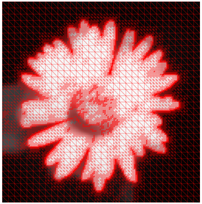
  

A posteriori error control for the binary Mumford–Shah model

Benjamin Berkels[1], Alexander Effland[2], Martin Rumpf[2]

[1]AICES Graduate School, RWTH Aachen University, Germany

[2]Institute for Numerical Simulation, University of Bonn

Minisymposium on Recent Advances in Convex Relaxations
SIAM Conference on IMAGING SCIENCE
May 23rd – 26th, Hotel Albuquerque at Old Town

The binary Mumford–Shah model

u_0

$\mathcal{O} \subset \Omega$

Binary Mumford–Shah model $\theta_1, \theta_2 \in L^1(\Omega)$

$$E[\mathcal{O}] = \int_{\mathcal{O}} \theta_1 \, \mathrm{d}x + \int_{\Omega \setminus \mathcal{O}} \theta_2 \, \mathrm{d}x + \mathrm{Per}[\mathcal{O}]$$

The binary Mumford–Shah model

u_0 $\qquad$ $\chi \in BV(\Omega, \{0,1\})$

Binary Mumford–Shah model $\quad \theta_1, \theta_2 \in L^1(\Omega)$

$$E[\chi] = \int_\Omega \theta_1 \chi + \theta_2 (1-\chi)\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega)$$

The binary Mumford–Shah model

u_0

$\chi \in BV(\Omega, \{0, 1\})$

Binary Mumford–Shah model $\theta_1, \theta_2 \in L^1(\Omega)$

$$E[\chi, c_1, c_2] = \int_\Omega \theta_1 \chi + \theta_2 (1 - \chi) \, \mathrm{d}x + |\mathrm{D}\chi|(\Omega)$$

- $\theta_i = \frac{1}{\nu}(c_i - u_0)^2$ $(i = 1, 2)$ for weight parameter $\nu > 0$

The binary Mumford–Shah model

u_0

$\chi \in BV(\Omega, \{0, 1\})$

Binary Mumford–Shah model $\theta_1, \theta_2 \in L^1(\Omega)$

$$E[\chi, c_1, c_2] = \int_\Omega \theta_1 \chi + \theta_2 (1 - \chi)\, \mathrm{d}x + |\mathrm{D}\chi|(\Omega)$$

- $\theta_i = \frac{1}{\nu}(c_i - u_0)^2$ $(i = 1, 2)$ for weight parameter $\nu > 0$
- $c_1 = (\int_\Omega \chi \, \mathrm{d}x)^{-1} \int_\Omega \chi u_0 \, \mathrm{d}x$, $c_2 = (\int_\Omega 1 - \chi \, \mathrm{d}x)^{-1} \int_\Omega (1 - \chi) u_0 \, \mathrm{d}x$

The binary Mumford–Shah model (cont.)

$$E[\chi, c_1, c_2] = \int_\Omega \theta_1 \chi + \theta_2 (1 - \chi)\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega)$$

minimize E over the set $\chi \in BV(\Omega, \{0,1\})$ (c_1, c_2 fixed)

The binary Mumford–Shah model (cont.)

$$E[\chi, c_1, c_2] = \int_\Omega \theta_1 \chi + \theta_2 (1 - \chi) \, \mathrm{d}x + |\mathrm{D}\chi|(\Omega)$$

minimize E over the set $\chi \in BV(\Omega, \{0, 1\})$ (c_1, c_2 fixed)

Goal: derive functional a posteriori error estimates for

$$\|\chi - \chi_h\|_{L^1(\Omega)}$$

The binary Mumford–Shah model (cont.)

$$E[\chi, c_1, c_2] = \int_\Omega \theta_1 \chi + \theta_2 (1-\chi)\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega)$$

minimize E over the set $\chi \in BV(\Omega, \{0,1\})$ (c_1, c_2 fixed)

Goal: derive functional a posteriori error estimates for

$$\|\chi - \chi_h\|_{L^1(\Omega)}$$

Major problems:

- optimization over a non convex set

The binary Mumford–Shah model (cont.)

$$E[\chi, c_1, c_2] = \int_\Omega \theta_1 \chi + \theta_2(1-\chi)\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega)$$

minimize E over the set $\chi \in BV(\Omega, \{0,1\})$ (c_1, c_2 fixed)

Goal: derive functional a posteriori error estimates for

$$\|\chi - \chi_h\|_{L^1(\Omega)}$$

Major problems:

- optimization over a non convex set
- no uniform convexity of E

The binary Mumford–Shah model (cont.)

$$E[\chi, c_1, c_2] = \int_\Omega \theta_1 \chi + \theta_2 (1-\chi)\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega)$$

minimize E over the set $\chi \in BV(\Omega, \{0,1\})$ (c_1, c_2 fixed)

Goal: derive functional a posteriori error estimates for

$$\|\chi - \chi_h\|_{L^1(\Omega)}$$

Major problems:

- optimization over a non convex set
- no uniform convexity of E

Strategy:

- use a suitable uniformly convex relaxation

The binary Mumford–Shah model (cont.)

$$E[\chi, c_1, c_2] = \int_\Omega \theta_1 \chi + \theta_2 (1 - \chi)\, \mathrm{d}x + |\mathrm{D}\chi|(\Omega)$$

minimize E over the set $\chi \in BV(\Omega, \{0,1\})$ (c_1, c_2 fixed)

Goal: derive functional a posteriori error estimates for

$$\|\chi - \chi_h\|_{L^1(\Omega)}$$

Major problems:

- optimization over a non convex set
- no uniform convexity of E

Strategy:

- use a suitable uniformly convex relaxation
- error estimate for the relaxation based on the duality gap

The binary Mumford–Shah model (cont.)

$$E[\chi, c_1, c_2] = \int_\Omega \theta_1 \chi + \theta_2(1-\chi)\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega)$$

minimize E over the set $\chi \in BV(\Omega, \{0,1\})$ (c_1, c_2 fixed)

Goal: derive functional a posteriori error estimates for

$$\|\chi - \chi_h\|_{L^1(\Omega)}$$

Major problems:

- optimization over a non convex set
- no uniform convexity of E

Strategy:

- use a suitable uniformly convex relaxation
- error estimate for the relaxation based on the duality gap
- thresholding of the relaxed solution and cut out argument

Convex relaxation and thresholding - CEN

Idea Relax range of χ to $[0,1]$ [Nikolova, Esedoḡlu, Chan '06]

$$E_{\mathsf{CEN}}[u] = \int_\Omega u\theta_1 \,\mathrm{d}x + \int_\Omega (1-u)\,\theta_2 \,\mathrm{d}x + |\mathrm{D}u|(\Omega).$$

Convex relaxation and thresholding - CEN

Idea Relax range of χ to $[0,1]$ [Nikolova, Esedoḡlu, Chan '06]

$$E_{\mathsf{CEN}}[u] = \int_\Omega u\theta_1 \, \mathrm{d}x + \int_\Omega (1-u)\,\theta_2 \, \mathrm{d}x + |\mathrm{D}u|(\Omega).$$

Theorem

$$u^* \in \operatorname*{argmin}_{u \in BV(\Omega,[0,1])} E_{\mathsf{CEN}}[u] \quad \Rightarrow \quad [u^* > s] \in \operatorname*{argmin}_{\chi \in BV(\Omega,\{0,1\})} E[\chi]$$

for all a.e. $s \in [0,1]$,

where $[u > s]$ is the s-superlevel set of u, i.e. $\{x : u(x) > s\}$.

Convex relaxation and thresholding - CEN

Idea Relax range of χ to $[0,1]$ [Nikolova, Esedoḡlu, Chan '06]

$$E_{\mathsf{CEN}}[u] = \int_\Omega u\theta_1 \, \mathrm{d}x + \int_\Omega (1-u)\,\theta_2 \, \mathrm{d}x + |\mathrm{D}u|(\Omega).$$

Theorem

$$u^* \in \operatorname*{argmin}_{u \in BV(\Omega,[0,1])} E_{\mathsf{CEN}}[u] \quad \Rightarrow \quad [u^* > s] \in \operatorname*{argmin}_{\chi \in BV(\Omega,\{0,1\})} E[\chi]$$

for all a.e. $s \in [0,1]$,

where $[u > s]$ is the s-superlevel set of u, i.e. $\{x : u(x) > s\}$.

- ✔ $BV(\Omega,[0,1])$ convex
- ✘ E_{CEN} convex, but not uniformly convex in u

Uniformly convex approach

$$E^{\mathsf{rel}}[u] = \int_\Omega u^2 \theta_1 \,\mathrm{d}x + \int_\Omega (1-u)^2\, \theta_2 \,\mathrm{d}x + |\mathrm{D}u|(\Omega)$$

Convex relaxation and thresholding - UC

Uniformly convex approach

$$E^{\text{rel}}[u] = \int_\Omega u^2 \theta_1 \, \mathrm{d}x + \int_\Omega (1-u)^2 \, \theta_2 \, \mathrm{d}x + |\mathrm{D}u|(\Omega)$$

Theorem

[Chambolle/Darbon '08][B. SSVM '09]

Let $\theta_1, \theta_2 \in L^1(\Omega)$, $\theta_1, \theta_2 \geq 0$ a. e.. Then,

- E^{rel} has a unique minimizer on $BV(\Omega)$.
- $u = \underset{v \in BV(\Omega)}{\operatorname{argmin}} E^{\text{rel}}[v] \Rightarrow \chi_{[u>0.5]} \in \underset{\chi \in BV(\Omega, \{0,1\})}{\operatorname{argmin}} E[\chi]$.
  Moreover, $u(x) \in [0,1]$ for a.e. $x \in \Omega$.

Convex relaxation and thresholding - UC

Uniformly convex approach

$$E^{\mathrm{rel}}[u] = \int_\Omega u^2 \theta_1 \,\mathrm{d}x + \int_\Omega (1-u)^2 \,\theta_2 \,\mathrm{d}x + |\mathrm{D}u|(\Omega)$$

Theorem

[Chambolle/Darbon '08][B. SSVM '09]

Let $\theta_1, \theta_2 \in L^1(\Omega)$, $\theta_1, \theta_2 \geq 0$ a. e.. Then,

- E^{rel} has a unique minimizer on $BV(\Omega)$.
- $u = \underset{v \in BV(\Omega)}{\operatorname{argmin}} E^{\mathrm{rel}}[v] \Rightarrow \chi_{[u>0.5]} \in \underset{\chi \in BV(\Omega,\{0,1\})}{\operatorname{argmin}} E[\chi]$.
  Moreover, $u(x) \in [0,1]$ for a.e. $x \in \Omega$.

✔ $BV(\Omega)$ convex (even unconstrained)

✔ E_{UC} uniformly convex in u

Convex relaxation and thresholding - UC (cont.)

Proof sketch: [Chambolle/Darbon '09]

- $E^{\text{rel}}[\min\{\max\{0,v\},1\}] \leq E^{\text{rel}}[v]$ for any $v \in BV(\Omega)$.

Convex relaxation and thresholding - UC (cont.)

Proof sketch: [Chambolle/Darbon '09]

- $E^{\text{rel}}[\min\{\max\{0,v\},1\}] \le E^{\text{rel}}[v]$ for any $v \in BV(\Omega)$.
- $\Psi(x,t) := t^2\theta_1(x) + (1-t)^2\theta_2(x)$, $C_\Psi := \int_\Omega \Psi(x,0)\,\mathrm{d}x$.

$$\int_\Omega \Psi(x,u)\,\mathrm{d}x = \int_\Omega \Psi(x,0)\,\mathrm{d}x + \int_0^1 \int_\Omega \partial_t \Psi(x,s)\chi_{[u>s]}\,\mathrm{d}x\,\mathrm{d}s$$

Convex relaxation and thresholding - UC (cont.)

Proof sketch: [Chambolle/Darbon '09]

- $E^{\mathrm{rel}}[\min\{\max\{0,v\},1\}] \leq E^{\mathrm{rel}}[v]$ for any $v \in BV(\Omega)$.
- $\Psi(x,t) := t^2\theta_1(x) + (1-t)^2\theta_2(x)$, $C_\Psi := \int_\Omega \Psi(x,0)\,\mathrm{d}x$.

$$\int_\Omega \Psi(x,u)\,\mathrm{d}x = \int_\Omega \Psi(x,0)\,\mathrm{d}x + \int_0^1 \int_\Omega \partial_t \Psi(x,s)\chi_{[u>s]}\,\mathrm{d}x\,\mathrm{d}s$$

- $h_1 \leq h_2$ a. e., $\chi_i \in \underset{\chi \in BV(\Omega,\{0,1\})}{\operatorname{argmin}} \int_\Omega h_i\chi\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega)$

Convex relaxation and thresholding - UC (cont.)

Proof sketch: [Chambolle/Darbon '09]

- $E^{\text{rel}}[\min\{\max\{0,v\},1\}] \leq E^{\text{rel}}[v]$ for any $v \in BV(\Omega)$.
- $\Psi(x,t) := t^2\theta_1(x) + (1-t)^2\theta_2(x)$, $C_\Psi := \int_\Omega \Psi(x,0)\,\mathrm{d}x$.

$$\int_\Omega \Psi(x,u)\,\mathrm{d}x = \int_\Omega \Psi(x,0)\,\mathrm{d}x + \int_0^1 \int_\Omega \partial_t \Psi(x,s)\chi_{[u>s]}\,\mathrm{d}x\,\mathrm{d}s$$

- $h_1 \leq h_2$ a. e., $\chi_i \in \underset{\chi \in BV(\Omega,\{0,1\})}{\operatorname{argmin}} \int_\Omega h_i \chi\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega) \Rightarrow \chi_1 \geq \chi_2$

Convex relaxation and thresholding - UC (cont.)

Proof sketch: [Chambolle/Darbon '09]

- $E^{\text{rel}}[\min\{\max\{0,v\},1\}] \leq E^{\text{rel}}[v]$ for any $v \in BV(\Omega)$.
- $\Psi(x,t) := t^2\theta_1(x) + (1-t)^2\theta_2(x)$, $C_\Psi := \int_\Omega \Psi(x,0)\,\mathrm{d}x$.

$$\int_\Omega \Psi(x,u)\,\mathrm{d}x = \int_\Omega \Psi(x,0)\,\mathrm{d}x + \int_0^1 \int_\Omega \partial_t \Psi(x,s)\chi_{[u>s]}\,\mathrm{d}x\,\mathrm{d}s$$

- $h_1 \leq h_2$ a. e., $\chi_i \in \underset{\chi\in BV(\Omega,\{0,1\})}{\operatorname{argmin}} \int_\Omega h_i\chi\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega) \Rightarrow \chi_1 \geq \chi_2$
- $\chi^s \in \underset{\chi\in BV(\Omega,\{0,1\})}{\operatorname{argmin}} \left\{E_s^{\text{rel}}[\chi] := \int_\Omega \partial_s\Psi(x,s)\chi\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega)\right\}$

Convex relaxation and thresholding - UC (cont.)

Proof sketch:

[Chambolle/Darbon '09]

- $E^{\mathsf{rel}}[\min\{\max\{0,v\},1\}] \leq E^{\mathsf{rel}}[v]$ for any $v \in BV(\Omega)$.
- $\Psi(x,t) := t^2\theta_1(x) + (1-t)^2\theta_2(x)$, $C_\Psi := \int_\Omega \Psi(x,0)\,\mathrm{d}x$.

$$\int_\Omega \Psi(x,u)\,\mathrm{d}x = \int_\Omega \Psi(x,0)\,\mathrm{d}x + \int_0^1 \int_\Omega \partial_t \Psi(x,s)\chi_{[u>s]}\,\mathrm{d}x\,\mathrm{d}s$$

- $h_1 \leq h_2$ a. e., $\chi_i \in \underset{\chi \in BV(\Omega,\{0,1\})}{\operatorname{argmin}} \int_\Omega h_i\chi\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega) \Rightarrow \chi_1 \geq \chi_2$
- $\chi^s \in \underset{\chi \in BV(\Omega,\{0,1\})}{\operatorname{argmin}} \left\{E_s^{\mathsf{rel}}[\chi] := \int_\Omega \partial_s\Psi(x,s)\chi\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega)\right\}$
- $\Psi(x,\cdot)$ strictly convex $\Rightarrow \partial_t\Psi(x,\cdot)$ is strictly increasing $\Rightarrow \chi^{s_1} \geq \chi^{s_2}$ a. e. for $s_1 \leq s_2$.

Convex relaxation and thresholding - UC (cont.)

Proof sketch:

[Chambolle/Darbon '09]

- $E^{\text{rel}}[\min\{\max\{0,v\},1\}] \le E^{\text{rel}}[v]$ for any $v \in BV(\Omega)$.
- $\Psi(x,t) := t^2\theta_1(x) + (1-t)^2\theta_2(x)$, $C_\Psi := \int_\Omega \Psi(x,0)\,\mathrm{d}x$.

$$\int_\Omega \Psi(x,u)\,\mathrm{d}x = \int_\Omega \Psi(x,0)\,\mathrm{d}x + \int_0^1 \int_\Omega \partial_t \Psi(x,s)\chi_{[u>s]}\,\mathrm{d}x\,\mathrm{d}s$$

- $h_1 \le h_2$ a. e., $\chi_i \in \operatorname*{argmin}_{\chi \in BV(\Omega,\{0,1\})} \int_\Omega h_i \chi\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega) \Rightarrow \chi_1 \ge \chi_2$
- $\chi^s \in \operatorname*{argmin}_{\chi \in BV(\Omega,\{0,1\})} \left\{E_s^{\text{rel}}[\chi] := \int_\Omega \partial_s \Psi(x,s)\chi\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega)\right\}$
- $\Psi(x,\cdot)$ strictly convex $\Rightarrow \partial_t\Psi(x,\cdot)$ is strictly increasing $\Rightarrow \chi^{s_1} \ge \chi^{s_2}$ a. e. for $s_1 \le s_2$.
- graph reconstruction $u^*(x) = \sup\{s \,|\, \chi^s(x) = 1\}$ is well-defined

Convex relaxation and thresholding - UC (cont.)

Proof sketch: [Chambolle/Darbon '09]

- $E^{\mathrm{rel}}[\min\{\max\{0,v\},1\}] \leq E^{\mathrm{rel}}[v]$ for any $v \in BV(\Omega)$.
- $\Psi(x,t) := t^2\theta_1(x) + (1-t)^2\theta_2(x)$, $C_\Psi := \int_\Omega \Psi(x,0)\,\mathrm{d}x$.

$$\int_\Omega \Psi(x,u)\,\mathrm{d}x = \int_\Omega \Psi(x,0)\,\mathrm{d}x + \int_0^1 \int_\Omega \partial_t \Psi(x,s)\chi_{[u>s]}\,\mathrm{d}x\,\mathrm{d}s$$

- $h_1 \leq h_2$ a. e., $\chi_i \in \operatorname*{argmin}_{\chi \in BV(\Omega,\{0,1\})} \int_\Omega h_i \chi\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega) \Rightarrow \chi_1 \geq \chi_2$
- $\chi^s \in \operatorname*{argmin}_{\chi \in BV(\Omega,\{0,1\})} \left\{ E_s^{\mathrm{rel}}[\chi] := \int_\Omega \partial_s \Psi(x,s)\chi\,\mathrm{d}x + |\mathrm{D}\chi|(\Omega) \right\}$
- $\Psi(x,\cdot)$ strictly convex $\Rightarrow \partial_t \Psi(x,\cdot)$ is strictly increasing $\Rightarrow \chi^{s_1} \geq \chi^{s_2}$ a. e. for $s_1 \leq s_2$.
- graph reconstruction $u^*(x) = \sup\{s \,|\, \chi^s(x) = 1\}$ is well-defined

$$\Rightarrow E^{\mathrm{rel}}[u] = C_\Psi + \int_0^1 E_s^{\mathrm{rel}}\left[\chi_{[u>s]}\right]\mathrm{d}s \geq C_\Psi + \int_0^1 E_s^{\mathrm{rel}}\left[\chi^s\right]\mathrm{d}s$$
$$= E^{\mathrm{rel}}[u^*] \geq E^{\mathrm{rel}}[u]$$

Predual functional

The dual of $BV(\Omega)$ is difficult to handle.

Predual functional

The dual of $BV(\Omega)$ is difficult to handle. $\Rightarrow$ Use predual of E^{rel}.

Predual functional

The dual of $BV(\Omega)$ is difficult to handle. $\Rightarrow$ Use predual of E^{rel}.

Starting point: $D^{\mathrm{rel}}[q] = F[q] + G[\Lambda q]$ [Hintermüller/Kunisch '04]

$$F[q] = I_{\overline{B}_1}[q] = \begin{cases} 0 & \text{if } |q| \leq 1 \text{ a.e.} \\ +\infty & \text{else} \end{cases}$$

$$G[v] = \int_\Omega \frac{\frac{1}{4}v^2 + v\theta_2 - \theta_1\theta_2}{\theta_1 + \theta_2} \, \mathrm{d}x$$

Predual functional

The dual of $BV(\Omega)$ is difficult to handle. $\Rightarrow$ Use predual of E^{rel}.

Starting point: $D^{\text{rel}}[q] = F[q] + G[\Lambda q]$ [Hintermüller/Kunisch '04]

$$F[q] = I_{\overline{B}_1}[q] = \begin{cases} 0 & \text{if } |q| \leq 1 \text{ a.e.} \\ +\infty & \text{else} \end{cases}$$

$$G[v] = \int_\Omega \frac{\frac{1}{4}v^2 + v\theta_2 - \theta_1\theta_2}{\theta_1 + \theta_2} \,\mathrm{d}x$$

D^{rel} is the predual functional of E^{rel}, i. e. $(D^{\text{rel}})^* = E^{\text{rel}}$.

Predual functional

The dual of $BV(\Omega)$ is difficult to handle. $\Rightarrow$ Use predual of E^{rel}.

Starting point: $D^{\text{rel}}[q] = F[q] + G[\Lambda q]$ [Hintermüller/Kunisch '04]

$$F[q] = I_{\overline{B}_1}[q] = \begin{cases} 0 & \text{if } |q| \leq 1 \text{ a.e.} \\ +\infty & \text{else} \end{cases}$$

$$G[v] = \int_\Omega \frac{\frac{1}{4}v^2 + v\theta_2 - \theta_1\theta_2}{\theta_1 + \theta_2} \, \mathrm{d}x$$

D^{rel} is the predual functional of E^{rel}, i. e. $(D^{\text{rel}})^* = E^{\text{rel}}$.

- $\Lambda \in \mathcal{L}(\mathcal{Q}, \mathcal{V})$, $\mathcal{Q} = H_N(\text{div}, \Omega)$ and $\mathcal{V} = L^2(\Omega)$

Predual functional

The dual of $BV(\Omega)$ is difficult to handle. $\Rightarrow$ Use predual of E^{rel}.

Starting point: $D^{\text{rel}}[q] = F[q] + G[\Lambda q]$ [Hintermüller/Kunisch '04]

$$F[q] = I_{\overline{B}_1}[q] = \begin{cases} 0 & \text{if } |q| \leq 1 \text{ a.e.} \\ +\infty & \text{else} \end{cases}$$

$$G[v] = \int_\Omega \frac{\frac{1}{4}v^2 + v\theta_2 - \theta_1\theta_2}{\theta_1 + \theta_2}\,\mathrm{d}x$$

D^{rel} is the predual functional of E^{rel}, i. e. $(D^{\text{rel}})^* = E^{\text{rel}}$.

- $\Lambda \in \mathcal{L}(\mathcal{Q}, \mathcal{V})$, $\mathcal{Q} = H_N(\mathrm{div}, \Omega)$ and $\mathcal{V} = L^2(\Omega)$
- $H(\mathrm{div}, \Omega) = \{q \in L^2(\Omega, \mathbb{R}^n) : \ \mathrm{div}\, q \in L^2(\Omega)\}$ and $H_N(\mathrm{div}, \Omega) = H(\mathrm{div}, \Omega) \cap \{q \cdot \nu = 0 \text{ on } \partial\Omega\}$

Predual functional

The dual of $BV(\Omega)$ is difficult to handle. $\Rightarrow$ Use predual of E^{rel}.

Starting point: $D^{\text{rel}}[q] = F[q] + G[\Lambda q]$ [Hintermüller/Kunisch '04]

$$F[q] = I_{\overline{B}_1}[q] = \begin{cases} 0 & \text{if } |q| \le 1 \text{ a.e.} \\ +\infty & \text{else} \end{cases}$$

$$G[v] = \int_\Omega \frac{\frac{1}{4}v^2 + v\theta_2 - \theta_1\theta_2}{\theta_1 + \theta_2}\,\mathrm{d}x$$

D^{rel} is the predual functional of E^{rel}, i. e. $(D^{\text{rel}})^* = E^{\text{rel}}$.

- $\Lambda \in \mathcal{L}(\mathcal{Q}, \mathcal{V})$, $\mathcal{Q} = H_N(\mathrm{div}, \Omega)$ and $\mathcal{V} = L^2(\Omega)$
- $H(\mathrm{div}, \Omega) = \{q \in L^2(\Omega, \mathbb{R}^n) : \mathrm{div}\, q \in L^2(\Omega)\}$ and $H_N(\mathrm{div}, \Omega) = H(\mathrm{div}, \Omega) \cap \{q \cdot \nu = 0 \text{ on } \partial\Omega\}$
- $\Lambda = \mathrm{div}$, $\Lambda^* = -\nabla$ holds in the sense

$$\langle \Lambda^* v, q \rangle = (v, \mathrm{div}\, q)_{L^2(\Omega)} \quad \forall v \in \mathcal{V},\ q \in \mathcal{Q}$$

Predual functional (cont.)

From general theory one knows

$$(D^{\mathsf{rel}})^*[v] = F^*[-\Lambda^* v] + G^*[v].$$

$$F[q] = I_{\overline{B}_1}[q], \quad G[v] = \int_\Omega \frac{\frac{1}{4}v^2 + v\theta_2 - \theta_1\theta_2}{\theta_1 + \theta_2}\,\mathrm{d}x$$

Predual functional (cont.)

From general theory one knows

$$(D^{\text{rel}})^*[v] = F^*[-\Lambda^* v] + G^*[v].$$

The Fenchel conjugates can be computed as follows:

$$F^*[-\Lambda^* v] = \sup_{q \in \mathcal{Q}} (\langle -\Lambda^* v, q \rangle - F[q])$$

$$F[q] = I_{\overline{B}_1}[q], \quad G[v] = \int_\Omega \frac{\frac{1}{4}v^2 + v\theta_2 - \theta_1\theta_2}{\theta_1 + \theta_2} \, \mathrm{d}x$$

Predual functional (cont.)

From general theory one knows

$$(D^{\mathsf{rel}})^*[v] = F^*[-\Lambda^* v] + G^*[v].$$

The Fenchel conjugates can be computed as follows:

$$F^*[-\Lambda^* v] = \sup_{q \in \mathcal{Q}} (\langle -\Lambda^* v, q \rangle - F[q]) = \sup_{q \in \mathcal{Q}} \left(-\int_\Omega v \operatorname{div} q \, \mathrm{d}x - I_{\overline{B}_1}[q] \right)$$

$$= \sup_{q \in \mathcal{Q}, \|q\|_\infty \leq 1} \int_\Omega v \operatorname{div} q \, \mathrm{d}x = |\mathrm{D}v|(\Omega)$$

$$F[q] = I_{\overline{B}_1}[q], \quad G[v] = \int_\Omega \frac{\frac{1}{4}v^2 + v\theta_2 - \theta_1\theta_2}{\theta_1 + \theta_2} \, \mathrm{d}x$$

Predual functional (cont.)

From general theory one knows

$$(D^{\text{rel}})^*[v] = F^*[-\Lambda^* v] + G^*[v].$$

The Fenchel conjugates can be computed as follows:

$$F^*[-\Lambda^* v] = \sup_{q\in\mathcal{Q}}(\langle -\Lambda^* v, q\rangle - F[q]) = \sup_{q\in\mathcal{Q}}\left(-\int_\Omega v\,\mathrm{div}\,q\,\mathrm{d}x - I_{\overline{B}_1}[q]\right)$$

$$= \sup_{q\in\mathcal{Q},\|q\|_\infty\leq 1}\int_\Omega v\,\mathrm{div}\,q\,\mathrm{d}x = |\mathrm{D}v|(\Omega)$$

$$G^*[v] = \sup_{w\in L^2(\Omega)}\left((v,w)_{L^2(\Omega)} - G[w]\right)$$

$$F[q] = I_{\overline{B}_1}[q], \quad G[v] = \int_\Omega \frac{\frac{1}{4}v^2 + v\theta_2 - \theta_1\theta_2}{\theta_1+\theta_2}\,\mathrm{d}x$$

Predual functional (cont.)

From general theory one knows

$$(D^{\mathsf{rel}})^*[v] = F^*[-\Lambda^* v] + G^*[v].$$

The Fenchel conjugates can be computed as follows:

$$\begin{aligned} F^*[-\Lambda^* v] &= \sup_{q\in\mathcal{Q}}(\langle -\Lambda^* v, q\rangle - F[q]) = \sup_{q\in\mathcal{Q}}\left(-\int_\Omega v \operatorname{div} q \,\mathrm{d}x - I_{\overline{B}_1}[q]\right) \\ &= \sup_{q\in\mathcal{Q},\|q\|_\infty\le 1} \int_\Omega v \operatorname{div} q \,\mathrm{d}x = |\mathrm{D}v|(\Omega) \\ G^*[v] &= \sup_{w\in L^2(\Omega)} \left((v,w)_{L^2(\Omega)} - G[w]\right) = \int_\Omega v^2\theta_1 + (1-v)^2\theta_2 \,\mathrm{d}x \end{aligned}$$

where last supremum is attained for $w = 2v(\theta_1+\theta_2) - 2\theta_2$.

$$F[q] = I_{\overline{B}_1}[q], \quad G[v] = \int_\Omega \frac{\frac{1}{4}v^2 + v\theta_2 - \theta_1\theta_2}{\theta_1+\theta_2}\,\mathrm{d}x$$

Predual functional (cont.)

From general theory one knows

$$(D^{\text{rel}})^*[v] = F^*[-\Lambda^* v] + G^*[v].$$

The Fenchel conjugates can be computed as follows:

$$F^*[-\Lambda^* v] = \sup_{q\in\mathcal{Q}} (\langle -\Lambda^* v, q\rangle - F[q]) = \sup_{q\in\mathcal{Q}} \left(- \int_\Omega v \,\mathrm{div}\, q \,\mathrm{d}x - I_{\overline{B}_1}[q] \right)$$

$$= \sup_{q\in\mathcal{Q}, \|q\|_\infty \leq 1} \int_\Omega v \,\mathrm{div}\, q \,\mathrm{d}x = |\mathrm{D}v|(\Omega)$$

$$G^*[v] = \sup_{w\in L^2(\Omega)} \left((v,w)_{L^2(\Omega)} - G[w] \right) = \int_\Omega v^2\theta_1 + (1-v)^2\theta_2 \,\mathrm{d}x$$

where last supremum is attained for $w = 2v(\theta_1+\theta_2) - 2\theta_2$. Thus,

$$(D^{\text{rel}})^*[v] = F^*[-\Lambda^* v] + G^*[v] = E^{\text{rel}}[v].$$

$$F[q] = I_{\overline{B}_1}[q], \quad G[v] = \int_\Omega \frac{\frac{1}{4}v^2 + v\theta_2 - \theta_1\theta_2}{\theta_1+\theta_2} \,\mathrm{d}x$$

Predual functional (cont.)

Central insight: Minimizers p of D^{rel} and u of $(D^{\mathsf{rel}})^*$ fulfill

$$D^{\mathsf{rel}}[p] = -(D^{\mathsf{rel}})^*[u]$$

Predual functional (cont.)

Central insight: Minimizers p of D^{rel} and u of $(D^{\text{rel}})^*$ fulfill

$$D^{\text{rel}}[p] = -(D^{\text{rel}})^*[u]$$

Can be seen by formally exchanging $\inf$ and $\sup$:

$$D^{\text{rel}}[p] = \inf_{q \in H_N(\mathrm{div},\Omega)} (F[q] + G[\Lambda q])$$

Predual functional (cont.)

Central insight: Minimizers p of D^{rel} and u of $(D^{\text{rel}})^*$ fulfill

$$D^{\text{rel}}[p] = -(D^{\text{rel}})^*[u]$$

Can be seen by formally exchanging $\inf$ and $\sup$:

$$D^{\text{rel}}[p] = \inf_{q \in H_N(\text{div},\Omega)} (F[q] + G[\Lambda q]) = \inf_{q \in H_N(\text{div},\Omega)} (F[q] + G^{**}[\Lambda q])$$

Predual functional (cont.)

Central insight: Minimizers p of D^{rel} and u of $(D^{\mathrm{rel}})^*$ fulfill

$$D^{\mathrm{rel}}[p] = -(D^{\mathrm{rel}})^*[u]$$

Can be seen by formally exchanging $\inf$ and $\sup$:

$$\begin{aligned} D^{\mathrm{rel}}[p] &= \inf_{q \in H_N(\mathrm{div},\Omega)} (F[q] + G[\Lambda q]) = \inf_{q \in H_N(\mathrm{div},\Omega)} (F[q] + G^{**}[\Lambda q]) \\ &= \inf_{q \in H_N(\mathrm{div},\Omega)} \sup_{v \in L^2(\Omega)} (F[q] + \langle v, \Lambda q \rangle - G^*[v]) \end{aligned}$$

Predual functional (cont.)

Central insight: Minimizers p of D^{rel} and u of $(D^{\text{rel}})^*$ fulfill

$$D^{\text{rel}}[p] = -(D^{\text{rel}})^*[u]$$

Can be seen by formally exchanging $\inf$ and $\sup$:

$$\begin{aligned}
D^{\text{rel}}[p] &= \inf_{q \in H_N(\text{div},\Omega)} \left(F[q] + G[\Lambda q]\right) = \inf_{q \in H_N(\text{div},\Omega)} \left(F[q] + G^{**}[\Lambda q]\right) \\
&= \inf_{q \in H_N(\text{div},\Omega)} \sup_{v \in L^2(\Omega)} \left(F[q] + \langle v, \Lambda q \rangle - G^*[v]\right) \\
&= \sup_{v \in L^2(\Omega)} \left(- \sup_{q \in H_N(\text{div},\Omega)} \left(\langle -\Lambda^* v, q \rangle - F[q]\right) - G^*[v] \right)
\end{aligned}$$

Predual functional (cont.)

Central insight: Minimizers p of D^{rel} and u of $(D^{\text{rel}})^*$ fulfill

$$D^{\text{rel}}[p] = -(D^{\text{rel}})^*[u]$$

Can be seen by formally exchanging $\inf$ and $\sup$:

$$\begin{aligned}
D^{\text{rel}}[p] &= \inf_{q \in H_N(\text{div},\Omega)} \left(F[q] + G[\Lambda q]\right) = \inf_{q \in H_N(\text{div},\Omega)} \left(F[q] + G^{**}[\Lambda q]\right) \\
&= \inf_{q \in H_N(\text{div},\Omega)} \sup_{v \in L^2(\Omega)} \left(F[q] + \langle v, \Lambda q \rangle - G^*[v]\right) \\
&= \sup_{v \in L^2(\Omega)} \left(- \sup_{q \in H_N(\text{div},\Omega)} \left(\langle -\Lambda^* v, q \rangle - F[q]\right) - G^*[v] \right) \\
&= \sup_{v \in L^2(\Omega)} \left(-F^*[-\Lambda^* v] - G^*[v]\right)
\end{aligned}$$

Predual functional (cont.)

Central insight: Minimizers p of D^{rel} and u of $(D^{\text{rel}})^*$ fulfill

$$D^{\text{rel}}[p] = -(D^{\text{rel}})^*[u]$$

Can be seen by formally exchanging $\inf$ and $\sup$:

$$\begin{aligned}
D^{\text{rel}}[p] &= \inf_{q\in H_N(\operatorname{div},\Omega)} (F[q] + G[\Lambda q]) = \inf_{q\in H_N(\operatorname{div},\Omega)} (F[q] + G^{**}[\Lambda q]) \\
&= \inf_{q\in H_N(\operatorname{div},\Omega)} \sup_{v\in L^2(\Omega)} (F[q] + \langle v, \Lambda q\rangle - G^*[v]) \\
&= \sup_{v\in L^2(\Omega)} \left(- \sup_{q\in H_N(\operatorname{div},\Omega)} (\langle -\Lambda^* v, q\rangle - F[q]) - G^*[v] \right) \\
&= \sup_{v\in L^2(\Omega)} (-F^*[-\Lambda^* v] - G^*[v]) \\
&= \sup_{v\in L^2(\Omega)} (-(D^{\text{rel}})^*[v]) = - \inf_{v\in L^2(\Omega)} (D^{\text{rel}})^*[v] = -(D^{\text{rel}})^*[u]\,.
\end{aligned}$$

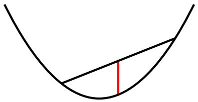
Functional a posteriori error estimates for E^{rel}

Two measures of uniform convexity for $J : X \to \mathbb{R}$

Functional a posteriori error estimates for E^{rel}

Two measures of uniform convexity for $J : X \to \mathbb{R}$

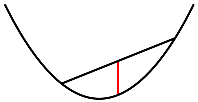
$$J\left[\frac{x_1+x_2}{2}\right] + \Phi_J(x_2 - x_1) \leq \frac{1}{2}(J[x_1] + J[x_2]) \text{ for all } x_1, x_2 \in X$$



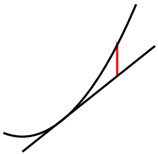
Functional a posteriori error estimates for E^{rel}

Two measures of uniform convexity for $J : X \to \mathbb{R}$

$$J\left[\tfrac{x_1+x_2}{2}\right] + \Phi_J(x_2 - x_1) \leq \tfrac{1}{2}(J[x_1] + J[x_2]) \text{ for all } x_1, x_2 \in X$$



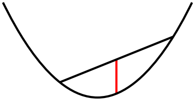
$$\langle x', x_2 - x_1\rangle + \Psi_J(x_2 - x_1) \leq J[x_2] - J[x_1] \text{ for all } x' \in \partial J[x_1]$$



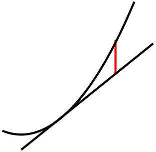
Functional a posteriori error estimates for E^{rel}

Two measures of uniform convexity for $J : X \to \mathbb{R}$

$$J\left[\tfrac{x_1+x_2}{2}\right] + \Phi_J(x_2 - x_1) \leq \tfrac{1}{2}(J[x_1] + J[x_2]) \text{ for all } x_1, x_2 \in X$$



$$\langle x', x_2 - x_1\rangle + \Psi_J(x_2 - x_1) \leq J[x_2] - J[x_1] \text{ for all } x' \in \partial J[x_1]$$



Theorem [Repin '00]: Let $u \in \operatorname{argmin}_{\tilde{v}\in\mathcal{V}} E^{\text{rel}}[\tilde{v}]$ and $q \in \mathcal{Q}$, $v \in \mathcal{V}' = \mathcal{V} = L^2(\Omega)$. Then,

$$\Phi_{G^*}(v-u) + \Phi_{F^*}(-\Lambda^*(v-u)) + \Psi_{E^{\text{rel}}}\left(\tfrac{v-u}{2}\right) \leq \tfrac{1}{2}(E^{\text{rel}}[v] + D^{\text{rel}}[q])\,.$$

Repin theorem - Proof sketch

The convexity property of Φ_{G^*} and Φ_{F^*} gives:

$$\begin{aligned}
&\Phi_{G^*}(v-u)+\Phi_{F^*}(-\Lambda^*(v-u)) \\
&\quad\leq \tfrac{1}{2}\left(F^*[-\Lambda^* v]+G^*[v]+F^*[-\Lambda^* u]+G^*[u]\right) \\
&\qquad -\left(F^*\left[-\Lambda^* \tfrac{u+v}{2}\right]+G^*\left[\tfrac{u+v}{2}\right]\right)
\end{aligned}$$

Repin theorem - Proof sketch

The convexity property of Φ_{G^*} and Φ_{F^*} gives:

$$\begin{aligned}
&\Phi_{G^*}(v-u) + \Phi_{F^*}(-\Lambda^*(v-u)) \\
&\quad \le \tfrac{1}{2}\left(F^*[-\Lambda^* v] + G^*[v] + F^*[-\Lambda^* u] + G^*[u]\right) \\
&\qquad - \left(F^*\left[-\Lambda^* \tfrac{u+v}{2}\right] + G^*\left[\tfrac{u+v}{2}\right]\right) = \tfrac{1}{2}(E^{\mathsf{rel}}[v] + E^{\mathsf{rel}}[u]) - E^{\mathsf{rel}}\left[\tfrac{u+v}{2}\right]
\end{aligned}$$

Repin theorem - Proof sketch

The convexity property of Φ_{G^*} and Φ_{F^*} gives:

$$\begin{aligned}&\Phi_{G^*}(v-u) + \Phi_{F^*}(-\Lambda^*(v-u))\\ &\quad\leq \tfrac{1}{2}\left(F^*[-\Lambda^* v] + G^*[v] + F^*[-\Lambda^* u] + G^*[u]\right)\\ &\qquad - \left(F^*\left[-\Lambda^* \tfrac{u+v}{2}\right] + G^*\left[\tfrac{u+v}{2}\right]\right) = \tfrac{1}{2}(E^{\mathsf{rel}}[v] + E^{\mathsf{rel}}[u]) - E^{\mathsf{rel}}\left[\tfrac{u+v}{2}\right]\end{aligned}$$

Since u is a minimizer of E^{rel} and thus $0 \in \partial E^{\mathsf{rel}}[u]$, we get

Repin theorem - Proof sketch

The convexity property of Φ_{G^*} and Φ_{F^*} gives:

$$\begin{aligned}
&\Phi_{G^*}(v-u) + \Phi_{F^*}(-\Lambda^*(v-u)) \\
&\quad \leq \tfrac{1}{2}\left(F^*[-\Lambda^* v] + G^*[v] + F^*[-\Lambda^* u] + G^*[u]\right) \\
&\qquad - \left(F^*\left[-\Lambda^* \tfrac{u+v}{2}\right] + G^*\left[\tfrac{u+v}{2}\right]\right) = \tfrac{1}{2}(E^{\mathsf{rel}}[v] + E^{\mathsf{rel}}[u]) - E^{\mathsf{rel}}\left[\tfrac{u+v}{2}\right]
\end{aligned}$$

Since u is a minimizer of E^{rel} and thus $0 \in \partial E^{\mathsf{rel}}[u]$, we get

$$\Psi_{E^{\mathsf{rel}}}\left(\tfrac{v-u}{2}\right) = \Psi_{E^{\mathsf{rel}}}\left(\tfrac{u+v}{2} - u\right) \leq E^{\mathsf{rel}}\left[\tfrac{u+v}{2}\right] - E^{\mathsf{rel}}[u].$$

Repin theorem - Proof sketch

The convexity property of Φ_{G^*} and Φ_{F^*} gives:

$$\begin{aligned}&\Phi_{G^*}(v-u)+\Phi_{F^*}(-\Lambda^*(v-u))\\ &\quad\leq \tfrac{1}{2}\left(F^*[-\Lambda^* v]+G^*[v]+F^*[-\Lambda^* u]+G^*[u]\right)\\ &\qquad-\left(F^*\left[-\Lambda^*\tfrac{u+v}{2}\right]+G^*\left[\tfrac{u+v}{2}\right]\right)=\tfrac{1}{2}(E^{\mathsf{rel}}[v]+E^{\mathsf{rel}}[u])-E^{\mathsf{rel}}[\tfrac{u+v}{2}]\end{aligned}$$

Since u is a minimizer of E^{rel} and thus $0\in\partial E^{\mathsf{rel}}[u]$, we get

$$\Psi_{E^{\mathsf{rel}}}\left(\tfrac{v-u}{2}\right)=\Psi_{E^{\mathsf{rel}}}\left(\tfrac{u+v}{2}-u\right)\leq E^{\mathsf{rel}}[\tfrac{u+v}{2}]-E^{\mathsf{rel}}[u].$$

Adding both inequalities gives

$$\Phi_{G^*}(v-u)+\Phi_{F^*}(-\Lambda^*(v-u))+\Psi_{E^{\mathsf{rel}}}\left(\tfrac{v-u}{2}\right)\leq\tfrac{1}{2}(E^{\mathsf{rel}}[v]-E^{\mathsf{rel}}[u]).$$

Repin theorem - Proof sketch

The convexity property of Φ_{G^*} and Φ_{F^*} gives:

$$\begin{aligned}
&\Phi_{G^*}(v-u) + \Phi_{F^*}(-\Lambda^*(v-u)) \\
&\quad \leq \tfrac{1}{2}\left(F^*[-\Lambda^* v] + G^*[v] + F^*[-\Lambda^* u] + G^*[u]\right) \\
&\qquad - \left(F^*\left[-\Lambda^* \tfrac{u+v}{2}\right] + G^*\left[\tfrac{u+v}{2}\right]\right) = \tfrac{1}{2}(E^{\mathsf{rel}}[v] + E^{\mathsf{rel}}[u]) - E^{\mathsf{rel}}[\tfrac{u+v}{2}]
\end{aligned}$$

Since u is a minimizer of E^{rel} and thus $0 \in \partial E^{\mathsf{rel}}[u]$, we get

$$\Psi_{E^{\mathsf{rel}}}\left(\tfrac{v-u}{2}\right) = \Psi_{E^{\mathsf{rel}}}\left(\tfrac{u+v}{2} - u\right) \leq E^{\mathsf{rel}}[\tfrac{u+v}{2}] - E^{\mathsf{rel}}[u].$$

Adding both inequalities gives

$$\Phi_{G^*}(v-u) + \Phi_{F^*}(-\Lambda^*(v-u)) + \Psi_{E^{\mathsf{rel}}}\left(\tfrac{v-u}{2}\right) \leq \tfrac{1}{2}(E^{\mathsf{rel}}[v] - E^{\mathsf{rel}}[u]).$$

The claim follows using the *weak complementarity principle*

$$E^{\mathsf{rel}}[u] \geq -D^{\mathsf{rel}}[q].$$

Functional a posteriori error estimates for E^{rel}

Theorem [Repin '00]: Let $u \in \operatorname{argmin}_{\tilde{v}\in\mathcal{V}} E^{\mathsf{rel}}[\tilde{v}]$ and $q \in \mathcal{Q}$, $v \in \mathcal{V}' = \mathcal{V} = L^2(\Omega)$. Then,

$$\Phi_{G^*}(v-u) + \Phi_{F^*}(-\Lambda^*(v-u)) + \Psi_{E^{\mathsf{rel}}}\left(\frac{v-u}{2}\right) \leq \frac{1}{2}(E^{\mathsf{rel}}[v] + D^{\mathsf{rel}}[q])\,.$$

Functional a posteriori error estimates for E^{rel}

Theorem [Repin '00]: Let $u \in \operatorname{argmin}_{\tilde{v}\in\mathcal{V}} E^{\mathsf{rel}}[\tilde{v}]$ and $q \in \mathcal{Q}$, $v \in \mathcal{V}' = \mathcal{V} = L^2(\Omega)$. Then,

$$\Phi_{G^*}(v-u) + \Phi_{F^*}(-\Lambda^*(v-u)) + \Psi_{E^{\mathsf{rel}}}\left(\tfrac{v-u}{2}\right) \le \tfrac{1}{2}(E^{\mathsf{rel}}[v] + D^{\mathsf{rel}}[q])\,.$$

In our case, one gets

$$\Phi_{F^*} \equiv 0,\ \Phi_{G^*}(v) = \frac{1}{4}\int_\Omega v^2(\theta_1+\theta_2)\,\mathrm{d}x,\ \Psi_{E^{\mathsf{rel}}}(v) = \int_\Omega v^2(\theta_1+\theta_2)\,\mathrm{d}x.$$

Functional a posteriori error estimates for E^{rel}

Theorem [Repin '00]: Let $u \in \operatorname{argmin}_{\tilde{v} \in \mathcal{V}} E^{\text{rel}}[\tilde{v}]$ and $q \in \mathcal{Q}$, $v \in \mathcal{V}' = \mathcal{V} = L^2(\Omega)$. Then,

$$\Phi_{G^*}(v-u) + \Phi_{F^*}(-\Lambda^*(v-u)) + \Psi_{E^{\text{rel}}}\left(\frac{v-u}{2}\right) \leq \frac{1}{2}(E^{\text{rel}}[v] + D^{\text{rel}}[q]) \,.$$

In our case, one gets

$$\Phi_{F^*} \equiv 0, \;\; \Phi_{G^*}(v) = \frac{1}{4}\int_\Omega v^2(\theta_1+\theta_2)\,\mathrm{d}x, \;\; \Psi_{E^{\text{rel}}}(v) = \int_\Omega v^2(\theta_1+\theta_2)\,\mathrm{d}x.$$

Using $\frac{1}{2\nu}(c_1-c_2)^2 \leq \theta_1+\theta_2$ this leads to:

Theorem [A posteriori error estimate for E^{rel}]: Let $u \in \mathcal{V}$ be the minimizer of E^{rel}. Then, for any $v \in \mathcal{V}$ and $q \in \mathcal{Q}$ it holds that

$$\|u-v\|^2_{L^2(\Omega)} \leq \operatorname{err}^2_u[v,q] := \frac{2\nu}{(c_1-c_2)^2}\left(E^{\text{rel}}[v] + D^{\text{rel}}[q]\right)$$

Functional a posteriori error estimates for E^{rel}

Theorem [Repin '00]: Let $u \in \operatorname{argmin}_{\tilde{v} \in \mathcal{V}} E^{\text{rel}}[\tilde{v}]$ and $q \in \mathcal{Q}$, $v \in \mathcal{V}' = \mathcal{V} = L^2(\Omega)$. Then,

$$\Phi_{G^*}(v-u) + \Phi_{F^*}(-\Lambda^*(v-u)) + \Psi_{E^{\text{rel}}}\left(\tfrac{v-u}{2}\right) \leq \tfrac{1}{2}(E^{\text{rel}}[v] + D^{\text{rel}}[q])\,.$$

In our case, one gets

$$\Phi_{F^*} \equiv 0,\ \Phi_{G^*}(v) = \frac{1}{4}\int_\Omega v^2(\theta_1+\theta_2)\,\mathrm{d}x,\ \Psi_{E^{\text{rel}}}(v) = \int_\Omega v^2(\theta_1+\theta_2)\,\mathrm{d}x.$$

Using $\frac{1}{2\nu}(c_1-c_2)^2 \leq \theta_1+\theta_2$ this leads to:

Theorem [A posteriori error estimate for E^{rel}]: Let $u \in \mathcal{V}$ be the minimizer of E^{rel}. Then, for any $v \in \mathcal{V}$ and $q \in \mathcal{Q}$ it holds that

$$\|u-v\|^2_{L^2(\Omega)} \leq \operatorname{err}^2_u[v,q] := \frac{2\nu}{(c_1-c_2)^2}\Big(E^{\text{rel}}[v] + D^{\text{rel}}[q]\Big)$$

$$= \frac{2\nu}{(c_1-c_2)^2}\int_\Omega v^2\theta_1 + (1-v)^2\theta_2 + |\nabla v| + \frac{\frac{1}{4}(\operatorname{div} q)^2 + \operatorname{div} q\theta_2 - \theta_1\theta_2}{\theta_1+\theta_2}\,\mathrm{d}x\,.$$

A posteriori error estimates for the binary model

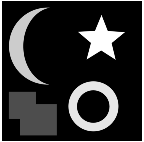
Key observation: solutions of E^{rel} are characterized by steep profiles

A posteriori error estimates for the binary model

Key observation: solutions of E^{rel} are characterized by steep profiles $\mathcal{S}_\eta = \left[\frac{1}{2} - \eta \leq v \leq \frac{1}{2} + \eta\right]$ $\left(\eta \in \left(0, \frac{1}{2}\right)\right)$ set of non properly identified phase regions and $\mathrm{a}[v, \eta] = |\mathcal{S}_\eta|$

A posteriori error estimates for the binary model

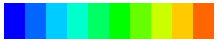
Key observation: solutions of E^{rel} are characterized by steep profiles $\mathcal{S}_\eta = \left[\frac{1}{2} - \eta \leq v \leq \frac{1}{2} + \eta\right]$ $\left(\eta \in \left(0, \frac{1}{2}\right)\right)$ set of non properly identified phase regions and $\mathrm{a}[v, \eta] = |\mathcal{S}_\eta|$

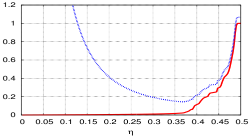


u_0

u

$\chi_{[u>0.5]}$

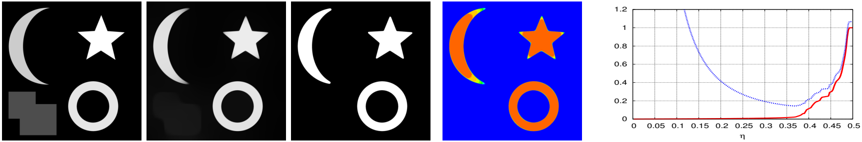
0  1



$\mathrm{a}[v, \eta]$ and $\mathrm{err}_\chi[v, q, \eta]$

A posteriori error estimates for the binary model

Key observation: solutions of E^{rel} are characterized by steep profiles $\mathcal{S}_\eta = \left[\frac{1}{2} - \eta \leq v \leq \frac{1}{2} + \eta\right]$ $\left(\eta \in \left(0, \frac{1}{2}\right)\right)$ set of non properly identified phase regions and $\mathrm{a}[v,\eta] = |\mathcal{S}_\eta|$



u_0 u $\chi_{[u>0.5]}$ 0 1 $\mathrm{a}[v,\eta]$ and $\mathrm{err}_\chi[v,q,\eta]$

Theorem [A posteriori error estimate for E]: Let $\chi \in BV(\Omega, \{0,1\})$ be the minimizer of the binary Mumford–Shah functional obtained from E^{rel}. Then for all $v \in \mathcal{V} = L^2(\Omega)$ and $q \in \mathcal{Q} = H_N(\mathrm{div}, \Omega)$ we have

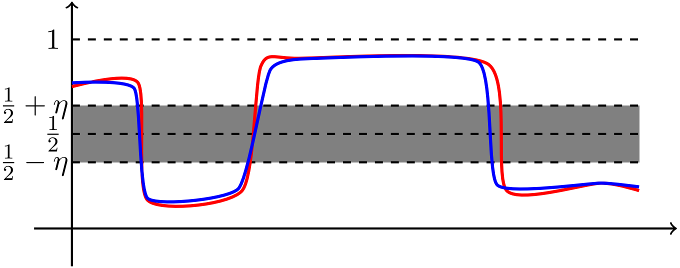
$$\left\|\chi - \chi_{[v>\frac{1}{2}]}\right\|_{L^1(\Omega)} \leq \inf_{\eta \in (0,\frac{1}{2})} \left\{ \mathrm{err}_\chi[v,q,\eta] = \left(\mathrm{a}[v,\eta] + \frac{1}{\eta^2} \mathrm{err}_u^2[v,q] \right) \right\}.$$

A posteriori error estimates for the binary model

Theorem [A posteriori error estimate for E]: Let $\chi \in BV(\Omega, \{0,1\})$ be the minimizer of the binary Mumford–Shah functional obtained from E^{rel}. Then for all $v \in \mathcal{V} = L^2(\Omega)$ and $q \in \mathcal{Q} = H_N(\text{div}, \Omega)$ we have

$$\left\|\chi - \chi_{[v>\frac{1}{2}]}\right\|_{L^1(\Omega)} \leq \inf_{\eta \in (0,\frac{1}{2})} \left\{ \text{err}_\chi[v,q,\eta] = \left(\text{a}[v,\eta] + \tfrac{1}{\eta^2}\text{err}_u^2[v,q] \right) \right\}.$$

Proof: $\left[u > \frac{1}{2}\right] \Delta \left[v > \frac{1}{2}\right] \subseteq \left[\frac{1}{2} - \eta \leq v \leq \frac{1}{2} + \eta\right] \cup \left[|u - v| > \eta\right]$



u minimizer of E^{rel} $\qquad$ $v \in \mathcal{V}$

A posteriori error estimates for the binary model

Theorem [A posteriori error estimate for E]: Let $\chi \in BV(\Omega, \{0,1\})$ be the minimizer of the binary Mumford–Shah functional obtained from E^{rel}. Then for all $v \in \mathcal{V} = L^2(\Omega)$ and $q \in \mathcal{Q} = H_N(\operatorname{div}, \Omega)$ we have

$$\left\|\chi - \chi_{[v>\frac{1}{2}]}\right\|_{L^1(\Omega)} \leq \inf_{\eta \in (0,\frac{1}{2})} \left\{ \operatorname{err}_\chi[v,q,\eta] = \left(\mathrm{a}[v,\eta] + \tfrac{1}{\eta^2} \operatorname{err}_u^2[v,q] \right) \right\} .$$

Proof: $\left[u > \frac{1}{2}\right] \Delta \left[v > \frac{1}{2}\right] \subseteq \left[\frac{1}{2} - \eta \leq v \leq \frac{1}{2} + \eta\right] \cup [|u - v| > \eta]$

Using the estimate for $\|u - v\|_{L^2(\Omega)}^2$, we get

$$\mathcal{L}^n([|u-v| > \eta]) \leq \int_{\{|u-v|>\eta\}} \tfrac{1}{\eta^2} |u-v|^2 \, \mathrm{d}x \leq \tfrac{1}{\eta^2} \operatorname{err}_u^2[v,q] \, .$$

The above holds for any $\eta \in (0, \frac{1}{2})$, so also for $\inf_{\eta \in (0,\frac{1}{2})}$.

Two discretizations and primal-dual algorithm

$\mathscr{T}$ adaptive simplicial mesh
$\mathcal{V}_h^0$ space of piecewise constant finite element functions on $\mathscr{T}$
$\mathcal{V}_h^1$ space of continuous and affine finite element functions on $\mathscr{T}$

Two discretizations and primal-dual algorithm

$\mathscr{T}$ adaptive simplicial mesh
$\mathcal{V}_h^0$ space of piecewise constant finite element functions on $\mathscr{T}$
$\mathcal{V}_h^1$ space of continuous and affine finite element functions on $\mathscr{T}$

$$G_h[v_h] = \int_\Omega \frac{\frac{1}{4}v_h^2 + v_h - \theta_{1,h}\theta_{2,h}}{\theta_{1,h} + \theta_{2,h}}\,\mathrm{d}x\,, \quad F_h[q_h] = I_{\bar{B}_1}[q_h]\,,$$

$$G_h^*[v_h] = \int_\Omega v_h^2\theta_{1,h} + (1-v_h)^2\theta_{2,h}\,\mathrm{d}x\,, \quad F_h^*[q_h] = \int_\Omega \mathcal{I}_h(|q_h|)\,\mathrm{d}x\,,$$

for $\theta_{i,h} = \mathcal{I}_h(\frac{1}{\nu}(c_i - u_0)^2)$ for $i = 1, 2$, $\mathcal{I}_h$ Lagrange interpolation

Two discretizations and primal-dual algorithm

$\mathscr{T}$ adaptive simplicial mesh
$\mathcal{V}_h^0$ space of piecewise constant finite element functions on $\mathscr{T}$
$\mathcal{V}_h^1$ space of continuous and affine finite element functions on $\mathscr{T}$

$$G_h[v_h] = \int_\Omega \frac{\frac{1}{4}v_h^2 + v_h - \theta_{1,h}\theta_{2,h}}{\theta_{1,h} + \theta_{2,h}}\,\mathrm{d}x\,, \quad F_h[q_h] = I_{\bar{B}_1}[q_h]\,,$$

$$G_h^*[v_h] = \int_\Omega v_h^2\theta_{1,h} + (1-v_h)^2\theta_{2,h}\,\mathrm{d}x\,, \quad F_h^*[q_h] = \int_\Omega \mathcal{I}_h(|q_h|)\,\mathrm{d}x\,,$$

for $\theta_{i,h} = \mathcal{I}_h(\frac{1}{\nu}(c_i - u_0)^2)$ for $i = 1,2$, $\mathcal{I}_h$ Lagrange interpolation

(FE) discretization $\mathcal{V}_h = \mathcal{V}_h^1$ with L^2-scalar product and $\mathcal{Q}_h = (\mathcal{V}_h^1)^2$ with lumped mass scalar product,
$-\Lambda_h^* : \mathcal{V}_h \to \mathcal{Q}_h$ implicitly defined via (cf. [Bartels '14])

$$\int_\Omega \mathcal{I}_h(-\Lambda_h^* v_h \cdot q_h)\,\mathrm{d}x = \int_\Omega v_h \mathcal{P}_h \operatorname{div} q_h\,\mathrm{d}x \quad \forall q_h \in \mathcal{Q}_h, v_h \in \mathcal{V}_h$$

$\mathcal{P}_h : L^2(\Omega) \to \mathcal{V}_h$ denotes L^2-projection

Two discretizations and primal-dual algorithm

$\mathscr{T}$ adaptive simplicial mesh
$\mathcal{V}_h^0$ space of piecewise constant finite element functions on $\mathscr{T}$
$\mathcal{V}_h^1$ space of continuous and affine finite element functions on $\mathscr{T}$

$$G_h[v_h] = \int_\Omega \frac{\frac{1}{4}v_h^2 + v_h - \theta_{1,h}\theta_{2,h}}{\theta_{1,h}+\theta_{2,h}}\,\mathrm{d}x\,, \quad F_h[q_h] = I_{\bar{B}_1}[q_h]\,,$$

$$G_h^*[v_h] = \int_\Omega v_h^2\theta_{1,h} + (1-v_h)^2\theta_{2,h}\,\mathrm{d}x\,, \quad F_h^*[q_h] = \int_\Omega |q_h|\,\mathrm{d}x\,,$$

for $\theta_{i,h} = \mathcal{I}_h(\frac{1}{\nu}(c_i - u_0)^2)$ for $i = 1, 2$, $\mathcal{I}_h$ Lagrange interpolation

(FE') discretization $\mathcal{V}_h = \mathcal{V}_h^1$ and $\mathcal{Q}_h = (\mathcal{V}_h^0)^2$ with L^2-scalar product
$\Lambda_h : \mathcal{Q}_h \to \mathcal{V}_h$ given by

$$\int_\Omega \mathcal{I}_h(\Lambda_h q_h v_h)\,\mathrm{d}x = -\int_\Omega q_h \cdot \nabla v_h\,\mathrm{d}x \quad \forall q_h \in \mathcal{Q}_h, v_h \in \mathcal{V}_h$$

$\Rightarrow -\Lambda_h^* v_h = \nabla v_h$
To evaluate err_u^2 an L^2-projection of the dual solution is performed

The primal-dual algorithm

Primal-dual algorithm

[Chambolle, Pock '11]

while $\|u_h^{k+1} - u_h^k\|_\infty > \mathit{THRESHOLD}$ **do**

$$p_h^{k+1} = (\mathrm{Id} + \sigma\partial F_h)^{-1}(p_h^k - \sigma\Lambda_h^* \bar{u}_h^k);$$
$$u_h^{k+1} = (\mathrm{Id} + \tau\partial G_h^*)^{-1}(u_h^k + \tau\Lambda_h p_h^{k+1});$$
$$\bar{u}_h^{k+1} = 2u_h^{k+1} - u_h^k;$$

end

Primal-dual algorithm [Chambolle, Pock '11]

while $\|u_h^{k+1} - u_h^k\|_\infty >$ *THRESHOLD* **do**

$$p_h^{k+1} = (\mathrm{Id} + \sigma \partial F_h)^{-1}(p_h^k - \sigma \Lambda_h^* \bar{u}_h^k);$$
$$u_h^{k+1} = (\mathrm{Id} + \tau \partial G_h^*)^{-1}(u_h^k + \tau \Lambda_h p_h^{k+1});$$
$$\bar{u}_h^{k+1} = 2u_h^{k+1} - u_h^k;$$

end

- The resolvent $(\mathrm{Id} + \sigma \partial F_h)^{-1}$ for F_h is well known.

The primal-dual algorithm

Primal-dual algorithm

[Chambolle, Pock '11]

while $\|u_h^{k+1} - u_h^k\|_\infty >$ *THRESHOLD* **do**

$$p_h^{k+1} = (\mathrm{Id} + \sigma\partial F_h)^{-1}(p_h^k - \sigma\Lambda_h^*\bar{u}_h^k);$$
$$u_h^{k+1} = (\mathrm{Id} + \tau\partial G_h^*)^{-1}(u_h^k + \tau\Lambda_h p_h^{k+1});$$
$$\bar{u}_h^{k+1} = 2u_h^{k+1} - u_h^k;$$

end

- The resolvent $(\mathrm{Id} + \sigma\partial F_h)^{-1}$ for F_h is well known.
- The resolvent $(\mathrm{Id} + \tau\partial G_h^*)^{-1}$ for

$$G_h^*[v_h] = \int_\Omega v_h^2\theta_{1,h} + (1 - v_h)^2\theta_{2,h}\,\mathrm{d}x$$

can be computed in a straightforwardly, but involves mass matrices.

The primal-dual algorithm

Primal-dual algorithm

[Chambolle, Pock '11]

while $\|u_h^{k+1} - u_h^k\|_\infty > \mathit{THRESHOLD}$ **do**

$$p_h^{k+1} = (\mathrm{Id} + \sigma\partial F_h)^{-1}(p_h^k - \sigma\Lambda_h^*\bar{u}_h^k);$$
$$u_h^{k+1} = (\mathrm{Id} + \tau\partial G_h^*)^{-1}(u_h^k + \tau\Lambda_h p_h^{k+1});$$
$$\bar{u}_h^{k+1} = 2u_h^{k+1} - u_h^k;$$

end

- The resolvent $(\mathrm{Id} + \sigma\partial F_h)^{-1}$ for F_h is well known.
- The resolvent $(\mathrm{Id} + \tau\partial G_h^*)^{-1}$ for

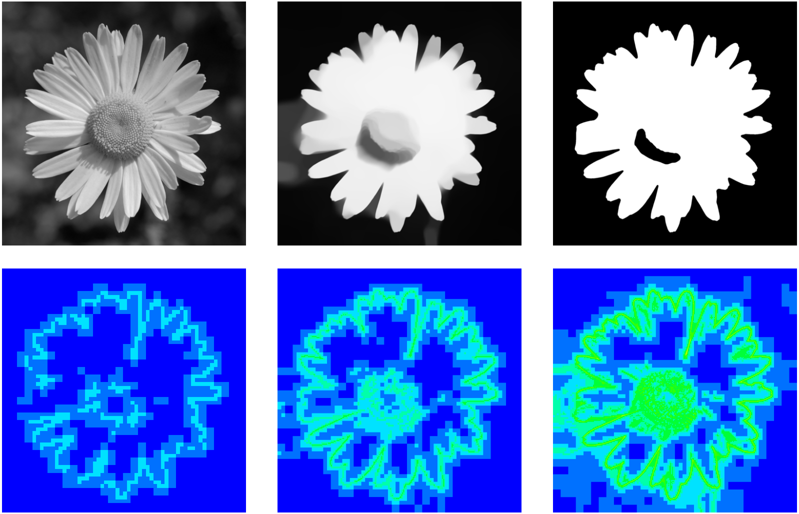
$$G_h^*[v_h] = \int_\Omega v_h^2\theta_{1,h} + (1 - v_h)^2\theta_{2,h}\,\mathrm{d}x$$

  can be computed in a straightforwardly, but involves mass matrices.
- The operator norm of Λ can be estimated as follows
  - $\|\Lambda_h\|^2 \leq 48(3 + 2\sqrt{2})h_{\mathrm{min}}^{-2}$ for (FE') and $n = 2$
  - $\|\Lambda_h\|^2 \leq 96(3 + 2\sqrt{2})h_{\mathrm{min}}^{-2}$ for (FE)

Numerical results - "Flower"

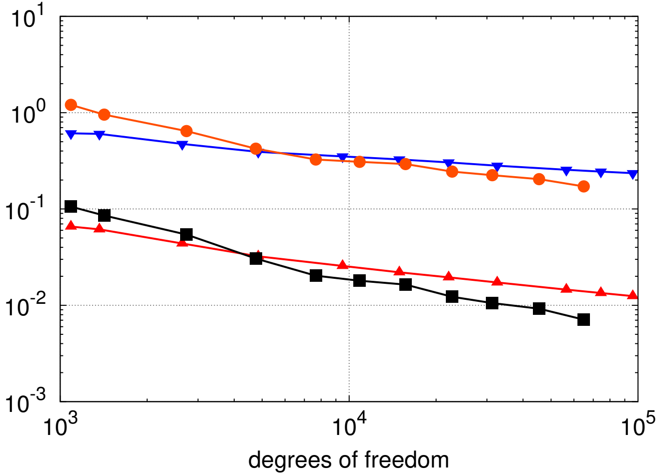
Numerical results - “Flower”

u_0 u_h $\chi_{[u_h > 0.5]}$



mesh after 2^{nd}, 5^{th}, 10^{th} iteration

Numerical results - "Flower" (cont.)



err_u^2 estimator for $\|u - u_h\|_{L^2(\Omega)}^2$ ((FE) (black), (FE') (red))

err_χ estimator for $\|\chi - \chi_{[u_h>0.5]}\|_{L^1(\Omega)}$ ((FE) (orange), (FE') (blue))

Numerical results - “Cameraman”

u_0 u_h $\chi_{[u_h>0.5]}$

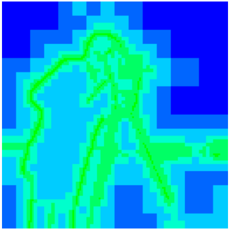
Image source: MATLAB

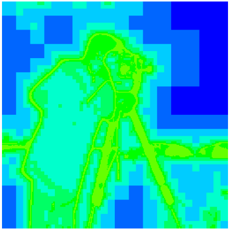
Numerical results - “Cameraman”

u_0

u_h

$\chi_{[u_h>0.5]}$

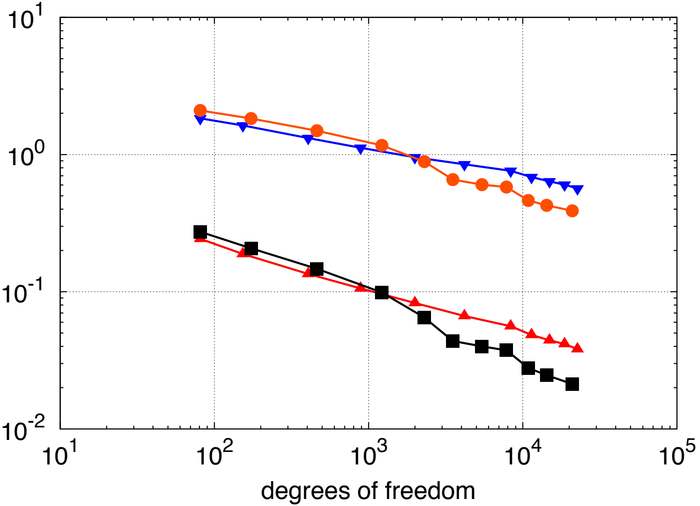




mesh after 5^{th}, 10^{th} iteration

Image source: MATLAB

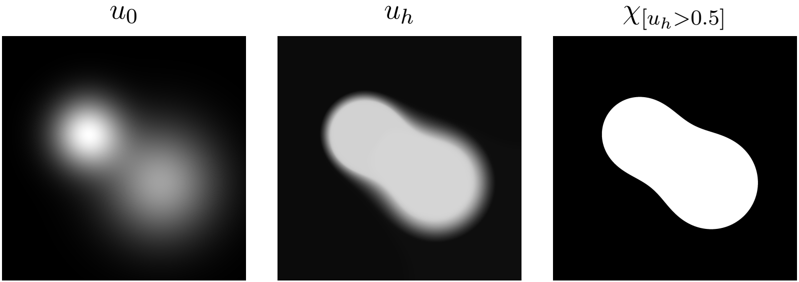
Numerical results - "Cameraman" (cont.)



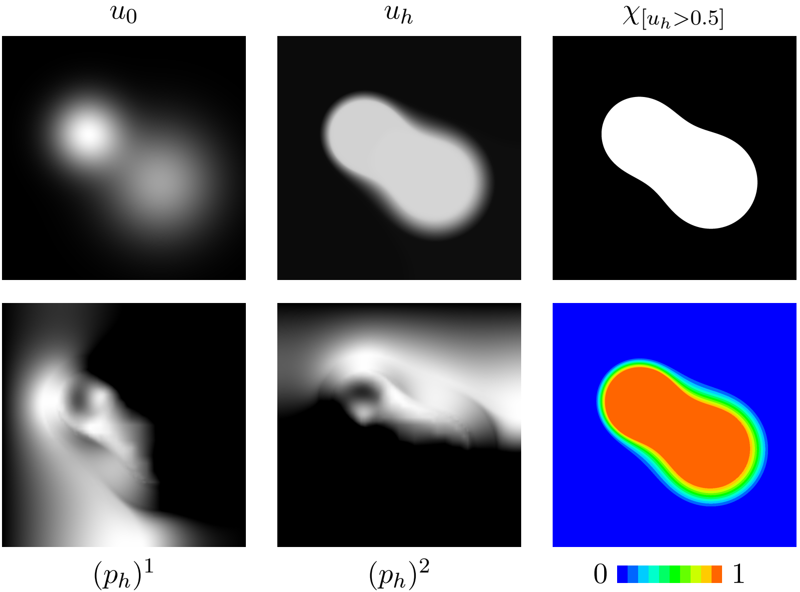
err_u^2 estimator for $\|u - u_h\|^2_{L^2(\Omega)}$ ((FE) (black), (FE') (red))

err_χ estimator for $\|\chi - \chi_{[u_h>0.5]}\|_{L^1(\Omega)}$ ((FE) (orange), (FE') (blue))

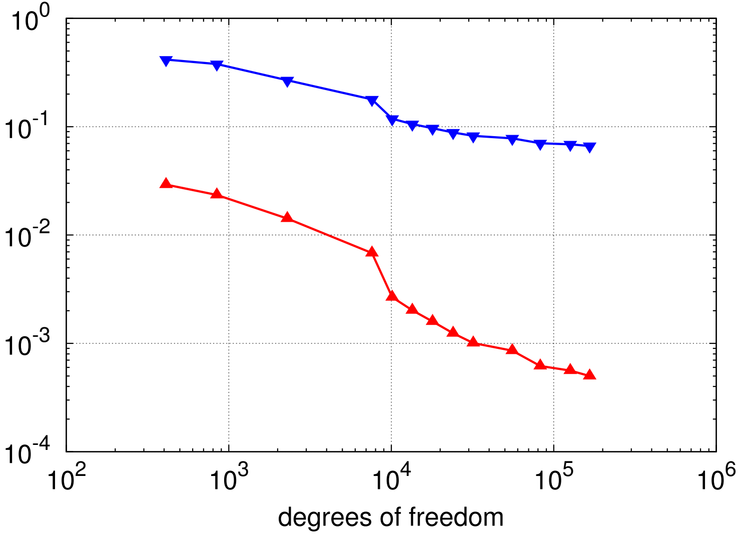
Numerical results - "Gaussians" using (FE')



Numerical results - "Gaussians" using (FE')



Numerical results - "Gaussians" using (FE')



err_u^2 estimator for $\|u - u_h\|^2_{L^2(\Omega)}$ err_χ estimator for $\|\chi - \chi_{[u_h > 0.5]}\|_{L^1(\Omega)}$

Conclusion and outlook

Summary

- A posteriori error estimates for the binary Mumford–Shah model

Conclusion and outlook

Summary

- A posteriori error estimates for the binary Mumford–Shah model
- Natural error quantity: area of the non-properly segmented region

Conclusion and outlook

Summary

- A posteriori error estimates for the binary Mumford–Shah model
- Natural error quantity: area of the non-properly segmented region
- Key ingredients:

Conclusion and outlook

Summary

- A posteriori error estimates for the binary Mumford–Shah model
- Natural error quantity: area of the non-properly segmented region
- Key ingredients:
  - Uniformly convex, unonstrained relaxation of the original problem

Conclusion and outlook

Summary

- A posteriori error estimates for the binary Mumford–Shah model
- Natural error quantity: area of the non-properly segmented region
- Key ingredients:
  - Uniformly convex, unonstrained relaxation of the original problem
  - Repin's theorem for a posteriori error estimation of the relaxation

Conclusion and outlook

Summary

- A posteriori error estimates for the binary Mumford–Shah model
- Natural error quantity: area of the non-properly segmented region
- Key ingredients:
  - Uniformly convex, unonstrained relaxation of the original problem
  - Repin's theorem for a posteriori error estimation of the relaxation
  - Cut out argument to estimate the area mismatch

Conclusion and outlook

Summary

- A posteriori error estimates for the binary Mumford–Shah model
- Natural error quantity: area of the non-properly segmented region
- Key ingredients:
  - Uniformly convex, unonstrained relaxation of the original problem
  - Repin's theorem for a posteriori error estimation of the relaxation
  - Cut out argument to estimate the area mismatch
- Estimate can be used for adaptive mesh refinement

Conclusion and outlook

Summary

- A posteriori error estimates for the binary Mumford–Shah model
- Natural error quantity: area of the non-properly segmented region
- Key ingredients:
  - Uniformly convex, unonstrained relaxation of the original problem
  - Repin's theorem for a posteriori error estimation of the relaxation
  - Cut out argument to estimate the area mismatch
- Estimate can be used for adaptive mesh refinement
- Two different adaptive primal-dual finite element schemes

Conclusion and outlook

Summary

- A posteriori error estimates for the binary Mumford–Shah model
- Natural error quantity: area of the non-properly segmented region
- Key ingredients:
  - Uniformly convex, unonstrained relaxation of the original problem
  - Repin's theorem for a posteriori error estimation of the relaxation
  - Cut out argument to estimate the area mismatch
- Estimate can be used for adaptive mesh refinement
- Two different adaptive primal-dual finite element schemes
- Numerical experiments show qualitative and quantitative properties

Conclusion and outlook

Summary

- A posteriori error estimates for the binary Mumford–Shah model
- Natural error quantity: area of the non-properly segmented region
- Key ingredients:
  - Uniformly convex, unonstrained relaxation of the original problem
  - Repin's theorem for a posteriori error estimation of the relaxation
  - Cut out argument to estimate the area mismatch
- Estimate can be used for adaptive mesh refinement
- Two different adaptive primal-dual finite element schemes
- Numerical experiments show qualitative and quantitative properties

Outlook

- Transfer the approach to more general computer vision problems

Conclusion and outlook

Summary

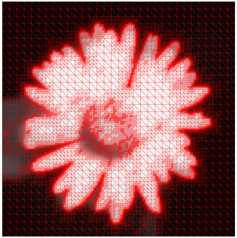
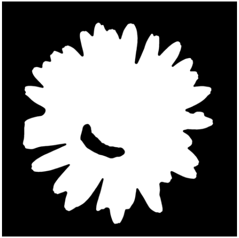
- A posteriori error estimates for the binary Mumford–Shah model
- Natural error quantity: area of the non-properly segmented region
- Key ingredients:
  - Uniformly convex, unonstrained relaxation of the original problem
  - Repin's theorem for a posteriori error estimation of the relaxation
  - Cut out argument to estimate the area mismatch
- Estimate can be used for adaptive mesh refinement
- Two different adaptive primal-dual finite element schemes
- Numerical experiments show qualitative and quantitative properties

Outlook

- Transfer the approach to more general computer vision problems
- Develop/use minimization algorithms that are tailored to the adaptive structure

References

- S. Bartels, Error control and adaptivity for a variational model problem defined on functions of bounded variation, Math. Comp. (online first), (2014).
- B. Berkels, A. Effland and M. Rumpf, A Posteriori Error Control for the Binary Mumford–Shah Model, Math. Comp. (accepted)
- B. Berkels, An unconstrained multiphase thresholding approach for image segmentation. In Proceedings of the Second SSVM, 2009.
- A. Chambolle and J. Darbon, On total variation minimization and surface evolution using parametric maximum flows, International Journal of Computer Vision, 84 (2009), pp. 288-307.
- A. Chambolle and T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, Journal of Mathematical Imaging and Vision, 40 (2011), pp. 120-145.
- M. Hintermüller and K. Kunisch, Total bounded variation regularization as a bilaterally constrained optimization problem, SIAM J. Appl. Math., 64 (2004), pp. 1311-1333.
- M. Nikolova, S. Esedoḡlu and T. F. Chan, Algorithms for finding global minimizers of image segmentation and denoising models, SIAM Journal on Applied Mathematics, 66 (2006), pp. 1632-1648.
- S. Repin, A posteriori error estimation for variational problems with uniformly convex functionals, Mathematics of Computation, 69 (2000), pp. 481-500.

Thank you for your attention!

Contact: `berkels@aices.rwth-aachen.de`