Reconstruction of piecewise constant signals by the L^1-Potts functional

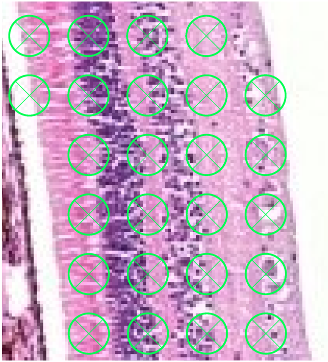
New Trends and Directions in Harmonic Analysis, Fractional Operator Theory, and Image Analysis

Laurent Demaret, Martin Storath, Andreas Weinmann

Institute of Biomathematics - Helmholtz Zentrum München
Supported by the BMBF project “Multimodal Proteome Imaging”

September 20, 2012

Motivation



Retinal tissue from Institute of Pathology, Helmholtz Center Munich

Retina tissue

- High-resolution image of microscopy (purple)
- Low-resolution mass spectrometry image “MALDI” (green)

Goal: Find protein distribution of the retinal tissue

Formulation as one-dimensional problem

Assumption

- Homogeneous mass distribution within retina layers

Idea (Binnig, Schmidt, Schönmeyer, '11)

- Mass distribution only dependent from distance to layer boundary

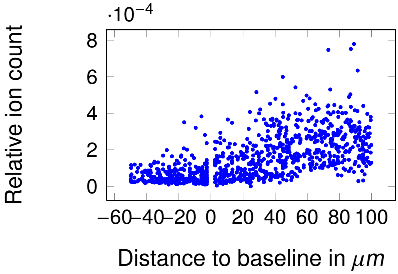
Formulation as one-dimensional problem

Assumption

- Homogeneous mass distribution within retina layers

Idea (Binnig, Schmidt, Schönmeyer, '11)

- Mass distribution only dependent from distance to layer boundary



Reconstruction of true piecewise constant signal from

- Non-Gaussian noise
- Non-equidistant sampling
- Blurred data

Reconstruction of piecewise constant signals by minimization of the L^1-Potts functional

Paradigm: Signal reconstruction ↔ Minimization of cost functional
“regularity + data fidelity”

L^1-Potts functional (Potts ’52, Friedrich et al. ’08)

$$P_\gamma(u) = \gamma \cdot \|\nabla u\|_0 + \|u - f\|_1$$

- Jump penalty $\|\nabla u\|_0$ enforces piecewise constant solutions
- Data term $\|u - f\|_1$ robust also to non-Gaussian noise

Reconstruction of piecewise constant signal
$\rightsquigarrow$
Optimization of the non-convex L^1-Potts functional

$f \in \mathbb{R}^n$ data vector, $\|\nabla u\|_0 = \#\{i : u_{i+1} \neq u_i\}$, $\gamma > 0$ trade-off parameter

Related approaches

L^1-TV (Rudin et al. '92, Fu et. al. '06, Clason et. al. '09, ...)

$$\min_u \ \gamma \cdot \|\nabla u\|_1 + \|u - f\|_1$$

- Penalization of total variation $\|\nabla u\|_1 = \sum_i |u_{i+1} - u_i|$
- Convex functional $\rightsquigarrow$ Convex optimization

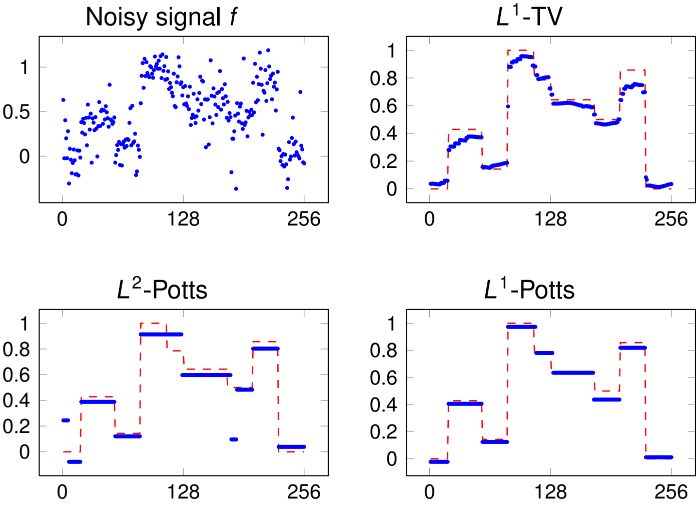
L^2-Potts (Potts '52, Friedrich et al. '08)

$$\min_u \ \gamma \cdot \|\nabla u\|_0 + \|u - f\|_2^2$$

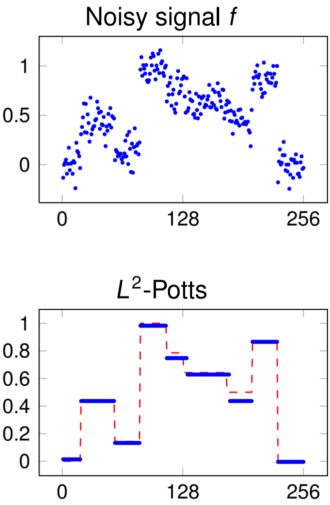
- Data deviation measured in L^2-norm
- Non-convex $\rightsquigarrow$ Dynamic programming

$f \in \mathbb{R}^n$ data vector, $\gamma > 0$ trade-off parameter

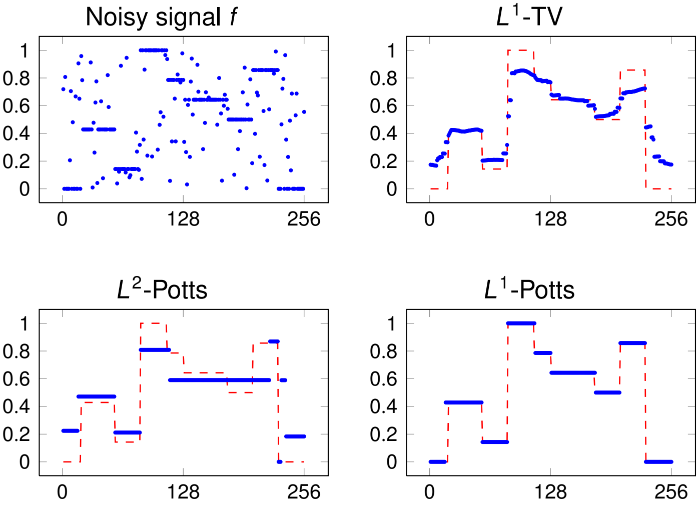
Denoising: Laplacian noise



Denoising: Gaussian noise



Denoising: Salt and pepper noise



Minimization of Potts functionals

Dynamic program for L^2-Potts (Friedrich, Kempe, Liebscher, Winkler, '08)

- Exact solution
- Fast, $O(n^2)$ time and $O(n)$ space

Question (Liebscher '11): Same complexity for L^1-Potts possible?

- L^2 data term $\rightsquigarrow$ mean value (closed formula, cheap)
- L^1 data term $\rightsquigarrow$ median (requires sorting, expensive!)

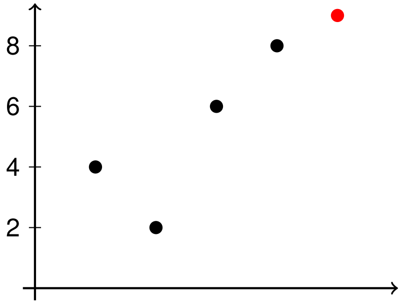
---

n length of data vector

Combinatorial optimization of Potts functionals

Dynamic program:

- Given: Minimizer for data of length $r-1$, $r-2$, ..., 1
- Then: Minimizer for data of length r in r steps

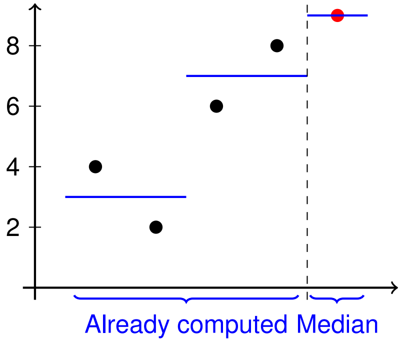


$f \in \mathbb{R}^n$, $P_\gamma(u) = \gamma \cdot \|\nabla u\|_0 + \|u - f\|_1$; here $\gamma = 1$ and $f = (4, 2, 6, 8, 9)$, $n = 5$

Combinatorial optimization of Potts functionals

Dynamic program:

- Given: Minimizer for data of length $r-1$, $r-2$, ..., 1
- Then: Minimizer for data of length r in r steps



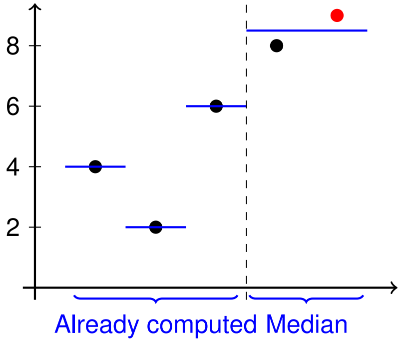
$$P_\gamma = 4 + \gamma + \gamma + 0 = 6$$

$f \in \mathbb{R}^n$, $P_\gamma(u) = \gamma \cdot \|\nabla u\|_0 + \|u - f\|_1$; here $\gamma = 1$ and $f = (4, 2, 6, 8, 9)$, $n = 5$

Combinatorial optimization of Potts functionals

Dynamic program:

- Given: Minimizer for data of length $r-1$, $r-2$, ..., 1
- Then: Minimizer for data of length r in r steps



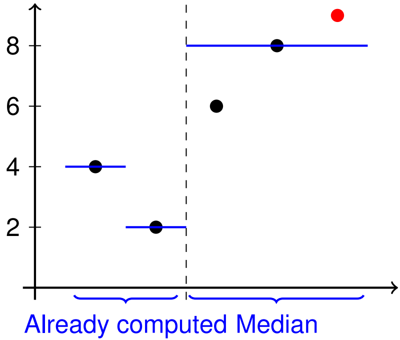
$$P_\gamma = 0 + 2\gamma + {\color{red}\gamma + 1} = 4$$

$f \in \mathbb{R}^n$, $P_\gamma(u) = \gamma \cdot \|\nabla u\|_0 + \|u - f\|_1$; here $\gamma = 1$ and $f = (4, 2, 6, 8, 9)$, $n = 5$

Combinatorial optimization of Potts functionals

Dynamic program:

- Given: Minimizer for data of length $r-1$, $r-2$, ..., 1
- Then: Minimizer for data of length r in r steps



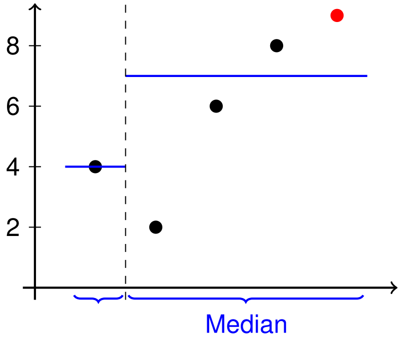
$$P_\gamma = 0 + \gamma + \gamma + 3 = 5$$

$f \in \mathbb{R}^n$, $P_\gamma(u) = \gamma \cdot \|\nabla u\|_0 + \|u - f\|_1$; here $\gamma = 1$ and $f = (4, 2, 6, 8, 9)$, $n = 5$

Combinatorial optimization of Potts functionals

Dynamic program:

- Given: Minimizer for data of length $r-1$, $r-2$, ..., 1
- Then: Minimizer for data of length r in r steps



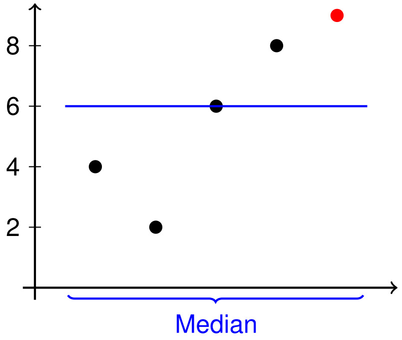
$$P_\gamma = 0 + \gamma + 9 = 10$$

$f \in \mathbb{R}^n$, $P_\gamma(u) = \gamma \cdot \|\nabla u\|_0 + \|u - f\|_1$; here $\gamma = 1$ and $f = (4, 2, 6, 8, 9)$, $n = 5$

Combinatorial optimization of Potts functionals

Dynamic program:

- Given: Minimizer for data of length $r-1$, $r-2$, ..., 1
- Then: Minimizer for data of length r in r steps



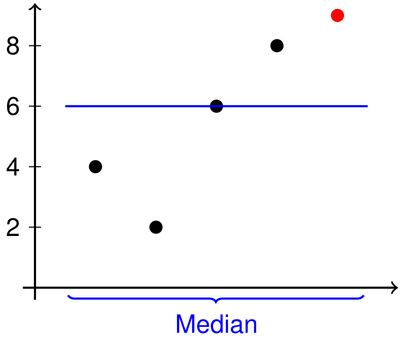
$P_\gamma = 11$

$f \in \mathbb{R}^n$, $P_\gamma(u) = \gamma \cdot \|\nabla u\|_0 + \|u - f\|_1$; here $\gamma = 1$ and $f = (4, 2, 6, 8, 9)$, $n = 5$

Combinatorial optimization of Potts functionals

Dynamic program:

- Given: Minimizer for data of length $r-1$, $r-2$, ..., 1
- Then: Minimizer for data of length r in r steps



$\rightsquigarrow$ For each $r = 1, \ldots, n$:
r median deviations required

$P_\gamma = 11$

$f \in \mathbb{R}^n$, $P_\gamma(u) = \gamma \cdot \|\nabla u\|_0 + \|u - f\|_1$; here $\gamma = 1$ and $f = (4, 2, 6, 8, 9)$, $n = 5$

A new fast L^1-Potts algorithm

Challenge: Compute $O(n^2)$ median deviations in $O(n^2)$ time

Idea

- Given: Median deviation of data $(f_l, ..., f_r)$
- Then: Median deviation of reduced data $(f_{l+1}, ..., f_r)$ in $O(1)$

New data structure: Indexed Linked Histogram

- Histogram values as linked list structure
- Extra indexing "data $f_i \mapsto$ histogram entry"

$\rightsquigarrow$ Reduction step of histogram in $O(1)$

Runtime of our L^1-Potts algorithm

Theorem (Demaret, Storath, Weinmann, '12)

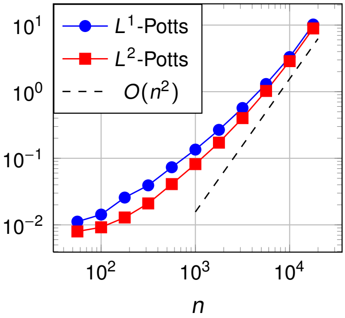
The proposed algorithm computes an exact minimizer of the discrete L^1-Potts functional within $O(n^2)$ time and $O(n)$ space.

Runtime of our L^1-Potts algorithm

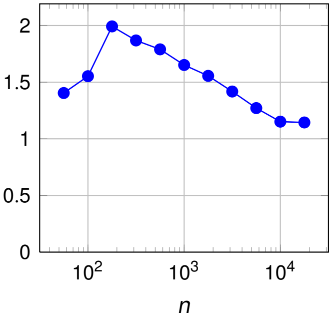
Theorem (Demaret, Storath, Weinmann, ’12)

The proposed algorithm computes an exact minimizer of the discrete L^1-Potts functional within $O(n^2)$ time and $O(n)$ space.

Runtime in seconds



$\frac{\text{Runtime } L^1\text{-Potts}}{\text{Runtime } L^2\text{-Potts}}$

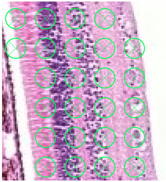


Blurred data

MALDI measurements f are averages over neighborhoods

$$f = K * g,$$

but kernel K unknown.



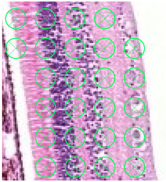
g true piecewise constant function, K convolution kernel

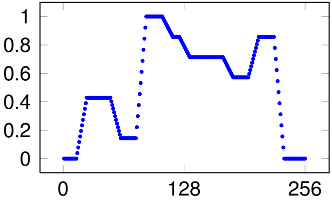
Blurred data

MALDI measurements f are averages over neighborhoods

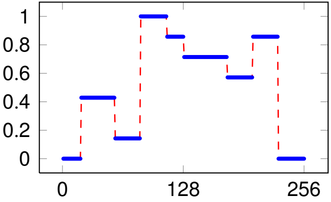
$$f = K * g,$$

but kernel K unknown.





Blurred data (synthetic)



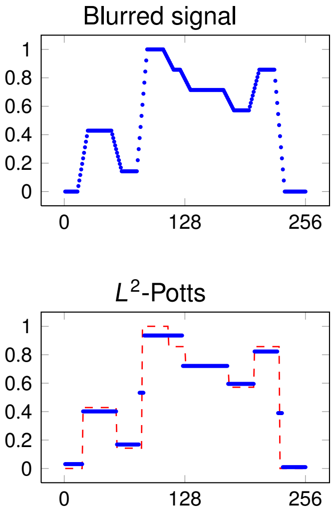
Minimizer of L^1-Potts

L^1-Potts restores the true signal g from the blurred data!

---

g true piecewise constant function, K convolution kernel

Comparison to other approaches



Blind deconvolution property of L^1-Potts

Theorem (Demaret, Storath, Weinmann, '12)

For data of the form

$$f = K * g, \quad g \text{ piecewise constant,}$$

there is a maximal size of the kernel's support and Potts parameters γ, such that g is the unique minimizer of the L^1-Potts functional, i.e.,

$$g = \arg\min_u \gamma \cdot \|\nabla u\|_0 + \|u - f\|_1.$$

K positive symmetric convolution kernel, compactly supported

Blind deconvolution property of L^1-Potts

Theorem (Demaret, Storath, Weinmann, '12)

For data of the form

$$f = K * g, \quad g \text{ piecewise constant,}$$

there is a maximal size of the kernel's support and Potts parameters γ, such that g is the unique minimizer of the L^1-Potts functional, i.e.,

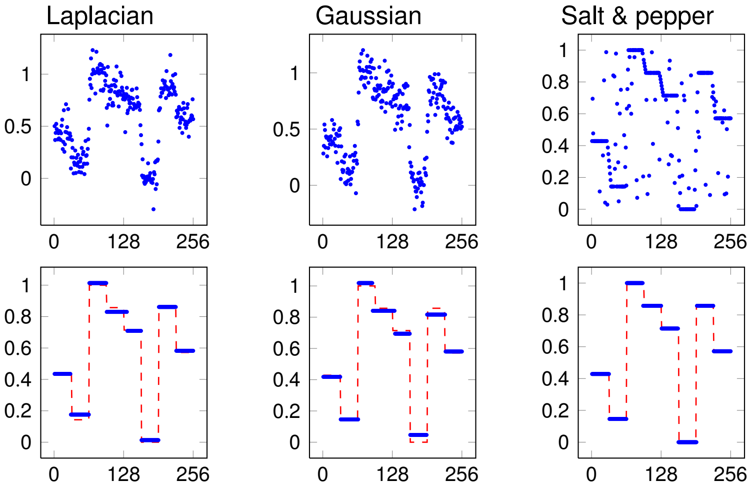
$$g = \arg\min_u \gamma \cdot \|\nabla u\|_0 + \|u - f\|_1.$$

In practice: data f is contaminated by noise

$$f = K * g + \text{noise}$$

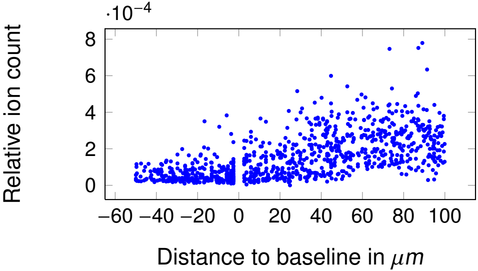
K positive symmetric convolution kernel, compactly supported

Blind deconvolution property under noise



Top: signal blurred by moving average of length 7 plus noise, Bottom: L^1-Potts solution

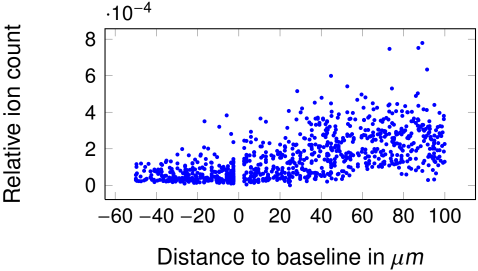
Reminder: Challenges of MALDI data



Reconstruction of true piecewise constant signal from

- Non-Gaussian noise
- Non-equidistant sampling
- Blurred data

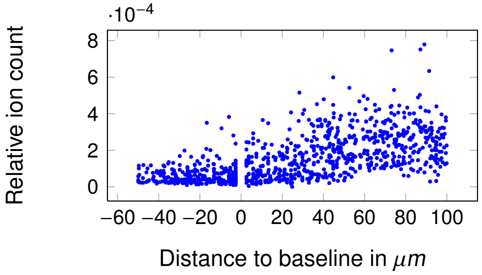
Reminder: Challenges of MALDI data

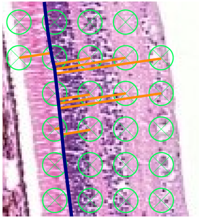


Reconstruction of true piecewise constant signal from

- Non-Gaussian noise ✓
- Non-equidistant sampling
- Blurred data

Reminder: Challenges of MALDI data

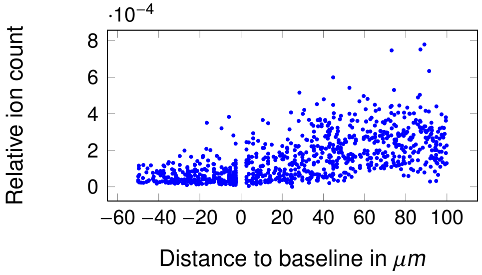


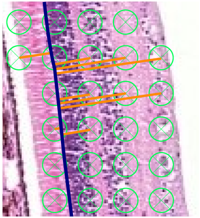


Reconstruction of true piecewise constant signal from

- Non-Gaussian noise ✓
- Non-equidistant sampling ✓
- Blurred data

Reminder: Challenges of MALDI data

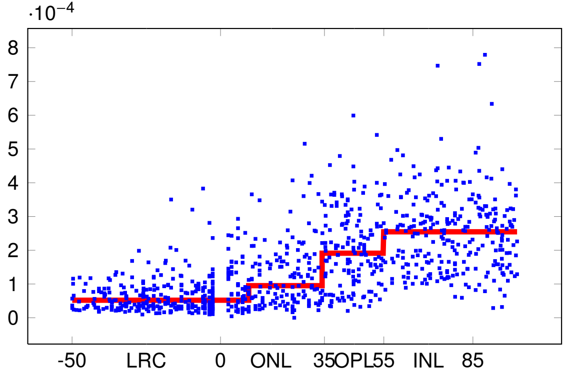


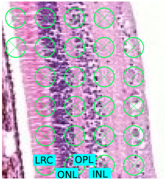


Reconstruction of true piecewise constant signal from

- Non-Gaussian noise ✓
- Non-equidistant sampling ✓
- Blurred data ✓

L^1-Potts to MALDI data





Jumps of L^1-Potts reconstruction correlate with layer boundaries
$\rightsquigarrow$ Confidence in MALDI measurements

LRC: Layer of rods and cones, ONL: Outer nuclear layer, OPL: Outer plexiform layer, INL: inner nuclear layer

Summary and references

Summary

- L^1-Potts for reconstruction of piecewise constant signal under various types of noise
- Deblurring without knowledge of the blurring operator
- Fast and exact algorithm

Summary and references

Summary

- L^1-Potts for reconstruction of piecewise constant signal under various types of noise
- Deblurring without knowledge of the blurring operator
- Fast and exact algorithm

Reference

- L. Demaret, M. Storath, A. Weinmann. *Reconstruction of piecewise constant signals by minimization of the L^1-Potts functional.* 2012. (arXiv preprint)
- Potts Toolbox for Matlab, http://purl.org/potts/toolbox (open source)

Further references

F. Friedrich, A. Kempe, V. Liebscher, and G. Winkler.
Complexity penalized M-Estimation.
Journal of Computational and Graphical Statistics, 17(1):201–224, March 2008.

L. Boysen, A. Kempe, V. Liebscher, A. Munk, and O. Wittich.
Consistencies and rates of convergence of jump-penalized least squares estimators.
The Annals of Statistics, 37(1):157–183, 2009.

C. Clason, B. Jin, and K. Kunisch.
A duality-based splitting method for l-tv image restoration with automatic regularization parameter choice.
SIAM J. Sci. Comput. v32 i3, pages 1484–1505, 2009.
http://math.uni-graz.at/optcon/projects/clason3/.

Sampling

- Integral sampling

$$f^k(j) = S_k f(j) = \frac{1}{|I_j|} \int_{I_j} f \, \mathrm{d}\lambda,$$

where I_j are the intervals given by the discretization sets X_k.

- Point sampling

$$f^k(j) = S_k f(j) = f(x_j)$$

where x_j might be taken as the midpoint of the interval I_j.

- Discrete Potts functionals

$$P_\gamma^k(u) = \begin{cases} \gamma \cdot J(u) + \left\| u - \sum_j S_k f(j) \cdot \mathbf{1}_{I_j} \right\|_1, & \text{if } u \in \mathrm{PC}^k[0,1], \\ \infty, & \text{else.} \end{cases}$$

Γ-Convergence

A sequence F^k of functionals on $L^1[0,1]$ taking values in the extended positive real line $[0,\infty]$, Γ*-converges* to a functional F_∞ if

(i) for each $u \in L^1[0,1]$ and each sequence $\{u^k\}_{k\in\mathbb{N}}$ in $L^1[0,1]$ with $u^k \to u$ as $k \to \infty$ holds

$$F^\infty(u) \le \liminf_{k\to\infty} F^k(u^k), \tag{1}$$

(ii) and for each $u \in L^1[0,1]$ there is a *recovery sequence* $\{u^k\}_{k\in\mathbb{N}}$ in $L^1[0,1]$ with $u^k \to u$ as $k \to \infty$ such that

$$F^\infty(u) \ge \limsup_{k\to\infty} F^k(u^k). \tag{2}$$

Γ-convergence to continuous L^1-Potts model

Continuous L^1-Potts functional

$$P_\gamma(u) = \begin{cases} \gamma \cdot \|\nabla u\|_0 + \|u - f\|_1\,, & \text{if } u \in \mathrm{PC}[0,1] \\ \infty, & \text{else.} \end{cases}$$

Interpretation of discrete data $(f_i)_i$ as sampling of the data $f \in L^1$.

$f \in L^1$ data, $\mathrm{PC}[0,1]$ piecewise constant functions on $[0,1]$, $\|\nabla u\|_0$ number of jumps of u

Γ-convergence to continuous L^1-Potts model

Continuous L^1-Potts functional

$$P_\gamma(u) = \begin{cases} \gamma \cdot \|\nabla u\|_0 + \|u - f\|_1 , & \text{if } u \in \mathrm{PC}[0,1] \\ \infty, & \text{else.} \end{cases}$$

Interpretation of discrete data $(f_i)_i$ as sampling of the data $f \in L^1$.

Theorem (Demaret, Storath, Weinmann, '12)

- *The discrete functionals Γ-converge to the continuous L^1-Potts functional P_γ as the sampling density approaches* 0.
- *Each accumulation point of a sequence of minimizers of the discrete approximations is a minimizer of the continuous model.*

$f \in L^1$ data, $\mathrm{PC}[0,1]$ piecewise constant functions on $[0,1]$, $\|\nabla u\|_0$ number of jumps of u

Reconstruction from severely blurred data

Convolution kernel K known $\rightsquigarrow$ “Deconvolution L^1-Potts functional”

$$\gamma \cdot \|\nabla u\|_0 + \|K * u - f\|_1 \to \min$$

K convolution kernel

Reconstruction from severely blurred data

Convolution kernel K known $\rightsquigarrow$ "Deconvolution L^1-Potts functional"

$$\gamma \cdot \|\nabla u\|_0 + \|K * u - f\|_1 \to \min$$

Splitting approach

$$\min_{u,v} \gamma \cdot \|\nabla u\|_0 + \mu \cdot \|u - v\|_1 + \|K * v - f\|_1.$$

K convolution kernel

Reconstruction from severely blurred data

Convolution kernel K known $\rightsquigarrow$ “Deconvolution L^1-Potts functional”

$$\gamma \cdot \|\nabla u\|_0 + \|K * u - f\|_1 \to \min$$

Splitting approach

$$\min_{u,v} \gamma \cdot \|\nabla u\|_0 + \mu \cdot \|u - v\|_1 + \|K * v - f\|_1 .$$

Heuristic algorithm: alternately solve

$$\min_{u} \frac{\gamma}{\mu} \cdot \|\nabla u\|_0 + \|u - v\|_1 ,$$

K convolution kernel

Reconstruction from severely blurred data

Convolution kernel K known $\rightsquigarrow$ "Deconvolution L^1-Potts functional"

$$\gamma \cdot \|\nabla u\|_0 + \|K * u - f\|_1 \to \min$$

Splitting approach

$$\min_{u,v} \gamma \cdot \|\nabla u\|_0 + \mu \cdot \|u - v\|_1 + \|K * v - f\|_1.$$

Heuristic algorithm: alternately solve

$$\min_{u} \frac{\gamma}{\mu} \cdot \|\nabla u\|_0 + \|u - v\|_1,$$
$$\min_{v} \mu \cdot \|u - v\|_1 + \|K * u - f\|_1,$$

K convolution kernel

Reconstruction from severely blurred data

Convolution kernel K known $\rightsquigarrow$ “Deconvolution L^1-Potts functional”

$$\gamma \cdot \|\nabla u\|_0 + \|K * u - f\|_1 \to \min$$

Splitting approach

$$\min_{u,v} \gamma \cdot \|\nabla u\|_0 + \mu \cdot \|u - v\|_1 + \|K * v - f\|_1 .$$

Heuristic algorithm: alternately solve

$$\min_u \frac{\gamma}{\mu} \cdot \|\nabla u\|_0 + \|u - v\|_1 ,$$
$$\min_v \mu \cdot \|u - v\|_1 + \|K * u - f\|_1 ,$$

and increase coupling parameter μ in each step.

K convolution kernel

Deconvolution by the iterative L^1-Potts