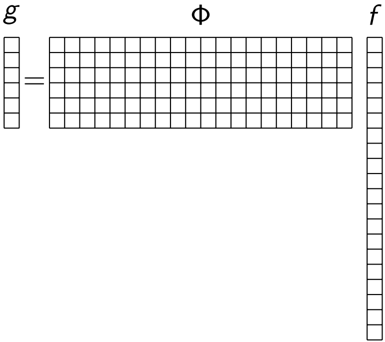
# Semismooth Newton and Active Set Methods for Sparse Reconstruction

Dirk Lorenz

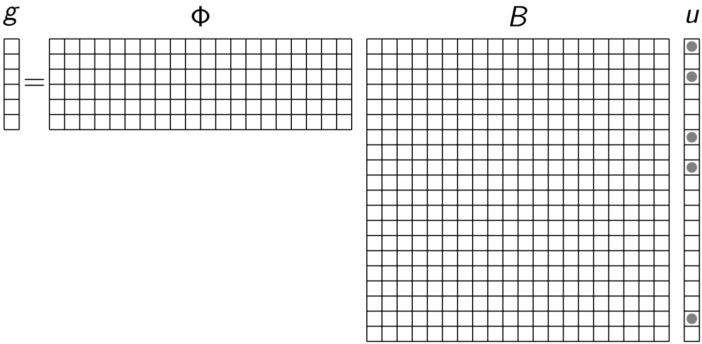
SIAM Conference on Imaging Science

San Diego, 2008-07-07

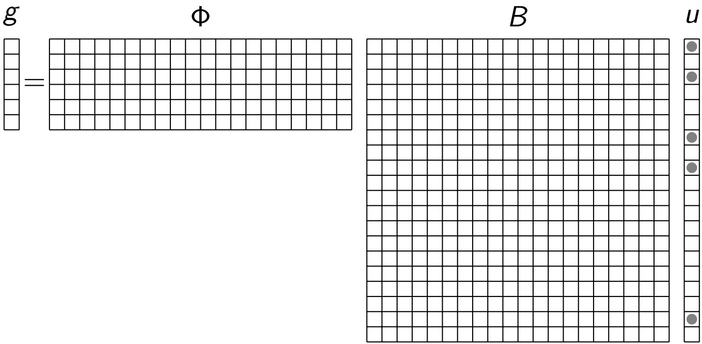
# Sparse reconstrucion



# Sparse reconstrucion



# Sparse reconstrucion

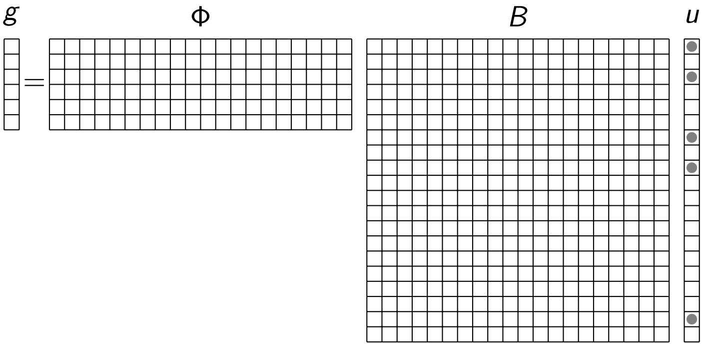



Reconstruct $u$ by

$$\min \|u\|_1 \quad \text{s.t.} \quad g = \Phi B u.$$

[Candès, Tao, Donoho,. . . ]

# Sparse reconstrucion



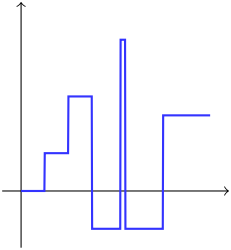

Reconstruct $u$ by

$$\min \|u\|_1 + \lambda \|\Phi B u - g\|^2.$$

[Candès, Tao, Donoho,...]

# Inverse ill-posed problems

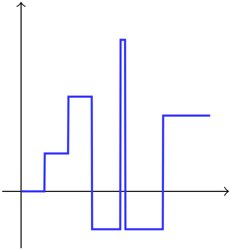
$$Af(x) = \int k(x,y)f(y)dy$$



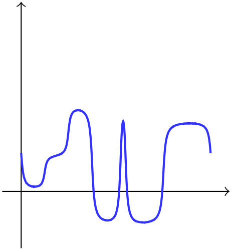
$f^\dagger$

# Inverse ill-posed problems

$$Af(x) = \int k(x, y) f(y) dy$$



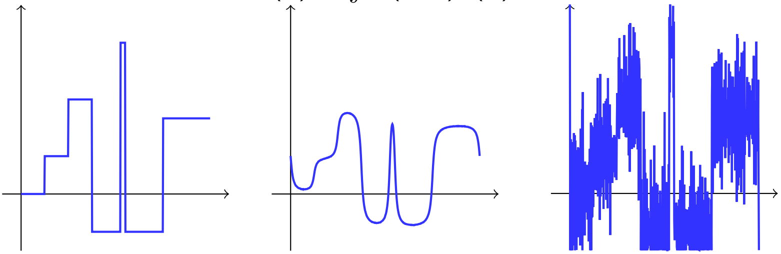
$f^\dagger$



$g^\delta = Af^\dagger + \text{noise}$

# Inverse ill-posed problems

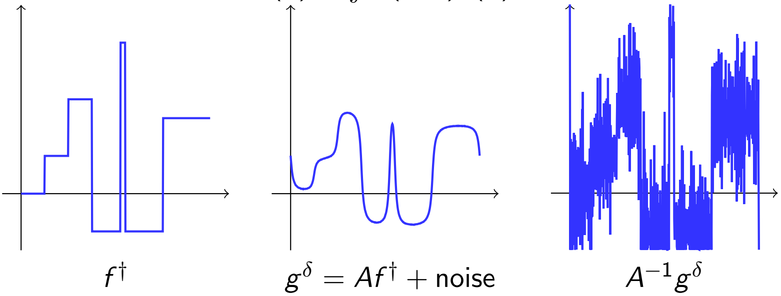
$$Af(x) = \int k(x,y)f(y)dy$$



$f^\dagger$ $\qquad$ $g^\delta = Af^\dagger + \text{noise}$ $\qquad$ $A^{-1}g^\delta$

# Inverse ill-posed problems

$$Af(x) = \int k(x,y)f(y)dy$$




If $f^\dagger$ has a sparse representation, $f^\dagger = Bu$, stabilize by

$$\min \tfrac{1}{2}\|ABu - g^\delta\|^2 + \alpha\|u\|_1.$$

[Daubechies, Defrise, DeMol,...]

# Same functionals but different issues

### Sparse reconstruction

- Large null-space
- Heavy non-uniqueness
- Well behaved singular values

$\ell^1$ makes solution unique

### Ill posed problmes

- No problem with null-space
- Ugly singular values
- Large instability

$\ell^1$ stabilizes the problem

# Stabilization with $\ell^1$ constraints

- $u^{\alpha,\delta} \in \operatorname{argmin} \frac{1}{2}\|Ku - g^\delta\|^2 + \alpha\|u\|_1$
- $\|Ku^\dagger - g^\delta\| \leq \delta$
- $K$ fulfills finite basis injectivity property
- $u^\dagger \in \operatorname{rg} K^*$

# Stabilization with $\ell^1$ constraints

- $u^{\alpha,\delta} \in \operatorname{argmin} \frac{1}{2}\|Ku - g^\delta\|^2 + \alpha\|u\|_1$
- $\|Ku^\dagger - g^\delta\| \leq \delta$
- $K$ fulfills finite basis injectivity property
- $u^\dagger \in \operatorname{rg} K^*$

**Theorem (L. 08)**

$$\|u^{\alpha,\delta} - u^\dagger\|_1 \leq \frac{\delta + \alpha\rho}{\sqrt{\lambda\alpha(1-\alpha)}} = \mathcal{O}(\sqrt{\delta})$$

# Stabilization with $\ell^1$ constraints

- $u^{\alpha,\delta} \in \operatorname{argmin} \frac{1}{2}\|Ku - g^\delta\|^2 + \alpha\|u\|_1$
- $\|Ku^\dagger - g^\delta\| \leq \delta$
- $K$ fulfills finite basis injectivity property
- $u^\dagger \in \operatorname{rg} K^*$

### Theorem (GHS 08)

$$\|u^{\alpha,\delta} - u^\dagger\|_2 \leq \frac{(\delta + \alpha\rho)^2}{\lambda\alpha} = \mathcal{O}(\delta)$$

*(even for nonlinear $K$).*

# Algorithms for $\ell^1$ minimization

- Iterated soft thresholding and extensions [DDD04, CW05]
- GPSR [FNW 07]
- l1_ls [KKLBG 07]
- LARS [OPT 00, EHJT 04, L 07]
- Iterated hard thresholding [BL 08]
- ...

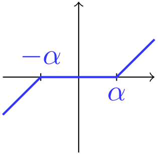
# Optimalty conditions

$$u^* \in \operatorname{argmin}\ \tfrac{1}{2}\|Ku - g\|^2 + \alpha \sum |u_k|$$
$$\Leftrightarrow u^* = \mathbf{S}_{\gamma\alpha}(u^* - \gamma K^*(Ku^* - g)) \text{ for any } \gamma > 0$$

# Optimalty conditions

$$u^* \in \operatorname{argmin} \; \tfrac{1}{2}\|Ku - g\|^2 + \alpha \sum |u_k|$$
$$\Leftrightarrow u^* = \mathbf{S}_{\gamma\alpha}(u^* - \gamma K^*(Ku^* - g)) \text{ for any } \gamma > 0$$

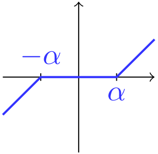
- $\mathbf{S}_\alpha(u)_k = S_\alpha(u_k)$
- Fixed point iteration
  $\rightsquigarrow$ iterative soft-thresholding [DDD04]
  $\rightsquigarrow$ converges linearly [BL08]
  $\rightsquigarrow$ slow...



# Optimalty conditions

$$u^* \in \operatorname{argmin} \ \tfrac{1}{2}\|Ku - g\|^2 + \alpha \sum |u_k|$$
$$\Leftrightarrow u^* = \mathbf{S}_{\gamma\alpha}(u^* - \gamma K^*(Ku^* - g)) \text{ for any } \gamma > 0$$

- $\mathbf{S}_\alpha(u)_k = S_\alpha(u_k)$
- Fixed point iteration
  $\rightsquigarrow$ iterative soft-thresholding [DDD04]
  $\rightsquigarrow$ converges linearly [BL08]
  $\rightsquigarrow$ slow... very slow



# Semismoothness of the optimality condition

The optimality condition

$$\mathcal{F}(u^*) = u^* - \mathbf{S}_{\gamma\alpha}(u^* - \gamma K^*(Ku^* - g)) = 0$$

is Lipschitz.

$\Longrightarrow$ $\mathcal{F}$ is semismooth aka Newton-differentiable [CNQ 00]

$\Longrightarrow$ Newton's method may be applicable

$$u^{n+1} = u^n - \mathcal{F}'(u^n)^{-1}\mathcal{F}(u^n)$$

# The Newton step can be calculated

$$\mathcal{F}(u) = u - \mathbf{S}_{\gamma\alpha}(u - \gamma K^*(Ku - g))$$

- The *active set* is

$$\mathcal{A} = \{k \in \mathbb{N} : |u^{n-1} - \gamma K^*(Ku^{n-1} - g)|_k > \gamma\alpha\}.$$

- The Newton-derivative of $\mathcal{F}$ is

$$\mathcal{F}'(u) = I - \begin{pmatrix} I_{\mathcal{A}} & 0 \\ 0 & 0 \end{pmatrix} (I - \gamma K^* K) = \begin{pmatrix} M_1 & M_2 \\ 0 & I \end{pmatrix}$$

- $M_1$ is a *finite dimensional square matrix*
- FBI $\implies$ $\mathcal{F}'(u)$ nonsingular.

# The SSN as an Active Set method

Calculate the *positive and negative active set* as

$$\mathcal{A}_n^+ = \{k \in \mathbb{N} : [u^{n-1} - \gamma K^*(Ku^{n-1} - g)]_k > \gamma\alpha\}$$
$$\mathcal{A}_n^- = \{k \in \mathbb{N} : [u^{n-1} - \gamma K^*(Ku^{n-1} - g)]_k < -\gamma\alpha\}.$$

and set

$$s_k^n = \begin{cases} 1, & k \in \mathcal{A}_n^+ \\ -1, & k \in \mathcal{A}_n^- \\ 0, & \text{else} \end{cases}.$$

Then the Newton step reads as

$$(K|_{\mathcal{A}_n})^* K|_{\mathcal{A}_n} u_{\mathcal{A}_n}^n = (K^* g - s^n \alpha)_{|\mathcal{A}_n}, \qquad u_k^n = 0, \text{ else}.$$

# Convergence of the SSN method

**Theorem**

*Let $K$ fulfill the FBI property and let $\gamma > 0$. Then the semismooth Newton method applied to the functional*

$$\frac{1}{2}\|Ku - g\|^2 + \alpha \sum |u_k|$$

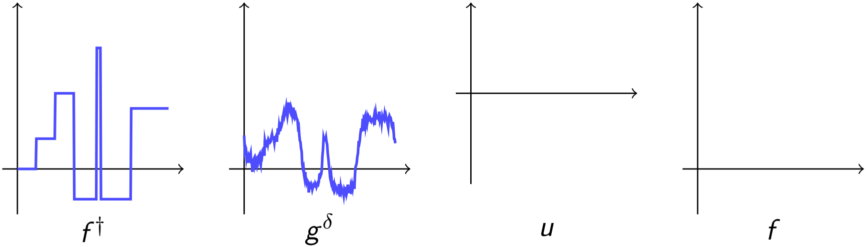
*converges locally superlinearly to a minimizer.*

*[Griesse, L. 08]*

- Convergence depends a lot on $\gamma$.
- Convergence is robust with respect to initial value.
- Moreover: Convergence in finitely many steps.

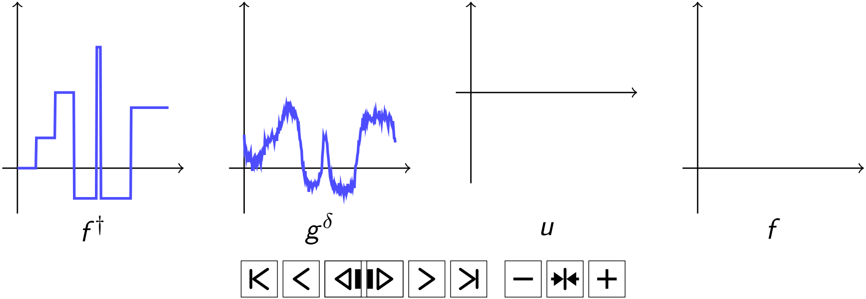
# Illustration for $\ell^1$-minimization

## Iterative thresholding

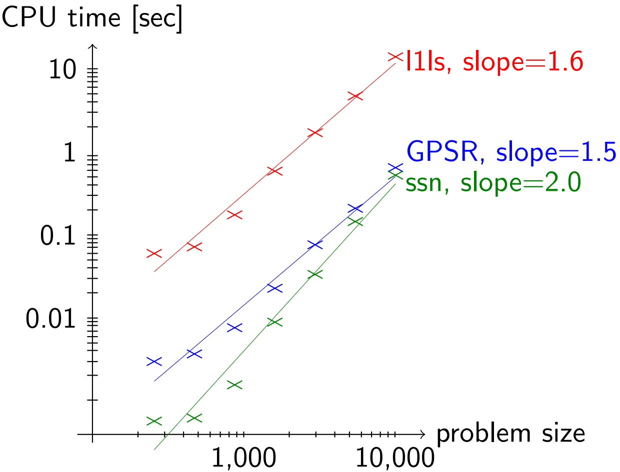


# Illustration for $\ell^1$-minimization

## SSN method

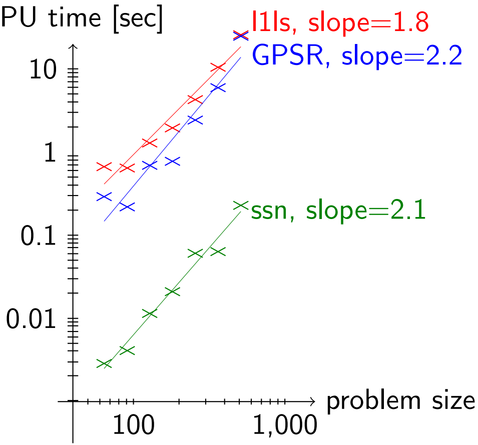


# Scaling: Compressed sensing



# Scaling: Ill posed problem of inverse integration

$$Af(x) = \int_0^x f(t)dt$$



# Globalization

1
- $\|\mathcal{F}(u^n)\|^2$ is a merit function
- $\mathcal{F}(u^n)^{-1}\mathcal{F}(u^n)$ is a descent direction for $\|\mathcal{F}\|^2$

$\rightsquigarrow$ use line-search that guarantees enough descent?

2 Reformulate as

$$u^+ = \max(u,0),\ \ u^- = \max(-u,0)$$
$$u = u^+ - u^-, \quad \bar{K} = [K\ -K]$$
$$\min \tfrac{1}{2}\|\bar{K}v - g\|^2 + \alpha \sum v_k \quad s.t.\ v \geq 0$$

and use SSN.

# Be more careful with the active set

## SSN:

$$\mathcal{A}^n = \{k \in \mathbb{N} : |u^{n-1} - \gamma K^*(Ku^{n-1} - g)|_k > \gamma\alpha\}.$$

Few iterations, local convergence, potentially large active sets

## Feature-sign search [LBRN06]:

$$\mathcal{A}^n = \mathcal{A}^{n-1} \cup \{k\} \quad \text{where} \quad k = \operatorname{argmax}\{|K^*(Ku^{n-1} - g)|_k - \alpha\}$$

$+$another step to reduce the active set if necessary.

Many iterations, global convergence, no error estimate available

# Conclusion and outlook

- $\ell^1$ minimization can also be used for the stabilization of ill-posed problem.
- SSN is a promising tool for $\ell^1$ minimization.
- Some work has to be done to make SSN "general purpose".