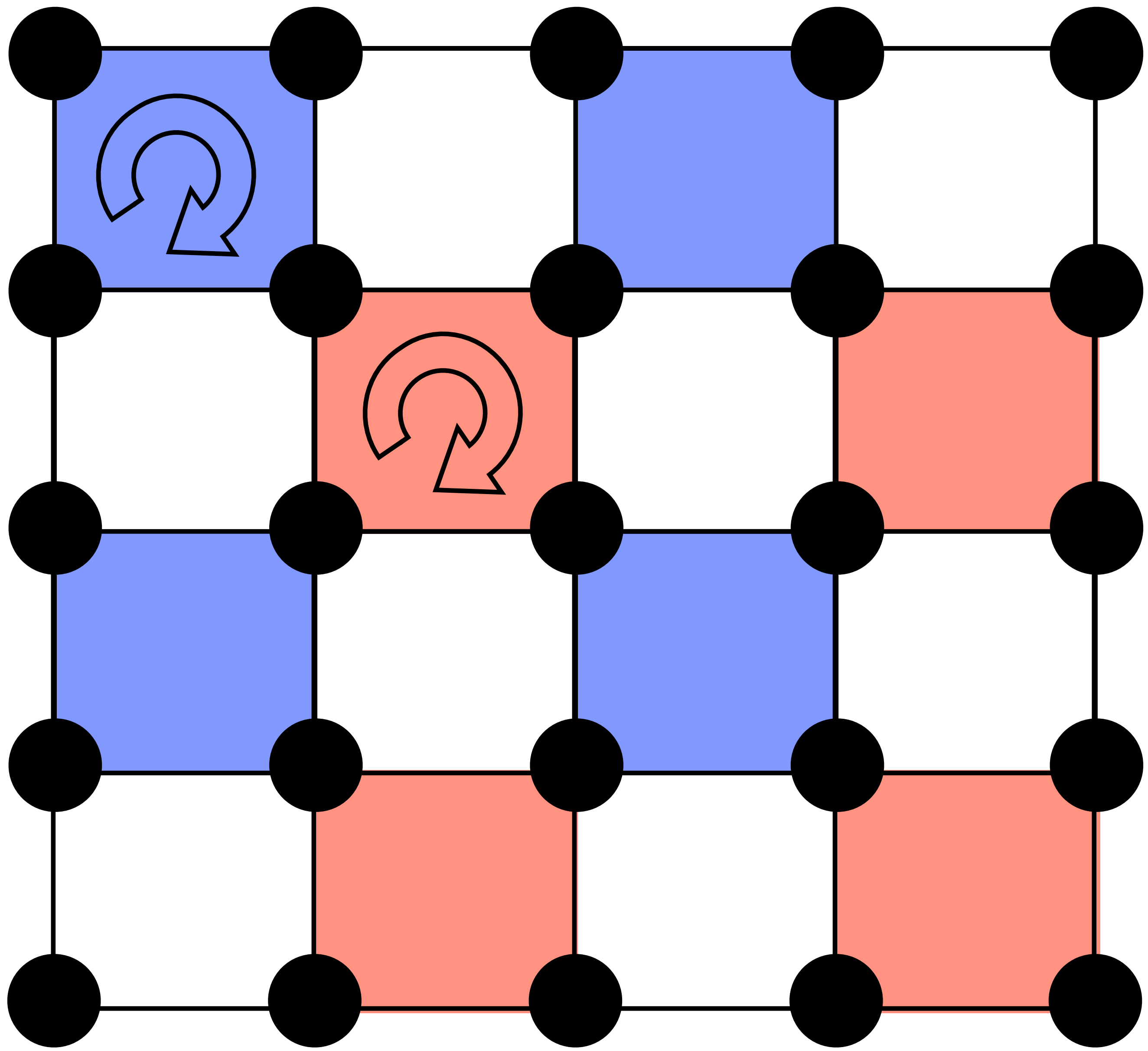
$$J(x) = \sum_{i,j} \sqrt{(\partial_{\rightarrow} x)_{i,j}^2 + (\partial_{\uparrow} x)_{i,j}^2}$$



An Alternative Total Variation

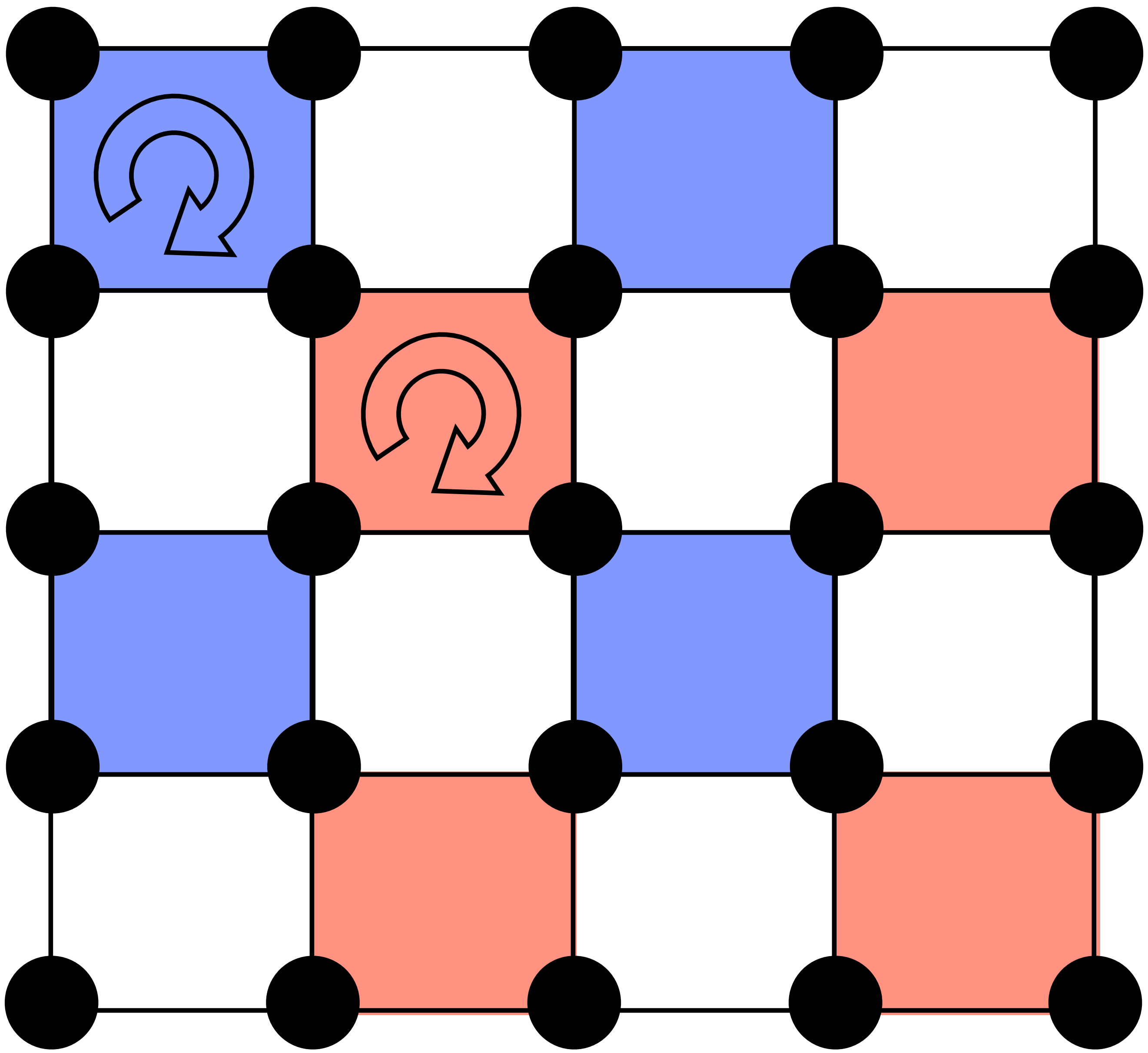
$$J(x) = \sum_{i,j} \mathrm{TV}^4_{2i,2j}(x) + \sum_{i,j} \mathrm{TV}^4_{2i+1,2j+1}(x) \xrightarrow{\Gamma} \text{continuous TV}$$

$$\mathrm{TV}^4_{i,j}(x) = \sqrt{(\partial_{\rightarrow} x)^2_{i,j} + (\partial_{\uparrow} x)^2_{i,j} + (\partial_{\leftarrow} x)^2_{i,j} + (\partial_{\downarrow} x)^2_{i,j}}$$



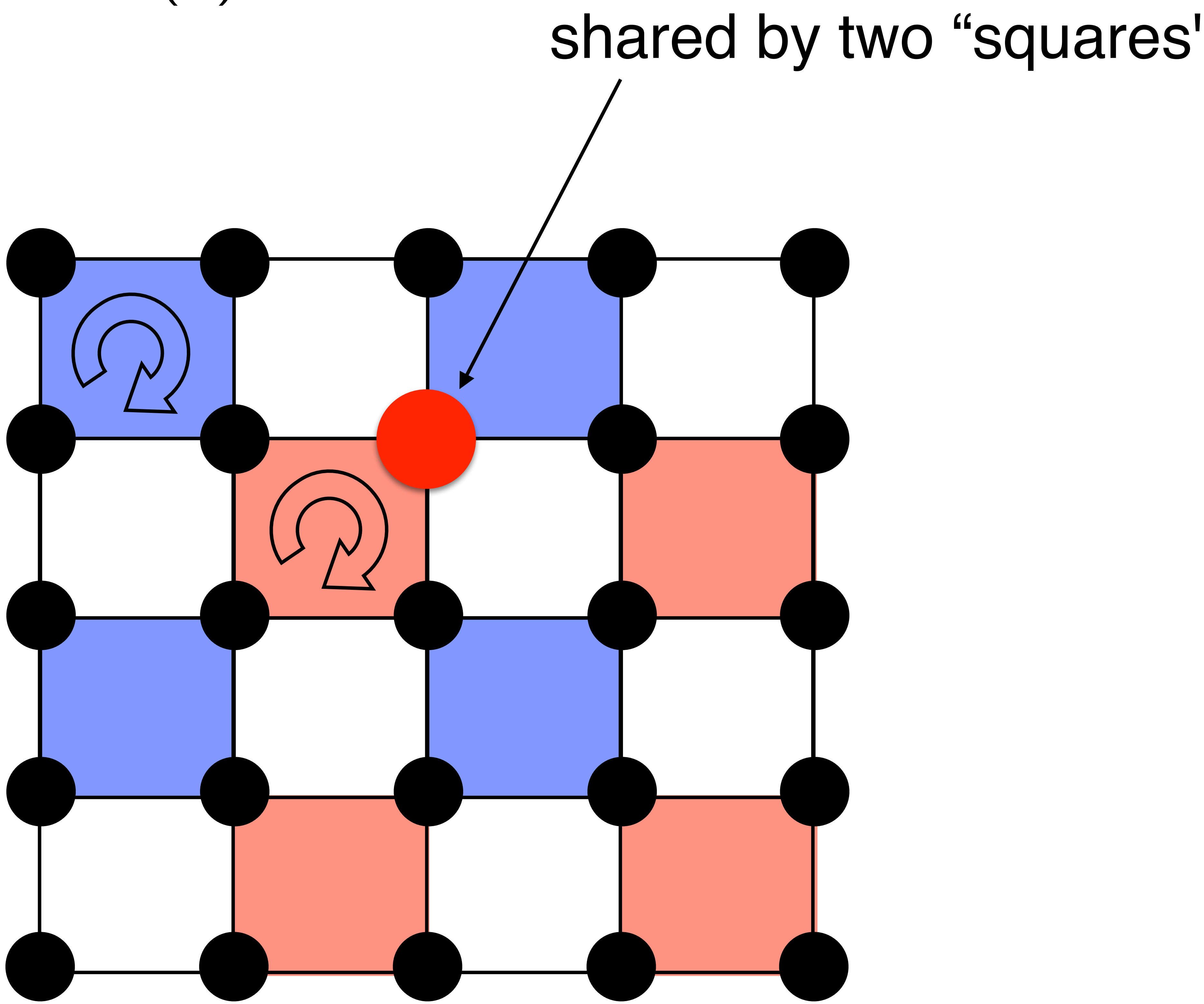
An Alternative Total Variation

$$J(x) = \textcolor{blue}{\mathrm{TV}^e}(x) + \textcolor{red}{\mathrm{TV}^o}(x)$$



An Alternative Total Variation

$$J(x) = \mathsf{TV}^e(x) + \mathsf{TV}^o(x)$$



Dual Problem

$$\mathrm{Prox}_{\lambda J}(x) = \operatorname*{argmin}_{z\in\mathbb{R}^n} \frac{1}{2\lambda}\|x-z\|_2^2 + J(z)$$

$$= \operatorname*{argmin}_{z\in\mathbb{R}^n} \frac{1}{2\lambda}\|x-z\|_2^2 + \mathrm{TV}^e(x) + \mathrm{TV}^o(x)$$

$$\min_{\xi^e,\xi^o} \|D^{o,*}\xi^o + D^{e,*}\xi^e - x\|_2^2 + f(\xi^e) + g(\xi^o)$$

adjoint of the “odd" part of the finite difference

indicator of the inf-unit ball restricted to even coordinates

General Setting

$$\min_{x,y} \|Ax + By\|_2^2 + f(x) + g(y)$$

f, g two proper lsc convex functions (possibily strongly convex)

A, B two linear operators

We will also need:

M, P two positive symmetric operators

K, L two integers ($\geqslant 1$)

Alternating Minimization

For $n > 0$

$$x^{n+1} = \underset{x}{\operatorname{argmin}}\ f(x) + \frac{1}{2}\|Ax + By^n\|_2^2$$

$$y^{n+1} = \underset{y}{\operatorname{argmin}}\ g(y) + \frac{1}{2}\|Ax^{n+1} + By\|_2^2$$

Proposed Scheme

For $n > 0$

$$x_0^{n+1} = x_K^n$$

For $k = 0, \ldots, K-1$

$$x_{k+1}^{n+1} = \operatorname*{argmin}_x f(x) + \frac{1}{2}\|Ax + By^n\|_2^2 + \frac{1}{2}\|x - x_k^{n+1}\|_M^2$$

$$x^{n+1} = \frac{1}{K}\sum_{k=1}^{K} x_k^{n+1}$$

$$y_0^{n+1} = y_L^n$$

For $l = 0, \ldots, L-1$

$$y_{l+1}^{n+1} = \operatorname*{argmin}_y g(y) + \frac{1}{2}\|Ax^{n+1} + By\|_2^2 + \frac{1}{2}\|y - y_l^{n+1}\|_P^2$$

$$y^{n+1} = \frac{1}{L}\sum_{l=1}^{L} y_l^{n+1}$$

Convergence Rates

Naive algorithm	$O(1/N)$
"FISTA" acceleration	$O(1/N^2)$
"FISTA"-acceleration w/ strongly convex	linear rate

Comment on Strongly Convex Case

$$\mathrm{TV}^4_{i,j}(x) = \sqrt{(\partial_{\rightarrow} x)^2_{i,j} + (\partial_{\uparrow} x)^2_{i,j} + (\partial_{\leftarrow} x)^2_{i,j} + (\partial_{\downarrow} x)^2_{i,j}}$$

$\longrightarrow$ not smooth

$$\mathrm{TV}^{4,\varepsilon}_{i,j}(x) = \begin{cases} \mathrm{TV}^4_{i,j}(x) - \varepsilon & \text{if } \mathrm{TV}^4_{i,j}(x) \geqslant 2\varepsilon \\ \frac{1}{4\varepsilon}\left(\mathrm{TV}^4_{i,j}(x)\right)^2 & \text{otherwise} \end{cases}$$

strongly convex

"Huber" version of TV

Performance

ε		FISTA	AAMM	AADM 1	AADM 3	AADM 5	AAMM-inexact 1	AAMM-inexact 3	AAMM-inexact 5
0	#iter	495	146	273	169	153	1654	271	**130**
	t (ms)	681	232	251	228	281	1115	320	**204**
0.1	#iter	142	**57**	89	62	59	513	100	58
	t (ms)	203	91	**89**	91	108	333	130	104
1.0	#iter	69	**27**	57	31	28	174	44	29
	t (ms)	101	**40**	62	50	57	127	54	43

Going on GPGPU (AAMM)

@Xeon E5-2670 / Tesla K20m (linux 3.12, CUDA 5.5)

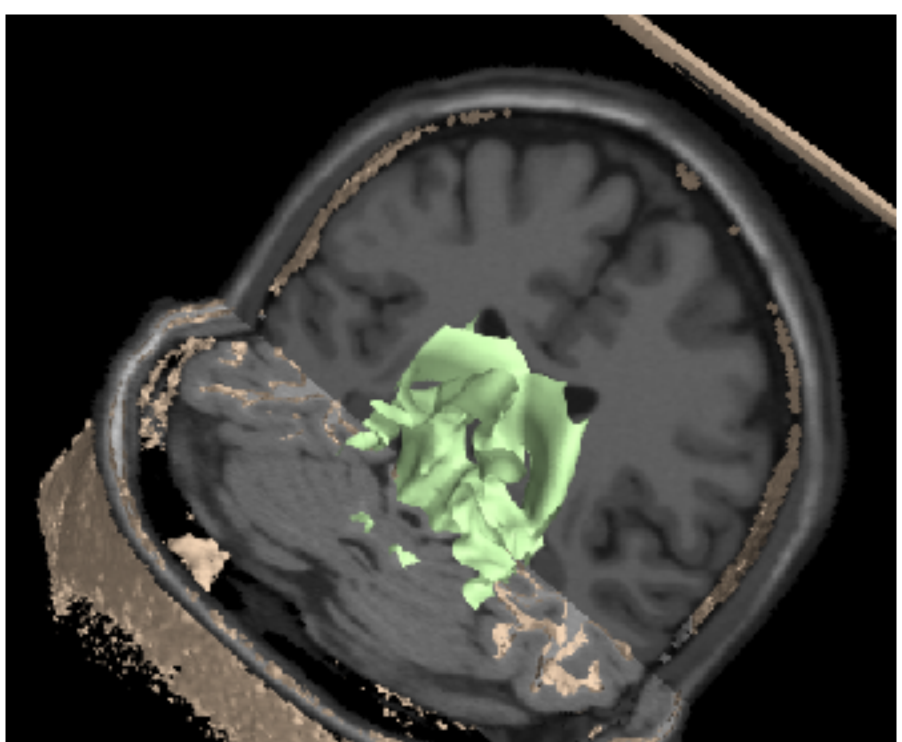
2D

dimension of the problem

parameter

(ms)	256^2	512^2	1024^2	2048^2	4096^2	8192^2
1.0	0.4	0.5	0.9	2.3	7.9	79
5.0	1.8	2.6	4.3	13	63	379
10.0	3.9	5.7	11	39	170	892
20.0	9.3	12	24	92	406	1961

3D



181 x 217 x 181
35 ms

Thanks for your attention!

Accelerated Alternating Descent Methods
for Dykstra-like Problems.
Antonin Chambolle, Pauline Tan and SV.
Journal of Mathematical Imaging and Vision (JMIV). 2017