# $L^1$ data fitting for parameter identification problems for PDEs

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

Applied Inverse Problems Conference
College Station, Texas, May 23, 2011

# $\mathrm{L}^1$ fitting problem

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

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

# $\mathrm{L}^1$ fitting problem

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

$\mathrm{L}^1$ fitting more robust for non-Gaussian noise:

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

$\rightsquigarrow$ Many applications in imaging

# $\mathrm{L}^1$ fitting problem

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

Here: parameter identification problems for PDEs

Main assumptions:

- $S : \mathcal{X} \to \mathcal{Y}$ sufficiently differentiable
- $\mathcal{X}$ Hilbert space (e.g., $\mathrm{L}^2$, $\mathrm{H}^1$)
- $\mathcal{Y}$ embeds compactly into $\mathrm{L}^q$, $q > 2$

Goal: Fast Newton-type methods for $\mathrm{L}^1$ fitting

# Elliptic model problems

1. Inverse potential problem: $S : \mathrm{L}^2(\Omega) \to \mathrm{H}^1(\Omega)$, $u \mapsto y$,

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

2. Inverse Robin problem: $S : \mathrm{L}^2(\Gamma_i) \to \mathrm{H}^{1/2}(\Gamma_c)$, $u \mapsto y|_{\Gamma_c}$,

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

3. Inverse conductivity problem, $S : \mathrm{H}^1(\Omega) \cap \mathrm{L}^\infty(\Omega) \to \mathrm{H}^1_0(\Omega)$, $u \mapsto y$,

$$\langle u \nabla y, \nabla v \rangle_{\mathrm{L}^2} = \langle f, v \rangle_{\mathrm{L}^2} \qquad \text{for all } v \in \mathrm{H}^1_0(\Omega)$$

# Common properties

(A1) $S$ uniformly bounded in $\mathcal{X}$, $u_n \rightharpoonup u$ in $\mathcal{X}$ implies

$$S(u_n) \to S(u) \quad \text{in } L^2(\Omega)$$

(A2) $S$ twice Fréchet differentiable

(A3) For all $u, h \in \mathcal{X}$,

$$\begin{aligned} \|S'(u)h\|_{L^2} &\leq C\|h\|_{\mathcal{X}} \\ \|S''(u)(h,h)\|_{L^2} &\leq C\|h\|_{\mathcal{X}}^2 \end{aligned}$$

$\rightsquigarrow$ sufficient conditions for approach, existence of minimizers $u_\alpha$

# Optimality conditions

## $L^1$ fitting problem

$$\min_u \left\{ \mathcal{J}_\alpha \equiv \mathcal{F}(u) + \mathcal{G}(S(u) - y^\delta) \right\}$$

with

$$\mathcal{F} : \mathcal{X} \to \mathbb{R}, \qquad u \mapsto \frac{\alpha}{2} \|u\|_{\mathcal{X}}^2,$$
$$\mathcal{G} : L^1(\Omega) \to \mathbb{R}, \quad v \mapsto \|v\|_{L^1},$$

Problem: $S$ nonlinear, Fenchel duality not applicable

But $S$ strictly diff., $\mathcal{G}$ convex, real-valued $\Rightarrow \mathcal{J}_\alpha$ is Lipschitz

# Optimality conditions

$\mathcal{J}_\alpha$ Lipschitz $\Rightarrow$ sum rule, chain rule for generalized gradient:

$$\partial\mathcal{J}_\alpha(u) = \mathcal{F}'(u) + S'(u)^*\partial\mathcal{G}(S(u) - y^\delta)$$

Thus: The necessary condition for any local minimizer $u_\alpha$,

$$0 \in \partial\mathcal{J}_\alpha(u_\alpha),$$

implies existence of $p_\alpha \in \partial\mathcal{G}(S(u_\alpha) - y^\delta) \subset L^\infty(\Omega)$ with

$$(\mathrm{OS}_1) \qquad 0 = \alpha j(u_\alpha) + S'(u_\alpha)^* p_\alpha$$

# Optimality conditions

$\mathcal{G} = \|\cdot\|_{L^1}$ convex:

$$p_\alpha \in \partial\mathcal{G}(S(u_\alpha) - y^\delta) \quad \Longleftrightarrow \quad S(u_\alpha) - y^\delta \in \partial\mathcal{G}^*(p_\alpha)$$

with Fenchel conjugate

$$\mathcal{G}^*(p) = I_{\{\|p\|_{L^\infty} \leq 1\}} := \begin{cases} 0 & \|p\|_{L^\infty} \leq 1 \\ \infty & \text{else} \end{cases}$$

$$v \in \partial\mathcal{G}^*(p) \Leftrightarrow \langle v, q - p\rangle_{L^{\infty *}, L^\infty} \leq 0$$

for all $q \in L^\infty(\Omega)$ with $\|q\|_{L^\infty} \leq 1$

# Optimality conditions

$\mathcal{G} = \|\cdot\|_{L^1}$ convex:

$$p_\alpha \in \partial\mathcal{G}(S(u_\alpha) - y^\delta) \quad \Longleftrightarrow \quad S(u_\alpha) - y^\delta \in \partial\mathcal{G}^*(p_\alpha)$$

Thus, $p_\alpha \in \partial\mathcal{G}(S(u_\alpha) - y^\delta)$ iff

$$\langle S(u_\alpha) - y^\delta, p - p_\alpha\rangle_{L^2} \leq 0 \tag{OS$_2$}$$

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

(Note $S(u_\alpha) - y^\delta \in L^2(\Omega)$ by assumption)

# Optimality conditions

## Theorem

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

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

($j$: duality mapping in $\mathcal{X}$)

# Optimality conditions

**Theorem**

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

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

Complementarity function for variational inequality: for any $c > 0$,

$$\begin{aligned} S(u_\alpha) - y^\delta = &\max(0, S(u_\alpha) - y^\delta + c(p_\alpha - 1)) \\ &+ \min(0, S(u_\alpha) - y^\delta + c(p_\alpha + 1)) \end{aligned}$$

# Optimality conditions

Complementarity function for variational inequality: for any $c > 0$,

$$S(u_\alpha) - y^\delta = \max(0, S(u_\alpha) - y^\delta + c(p_\alpha - 1)) + \min(0, S(u_\alpha) - y^\delta + c(p_\alpha + 1))$$

## Pointwise interpretation:

$$p_\alpha = \operatorname{sign}(S(u_\alpha) - y^\delta) = \begin{cases} 1 & S(u_\alpha) - y^\delta > 0 \\ -1 & S(u_\alpha) - y^\delta < 0 \\ \tau \in [-1, 1] & S(u_\alpha) - y^\delta = 0 \end{cases}$$

# Optimality conditions

## Pointwise interpretation:

$$p_\alpha = \operatorname{sign}(S(u_\alpha) - y^\delta) = \begin{cases} 1 & S(u_\alpha) - y^\delta > 0 \\ -1 & S(u_\alpha) - y^\delta < 0 \\ \tau \in [-1, 1] & S(u_\alpha) - y^\delta = 0 \end{cases}$$

## Reduced optimality system

(OS')
$$\alpha j(u_\alpha) + S'(u_\alpha)^*(\operatorname{sign}(S(u_\alpha) - y^\delta)) \ni 0$$

# Regularization

sign not differentiable in any sense $\rightsquigarrow$ replace by $\text{sign}_\beta$ for $\beta > 0$,

$$\text{sign}_\beta(u)(x) := \begin{cases} 1 & \text{if } u(x) > \beta, \\ -1 & \text{if } u(x) < -\beta, \\ \frac{1}{\beta} t & \text{if } |u(x)| \leq \beta, \end{cases}$$

(equivalent to Huber-smoothing of $(\mathcal{P})$, dual $L^2$ regularization)

## Regularized optimality system

$$(\text{OS}_\beta) \qquad \alpha j(u_\beta) + S'(u_\beta)^*(\text{sign}_\beta(S(u_\beta) - y^\delta)) = 0$$

# Regularization

## Regularized optimality system

$$(\mathrm{OS}_\beta) \qquad \alpha j(u_\beta) + S'(u_\beta)^*(\mathrm{sign}_\beta(S(u_\beta) - y^\delta)) = 0$$

## Theorem

$(\mathrm{OS}_\beta)$ *has a solution* $u_\beta$*, and sequence* $\{u_\beta\}_{\beta>0}$ *contains subsequence converging in* $\mathcal{X}$ *to solution* $u_\alpha$ *to* (OS').

$\rightsquigarrow$ Continuation strategy in $\beta \to 0$ for numerical solution

# Semi-smooth Newton method

Consider ($OS_\beta$) as $F(u) = 0$ for $F : \mathcal{X} \to \mathcal{X}^*$,

$$F(u) = \alpha j(u) + S'(u)^*(\text{sign}_\beta(S(u) - y^\delta))$$

$t \mapsto \text{sign}_\beta(t)$ semi-smooth, $S(u) - y^\delta \in L^q$, $S$ twice differentiable

$\Rightarrow P(u) = \text{sign}_\beta(S(u) - y^\delta)$ semi-smooth, Newton derivative

$$\begin{aligned} D_N P(u)h &= \beta^{-1}(S'(u)h)\chi_{\mathcal{I}} \\ &= \begin{cases} \beta^{-1}(S'(u)h) & \text{if } |(S(u) - y^\delta)| \leq \beta \\ 0 & \text{else} \end{cases} \end{aligned}$$

# Semi-smooth Newton method

$\mathcal{X}$ Hilbert space, $S'(u)$ linear operator $\Rightarrow$ $F$ semi-smooth

**Semi-smooth Newton step for $\delta u = u^{k+1} - u^k$**

$$\alpha j'(u^k)\delta u + (S''(u^k)\delta u)^* P(u^k) + \frac{1}{\beta} S'(u^k)^*(\chi_{\mathcal{I}^k} S'(u^k)\delta u) = -F(u^k)$$

Can be solved using matrix-free Krylov method
(given $u^k$, $\delta u$, rhs/lhs computed by solving forward, adjoint PDE)

# Semi-smooth Newton method

But: superlinear convergence requires regularity condition, $S$ nonlinear, functional not necessarily convex $\rightsquigarrow$ assume for $\gamma > 0$

**Second order condition**

(S) $\langle S''(u_\beta)(h,h), P(u_\beta)\rangle_{L^2} + \alpha\|h\|_{\mathcal{X}}^2 \geq \gamma\|h\|_{\mathcal{X}}^2$ for all $h \in \mathcal{X}$

(compare second order sufficient condition)

Here: (S) holds if either

- $\alpha$ large (large noise)
- $\beta$ large or residual small (small noise) ($\Rightarrow P(u_\beta)$ small)

# Semi-smooth Newton method

## Second order condition

$$\text{(S)} \qquad \langle S''(u_\beta)(h,h), P(u_\beta)\rangle_{L^2} + \alpha\|h\|_{\mathcal{X}}^2 \geq \gamma\|h\|_{\mathcal{X}}^2 \quad \text{for all } h \in \mathcal{X}$$

## Theorem

*If* (S) *holds and* $u^0$ *is sufficiently close to* $u_\beta$*, then the iterates of the semi-smooth Newton method converge superlinearly to the solution* $u_\beta$ *to* ($OS_\beta$).

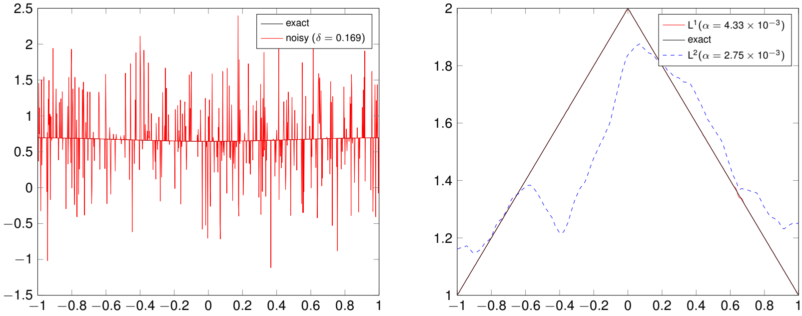
# Numerical results for model problems

- Discretization using uniform linear finite elements
  1d: $N = 1001$, 2d: $N = 128 \times 128$ grid points
- Random impulsive noise: $y^\dagger = S(u^\dagger)$,

$$y^\delta = \begin{cases} y^\dagger + \|y^\dagger\|_{L^\infty}\xi, & \text{with probability } r \\ y^\dagger, & \text{otherwise} \end{cases}$$

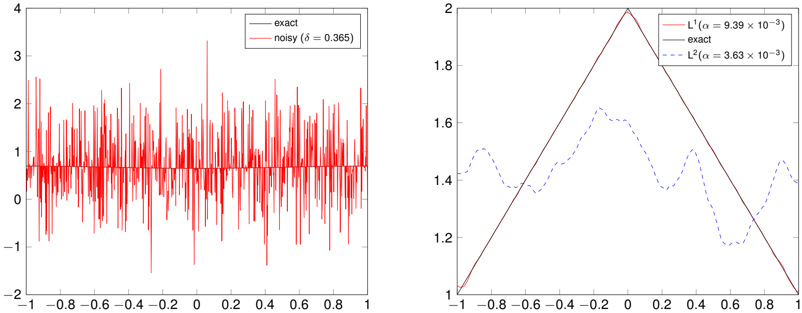
  $\xi(x)$ normally distributed random variable
- $\alpha$ chosen using fixed point iteration (2–4 its.)
- Comparison with standard $L^2$ fitting (Newton method)

# Inverse potential: $r = 0.3$




(a) data (b) reconstruction

# Inverse potential: $r = 0.6$




(a) data

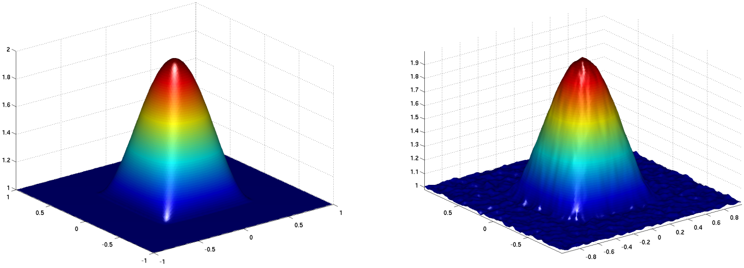
(b) reconstruction

# Inverse potential: Performance

| $N$ | 400 | 800 | 1600 | 3200 | 6400 | 12800 |
|---|---|---|---|---|---|---|
| $t_s$ | 5.28 | 12.09 | 19.40 | 29.66 | 55.33 | 107.87 |
| $t_b$ | 14.42 | 39.04 | 54.19 | 80.30 | 131.72 | 234.00 |
| $e$ | 2.88e-3 | 9.17e-4 | 6.22e-4 | 3.52e-4 | 2.76e-4 | 2.78e-4 |

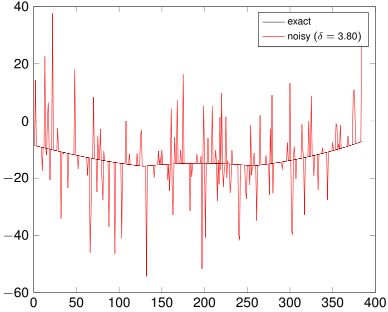
- $N$: number of elements
- $t_s$: computing time for semi-smooth Newton method including continuation in $\beta$ (seconds, average of 10)
- $t_b$: computing time for fixed point iteration (choice of $\alpha$)
- $e$: $L^2$ reconstruction error (average of 10)

# Inverse potential (2d): $r = 0.3$

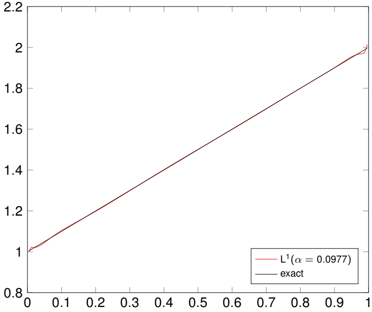



(a) exact (b) reconstruction

# Inverse Robin: $r = 0.3$

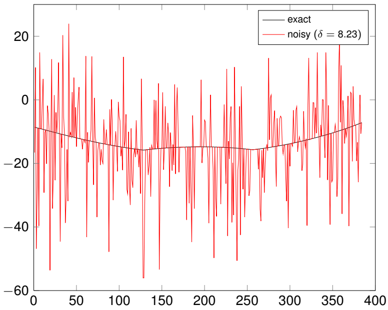



(a) data

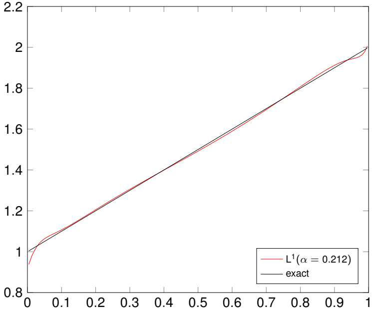



(b) reconstruction

# Inverse Robin: $r = 0.6$

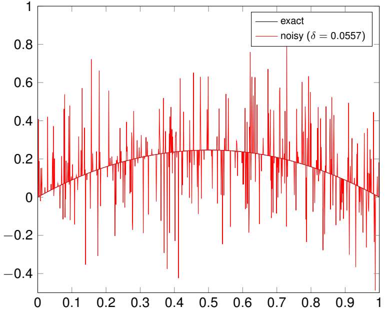



(a) data

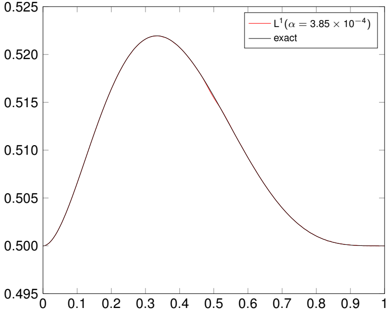



(b) reconstruction

# Inverse conductivity: $r = 0.3$

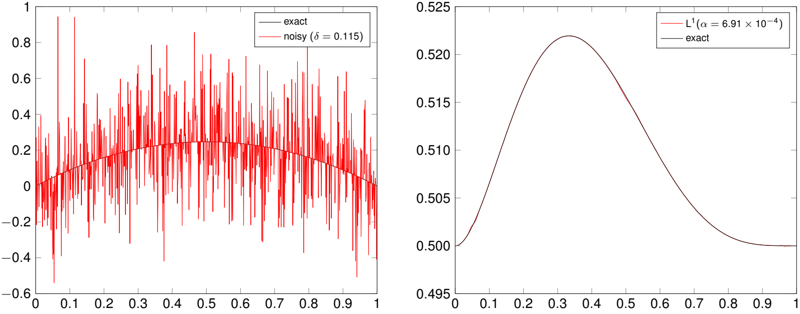



(a) data



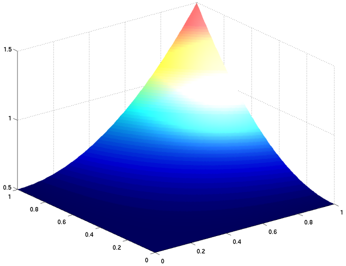

(b) reconstruction

# Inverse conductivity: $r = 0.6$

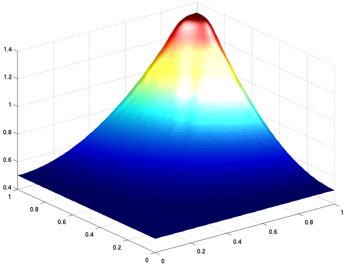



(a) data (b) reconstruction

# Inverse conductivity (2d): $r = 0.3$




(a) exact




(b) reconstruction

# Conclusion

- Semi-smooth Newton methods for numerical solution of non-smooth (Lipschitz) problems
- $L^1$ fitting very robust for impulsive noise

**Future work**

- Time dependent problems (require efficient FE solvers)
- Applications (magnetic induction, diffuse optical tomography)

Preprint, MATLAB code:

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