L^1 data fitting for nonlinear inverse problems

Christian Clason[1] Bangti Jin[2]

[1] Institute for Mathematics and Scientific Computing, Karl-Franzens-Universität Graz

[2] Department of Mathematics (IAMCS), Texas A&M University

Chemnitz Symposium on Inverse Problems
September 24, 2010

L^1 fitting problem

$$(\mathcal{P}) \qquad \min_{u \in L^2(\Omega)} \|S(u) - y^\delta\|_{L^1} + \frac{\alpha}{2}\|u\|_{L^2}^2$$

- $S : L^2(\Omega) \to Y \subset L^1(\Omega)$ **nonlinear** forward operator
- $y^\delta \in L^\infty(\Omega)$ noisy measurement, given
- $\alpha > 0$ regularization parameter
- $\Omega \subset \mathbb{R}^n$, $n = 1, 2, 3$, Lipschitz boundary $\partial\Omega$

L^1 fitting problem

$$(\mathcal{P}) \qquad \min_{u \in \mathrm{L}^2(\Omega)} \|S(u) - y^\delta\|_{\mathrm{L}^1} + \frac{\alpha}{2}\|u\|_{\mathrm{L}^2}^2$$

L^1 fitting more robust for non-Gaussian noise:

- large outliers
- Laplace-distributed noise
- impulsive noise (salt & pepper, random-valued)

$\Rightarrow$ Many applications in imaging

L^1 fitting problem

$$(\mathcal{P}) \qquad \min_{u\in \mathrm{L}^2(\Omega)} \|S(u) - y^\delta\|_{\mathrm{L}^1} + \frac{\alpha}{2}\|u\|_{\mathrm{L}^2}^2$$

Our interest: **Parameter identification problems for PDEs**

Main assumptions:

- $S : \mathrm{L}^2(\Omega) \to Y$ differentiable enough
- Y embeds compactly into $\mathrm{L}^2(\Omega)$

Goal: Fast Newton-type methods for L^1 fitting

Model problem

Consider $S : \mathrm{L}^2(\Omega) \to \mathrm{H}^1(\Omega)$, $S(u) =: y$ solution of

Elliptic BVP

$$\langle \nabla y, \nabla v \rangle_{\mathrm{L}^2} + \langle uy, v \rangle_{\mathrm{L}^2} = \langle f, v \rangle_{\mathrm{L}^2} \qquad \text{for all } v \in \mathrm{H}^1(\Omega)$$

for given $f \in \mathrm{L}^2(\Omega)$.

- Satisfies all assumptions
- Numerical results
- (Additional assumption: $u(x) \geq c > 0$, not enforced explicitly)

Well-posedness

Assumption (A1)

- S uniformly bounded in $L^2(\Omega)$
- $u_n \rightharpoonup u \implies S(u_n) \to S(u)$ in $L^2(\Omega)$

Theorem (Existence, consistency)

- *$(\mathcal{P})$ has solution u_α^δ*
- $y^n \to y^\delta \implies u_\alpha^n \to u_\alpha^\delta$ *(subsequence)*
- $\lim_{\delta\to 0} \alpha(\delta) = \lim_{\delta\to 0} \frac{\delta}{\alpha(\delta)} = 0 \implies u_{\alpha(\delta)}^\delta \to u^\dagger$ *(subsequence, $u^\dagger$ true parameter)*

Exact recovery

Assume

- S Fréchet differentiable
- S' Lipschitz continuous with constant $L > 0$
- There is $w \in \mathrm{L}^\infty(\Omega)$: $L\,\|w\|_{\mathrm{L}^2} \leq 1$, $u^\dagger - u_0 = S'(u^\dagger)^* w$

Then, for exact data $y^\dagger = S(u^\dagger)$ and $\alpha > 0$ small enough,

$$u_\alpha = u^\dagger$$

holds.

Compare with quadratic fitting term: $u_\alpha \neq u^\dagger$ for all $\alpha > 0$

Convergence rates

Assume

- S Fréchet differentiable
- S' Lipschitz continuous with constant $L > 0$
- There is $w \in \mathrm{L}^\infty(\Omega)$: $L\|w\|_{\mathrm{L}^2} \leq 1$, $u^\dagger - u_0 = S'(u^\dagger)^* w$

Then the following convergence rates hold:

- For **a priori choice** ($\alpha \approx \delta^\varepsilon$, $\varepsilon \in (0,1)$, $\|y^\dagger - y^\delta\|_{\mathrm{L}^1} \leq \delta$),

$$\|u_\alpha^\delta - u^\dagger\|_{\mathrm{L}^2} \leq C\delta^{\frac{1-\varepsilon}{2}}$$

- For **Morozov discrepancy principle** ($\|S(u_\alpha^\delta) - y^\delta\|_{\mathrm{L}^1} \approx \delta$),

$$\|u_\alpha^\delta - u^\dagger\|_{\mathrm{L}^2} \leq C\delta^{\frac{1}{2}}$$

Automatic parameter choice

Noise level δ unknown: choose α^* solving

Balancing equation

$$(\sigma - 1)\|S(u^\delta_{\alpha^*}) - y^\delta\|_{L^1} = \frac{\alpha^*}{2}\|u^\delta_{\alpha^*}\|^2_{L^2}$$

($\sigma > 1$ fixed)

Fixed point iteration

$$\alpha_{k+1} = (\sigma - 1)\frac{\|S(u^\delta_{\alpha_k}) - y^\delta\|_{L^1}}{\frac{1}{2}\|u^\delta_{\alpha_k}\|^2_{L^2}}.$$

Automatic parameter choice

Theorem (convergence)

If initial guess α_0 *satisfies*

$$(\sigma - 1)\|S(u^\delta_{\alpha_0}) - y^\delta\|_{L^1} - \frac{\alpha_0}{2}\|u^\delta_{\alpha_0}\|^2_{L^2} < 0$$

then $\{\alpha_k\}$

- *is monotonically decreasing*
- *converges to solution of balancing equation*

Further assumptions

There exists neighborhood U of u_α:

(A2) S is twice Fréchet differentiable

(A3) There exists $C > 0$ independent of u:

$$\left\|S'(u)h\right\|_{L^2} \leq C \left\|h\right\|_{L^2}$$

for all $u \in U$ and $h \in L^2(\Omega)$

(A4) There exists $C > 0$ independent of u:

$$\left\|S''(u)(h,h)\right\|_{L^2} \leq C \left\|h\right\|_{L^2}^2$$

all $u \in U$ and $h \in L^2(\Omega)$

Optimality conditions

$$(\mathcal{P}) \qquad \min_{u \in \mathrm{L}^2(\Omega)} \|S(u) - y^\delta\|_{\mathrm{L}^1} + \frac{\alpha}{2}\|u\|_{\mathrm{L}^2}^2$$

Theorem (Necessary optimality condition)

For any local minimizer $u_\alpha \in \mathrm{L}^2(\Omega)$ *of problem* $(\mathcal{P})$, *there exists a* $p_\alpha \in \mathrm{L}^\infty(\Omega)$ *such that*

$$(\text{OS}) \qquad \begin{cases} S'(u_\alpha)^* p_\alpha + \alpha u_\alpha = 0, \\ \langle S(u_\alpha) - y^\delta, p - p_\alpha \rangle_{\mathrm{L}^2} \leq 0 \quad \text{for all } \|p\|_{\mathrm{L}^\infty} \leq 1. \end{cases}$$

(S strictly diff., $\|\cdot\|_{\mathrm{L}^1}$ convex, real-valued $\Rightarrow$ $(\mathcal{P})$ is Lipschitz continuous)

Characterization of minimizer

For all $p \in L^\infty(\Omega)$, $p \geq 0$:

$$\begin{aligned} \langle S(u_\alpha) - y^\delta, p\rangle_{L^2} &= 0 && \text{if } \operatorname{supp} p \subset \{x : |p_\alpha(x)| < 1\}, \\ \langle S(u_\alpha) - y^\delta, p\rangle_{L^2} &\geq 0 && \text{if } \operatorname{supp} p \subset \{x : p_\alpha(x) = 1\}, \\ \langle S(u_\alpha) - y^\delta, p\rangle_{L^2} &\leq 0 && \text{if } \operatorname{supp} p \subset \{x : p_\alpha(x) = -1\}. \end{aligned}$$

Interpretation:

- Box constraint on p_α active where data is not attained by u_α
- Sign of p_α gives sign of noise
- $\Rightarrow$ p_α is **noise indicator**

Regularization

Problem:

- L^1 norm not strictly convex
- $\Rightarrow$ no local uniqueness (for p)
- $\Rightarrow$ no semi-smoothness

$\Rightarrow$ **Local smoothing**: consider for $\beta > 0$

Saddle point problem

$$(\mathcal{P}_\beta) \qquad \min_{u \in L^2(\Omega)} \max_{\|p\|_{L^\infty} \leq 1} \langle S(u) - y^\delta, p \rangle_{L^2} + \frac{\alpha}{2} \|u\|_{L^2}^2 - \frac{\beta}{2} \|p\|_{L^2}^2$$

Regularization

Saddle point problem

$$(\mathcal{P}_\beta) \qquad \min_{u\in L^2(\Omega)} \max_{\|p\|_{L^\infty}\le 1} \langle S(u)-y^\delta, p\rangle_{L^2} + \frac{\alpha}{2}\|u\|_{L^2}^2 - \frac{\beta}{2}\|p\|_{L^2}^2 .$$

Theorem

There exists a saddle point $(u_\beta, p_\beta) \in L^2(\Omega)\times L^\infty(\Omega)$ *of* $(\mathcal{P}_\beta)$ *satisfying*

$$(OS_\beta) \qquad \begin{cases} S'(u_\beta)^* p_\beta + \alpha u_\beta = 0, \\ \langle S(u_\beta) - y^\delta - \beta p_\beta, p - p_\beta\rangle_{L^2} \le 0. \end{cases}$$

for all $p \in L^\infty(\Omega)$ *with* $\|p\|_{L^\infty} \le 1$.

Regularization

Theorem (convergence)

As $\beta \to 0$, *the sequence of saddle points* (u_β, p_β) *has a subsequence converging to solution* (u_α, p_α) *of* (OS) *with*

$$\begin{aligned} u_\beta &\to u_\alpha && \text{in } L^2(\Omega) \\ p_\beta &\rightharpoonup^\star p_\alpha && \text{in } L^\infty(\Omega) \end{aligned}$$

⇒ **Continuation strategy:**

1. Solve for large β_k
2. Solve for smaller β_{k+1} , using $(u_{\beta_k}, p_{\beta_k})$ as initial guess
3. Repeat

Second order condition

S nonlinear $\Rightarrow$ require local convexity for (local) uniqueness:

Second order sufficient condition (SSC)

There exists $\gamma > 0$ such that

$$\langle S''(u_\beta)(h,h), p_\beta \rangle_{L^2} + \alpha \|h\|_{L^2}^2 \geq \gamma \|h\|_{L^2}^2$$

holds for all $h \in L^2(\Omega)$

Theorem (local uniqueness)

If SSC holds, the solution of (OS_β) is locally unique and is a strict saddle point of $(\mathcal{P}_\beta)$.

Second order condition

From (A4), we deduce

$$\left\langle S''(u_\beta)(h,h), p_\beta \right\rangle_{L^2} + \alpha \, \|h\|_{L^2}^2 \geq (\alpha - C\|p_\beta\|_{L^2}) \, \|h\|_{L^2}^2$$

Thus, SSC is satisfied if either

- α large or
- p_β small (p is noise indicator!)

$\Rightarrow$ reasonable assumption for parameter identification problems

Numerical solution

- Discretization by finite elements (u, p piecewise constant, $y = S(u)$ piecewise linear)
- Solution of (OS_β) by **semi-smooth Newton method** (SSC implies semi-smoothness)
- Results for model problem: $S : \mathrm{L}^2(\Omega) \to \mathrm{H}^1(\Omega)$, $S(u) =: y$ solution of

$$\langle \nabla y, \nabla v \rangle_{\mathrm{L}^2} + \langle uy, v \rangle_{\mathrm{L}^2} = \langle f, v \rangle_{\mathrm{L}^2}$$

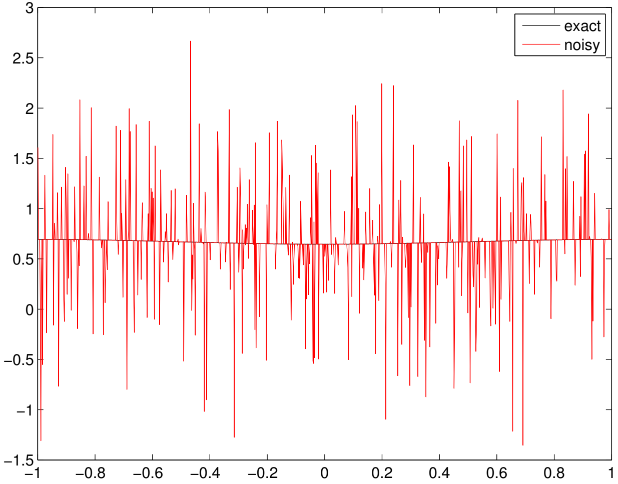
for all $v \in \mathrm{H}^1(\Omega)$ and given $f \in \mathrm{L}^2(\Omega)$.

(Satisfies Assumptions (A1)–(A4))

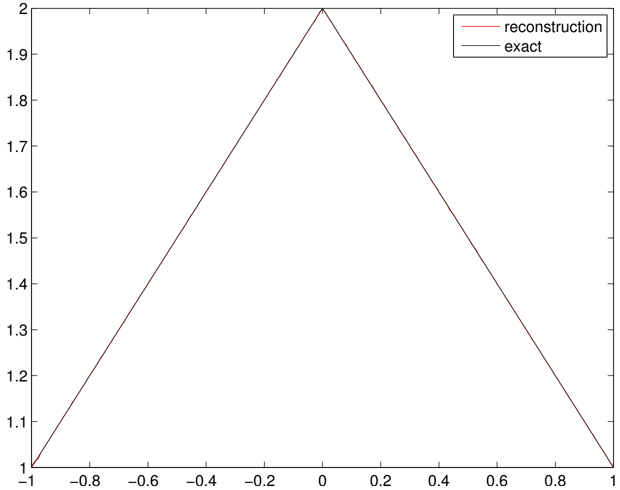
Numerical results (1D)

- $\Omega = [-1, 1]$, $f(x) = 1$, $N = 1000$ elements
- exact solution
$$u^\dagger = 2 - |x| \geq 1$$
- exact data $y^\dagger = S(u^\dagger)$, noisy data
$$y^\delta(x) = \begin{cases} y^\dagger(x) + \|y^\dagger\|_{L^\infty}\xi, & \text{with probability } r \\ y^\dagger(x), & \text{otherwise} \end{cases}$$
for all x, ξ normally distributed random variable
- α chosen using fixed point iteration (convergence in 4 its.)

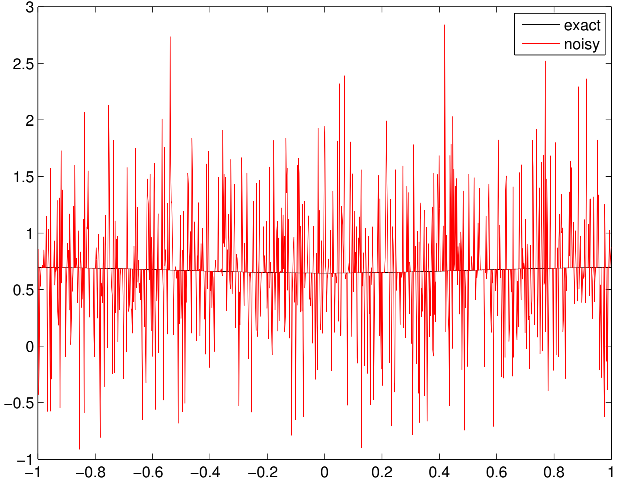
Data y^δ ($r = 0.3$)



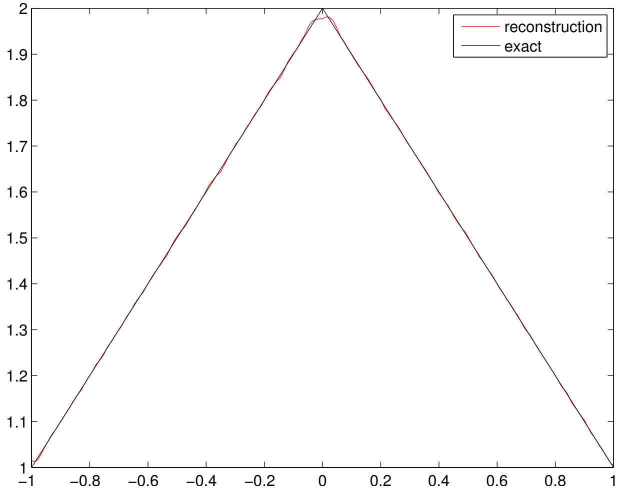
Reconstruction u^δ ($r = 0.3$, $\alpha = 4.781 \cdot 10^{-3}$)



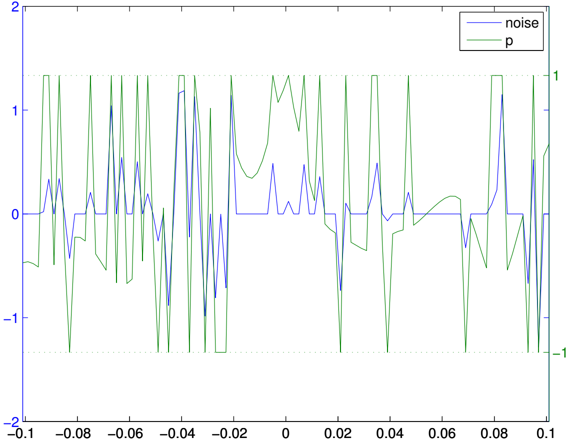
Data y^δ ($r = 0.6$)



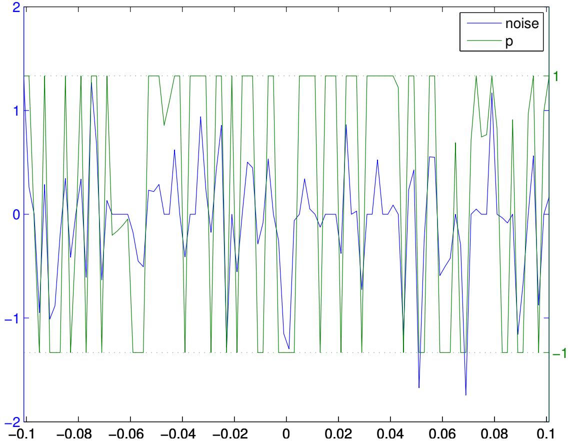
Reconstruction u^δ ($r = 0.6$, $\alpha = 8.236 \cdot 10^{-3}$)



Dual solution p^δ and noise ($r = 0.3$)



Dual solution p^δ and noise ($r = 0.6$)



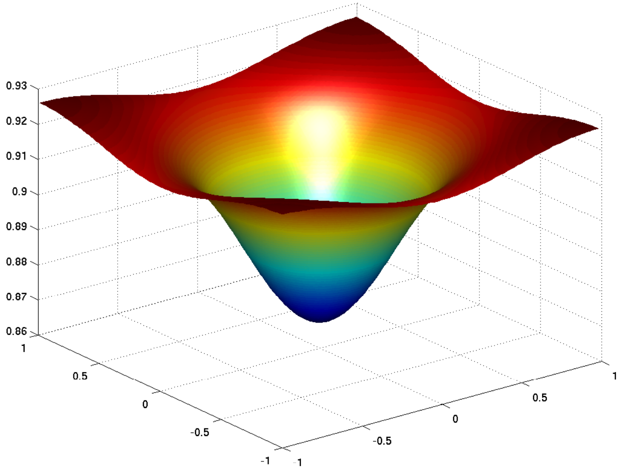
Numerical results (2D)

- $\Omega = [-1, 1]^2$
- standard uniform triangulation, $N = 32258$ elements
- exact solution

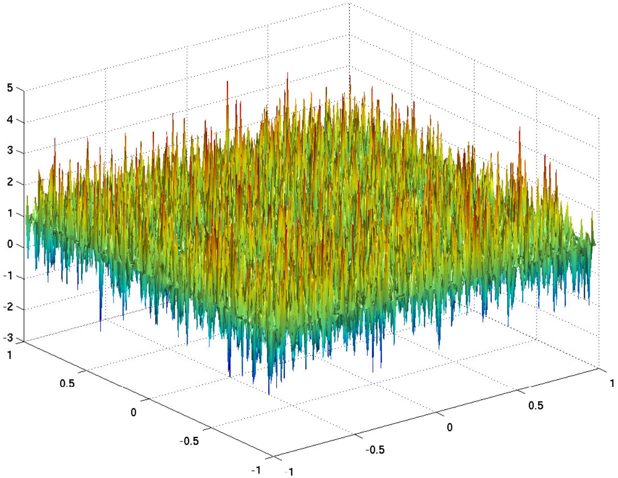
$$u^\dagger = \begin{cases} 1 + \cos(x_1\pi)\cos(x_2\pi), & \text{if } |x_1| \leq \frac{1}{2}, |x_2| \leq \frac{1}{2} \\ 1, & \text{otherwise} \end{cases}$$

- $f(x_1, x_2) = 1$
- exact data $y^\dagger = S(u^\dagger)$, noisy data y^δ
- α chosen using fixed point iteration (4 its.)

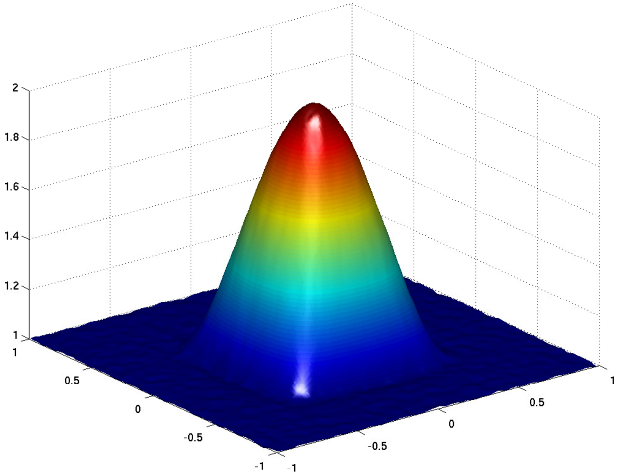
Exact data $y^\dagger$



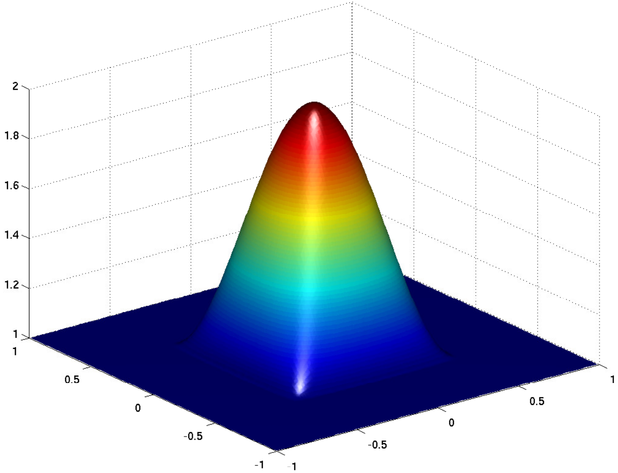
Noisy data y^δ ($r = 0.3$)



L^1 reconstruction u^δ ($r = 0.3$, $\alpha = 1.05 \cdot 10^{-2}$)



Exact solution $u^\dagger$



Conclusion

- L^1 fitting for nonlinear problems
- Very robust for impulsive noise
- Numerical solution by semi-smooth Newton method

Future work

- Time dependent problems (require efficient FE solvers)
- Higher-order coefficient identification problems (conductivity, wave-speed, . . .)

Preprint, MATLAB code:

`http://www.uni-graz.at/~clason/publications.html`