

# New Parameter Choice Rules for Regularization with Mixed Gaussian and Poissonian Noise

Elias Helou (joint work with Álvaro De Pierro)  
ICMC - USP  
[elias@icmc.usp.br](mailto:elias@icmc.usp.br)

1 de agosto de 2013

# Sumário

## Regularization

Conditioning and Regularization

A Deterministic Approach

A Frequentist Approach

## Frequentist Approaches to Parameter Selection

Existing Methods

The New Technique

## Numerical Results (preliminary)

Tomography

The (Still Preliminary) Results

Concluding Remarks

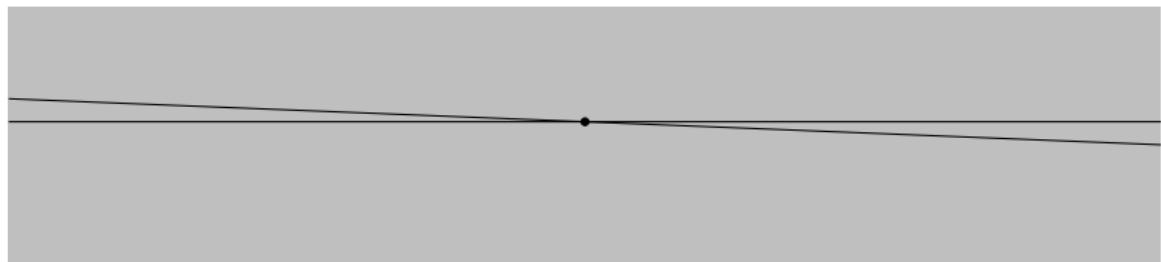
# Ill-Conditioned Linear System



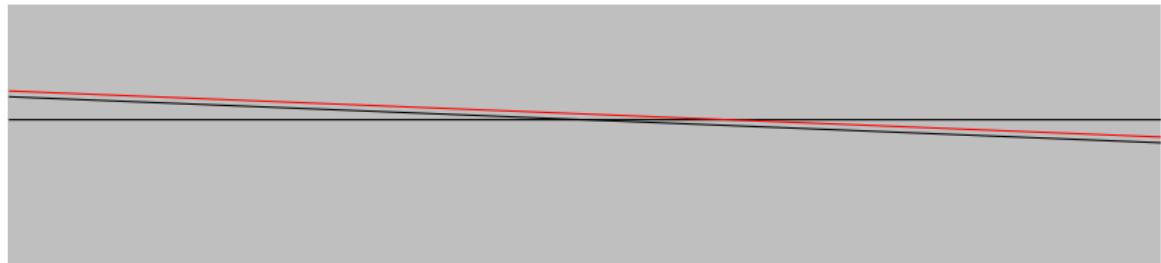
# Ill-Conditioned Linear System



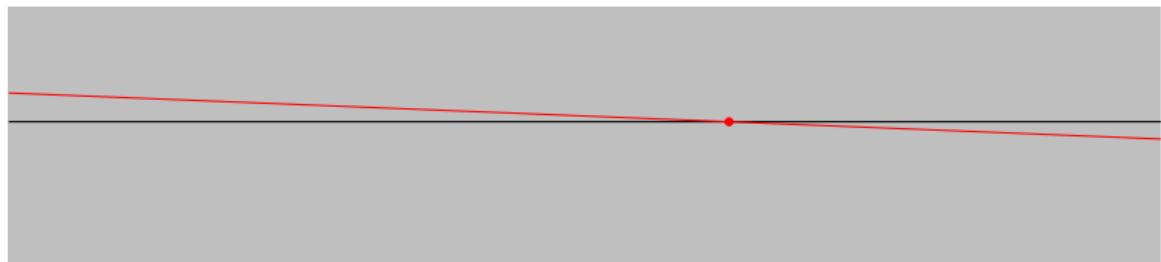
# Ill-Conditioned Linear System



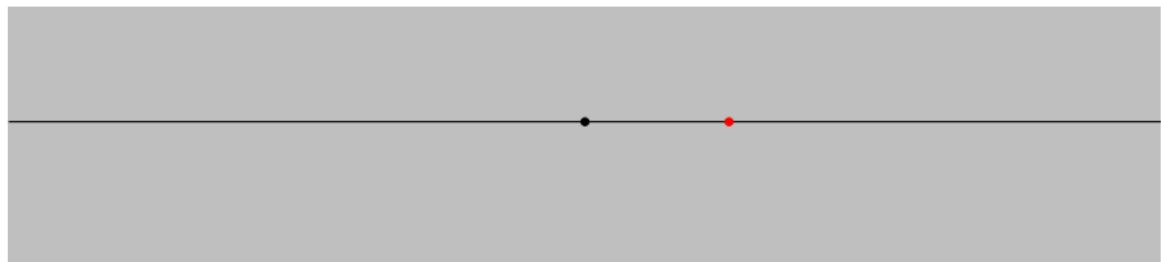
# Ill-Conditioned Linear System



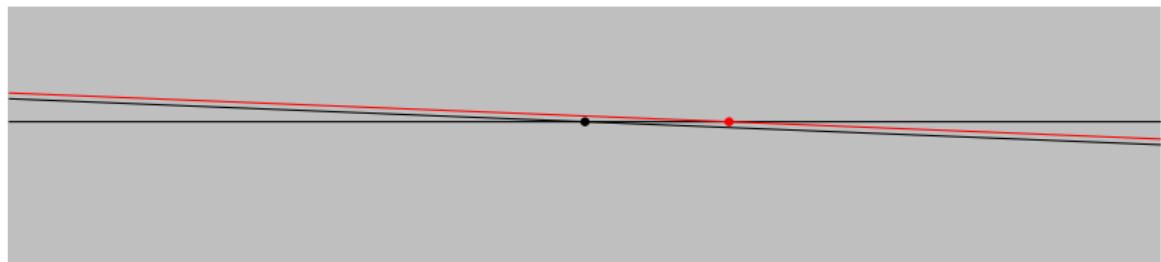
# Ill-Conditioned Linear System



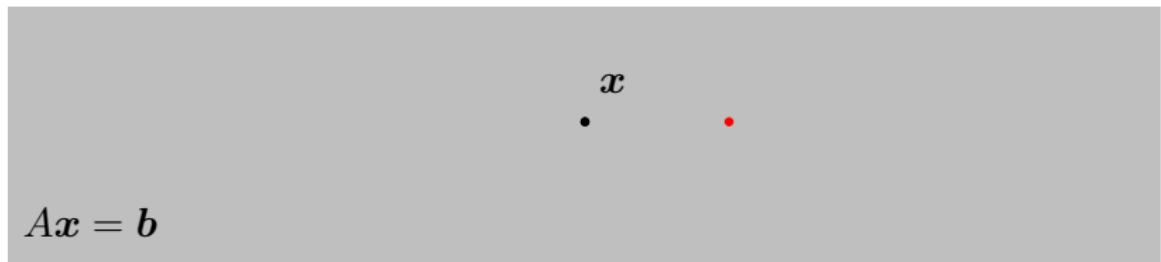
# Ill-Conditioned Linear System



# Ill-Conditioned Linear System



# Ill-Conditioned Linear System



# Ill-Conditioned Linear System



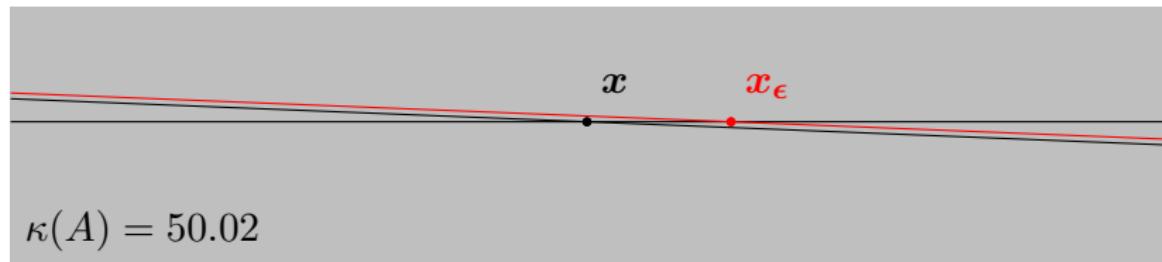
$$Ax_\epsilon = b + \epsilon =: b_\epsilon$$

# Ill-Conditioned Linear System



$$\frac{\|x_\epsilon - x\|}{\|\epsilon\|} = 25.00$$

## Ill-Conditioned Linear System



# Ill-Conditioned Linear System

- ▶ Ill-conditioned linear systems appear frequently in applications;

# Ill-Conditioned Linear System

- ▶ Ill-conditioned linear systems appear frequently in applications;
- ▶ Condition number will be much higher than example;

# Ill-Conditioned Linear System

- ▶ Ill-conditioned linear systems appear frequently in applications;
- ▶ Condition number will be much higher than example;
- ▶ Under the presence of noise, severe loss of precision.

# Regularization

- ▶ Replace the original problem by a stable perturbed version from a family  $\{P_\gamma(\mathbf{b}_\epsilon)\}_{\gamma \in \Gamma}$ ;

# Regularization

- ▶ Replace the original problem by a stable perturbed version from a family  $\{P_\gamma(\mathbf{b}_\epsilon)\}_{\gamma \in \Gamma}$ ;
- ▶ Example: Tikhonov Regularization ( $\Gamma = \mathbb{R}^+$ )

$$\mathbf{x}_\gamma^{\text{TIK}}(\mathbf{b}_\epsilon) := \underset{\mathbf{x} \in \mathbb{R}_n}{\operatorname{argmin}} \|A\mathbf{x} - \mathbf{b}_\epsilon\|_2^2 + \gamma \|\mathbf{x}\|_2^2$$

# Regularization

- ▶ Replace the original problem by a stable perturbed version from a family  $\{P_\gamma(\mathbf{b}_\epsilon)\}_{\gamma \in \Gamma}$ ;
- ▶ Example: Tikhonov Regularization ( $\Gamma = \mathbb{R}^+$ )

$$\begin{aligned}\mathbf{x}_\gamma^{\text{TIK}}(\mathbf{b}_\epsilon) &:= \underset{\mathbf{x} \in \mathbb{R}_n}{\operatorname{argmin}} \|A\mathbf{x} - \mathbf{b}_\epsilon\|_2^2 + \gamma \|\mathbf{x}\|_2^2 \\ &= (A^T A + \gamma I)^{-1} A^T \mathbf{b}_\epsilon;\end{aligned}$$

# Regularization

- ▶ Replace the original problem by a stable perturbed version from a family  $\{P_\gamma(\mathbf{b}_\epsilon)\}_{\gamma \in \Gamma}$ ;
- ▶ Example: Tikhonov Regularization ( $\Gamma = \mathbb{R}^+$ )

$$\begin{aligned}\mathbf{x}_\gamma^{\text{TIK}}(\mathbf{b}_\epsilon) &:= \underset{\mathbf{x} \in \mathbb{R}_n}{\operatorname{argmin}} \|A\mathbf{x} - \mathbf{b}_\epsilon\|_2^2 + \gamma \|\mathbf{x}\|_2^2 \\ &= (A^T A + \gamma I)^{-1} A^T \mathbf{b}_\epsilon;\end{aligned}$$

- ▶ How to choose the regularization parameter  $\gamma$ ?

# Deterministic Regularization Parameter Choice

- ▶ If  $\|\epsilon_k\| \rightarrow 0$ , then the parameter selection function  $\ell(b_{\epsilon_k}, \|\epsilon_k\|)$  must satisfy:

$$x_{\ell(b_{\epsilon_k}, \|\epsilon_k\|)}(b_{\epsilon_k}) \rightarrow x;$$

# Deterministic Regularization Parameter Choice

- ▶ If  $\|\epsilon_k\| \rightarrow 0$ , then the parameter selection function  $\ell(b_{\epsilon_k}, \|\epsilon_k\|)$  must satisfy:

$$x_{\ell(b_{\epsilon_k}, \|\epsilon_k\|)}(b_{\epsilon_k}) \rightarrow x;$$

- ▶ Drawback:

# Deterministic Regularization Parameter Choice

- ▶ If  $\|\epsilon_k\| \rightarrow 0$ , then the parameter selection function  $\ell(b_{\epsilon_k}, \|\epsilon_k\|)$  must satisfy:

$$x_{\ell(b_{\epsilon_k}, \|\epsilon_k\|)}(b_{\epsilon_k}) \rightarrow x;$$

- ▶ Drawback:
  - ▶ Impossible if  $\|\epsilon\|$  is not available (is it ever available?);

# Deterministic Regularization Parameter Choice

- ▶ If  $\|\epsilon_k\| \rightarrow 0$ , then the parameter selection function  $\ell(b_{\epsilon_k}, \|\epsilon_k\|)$  must satisfy:

$$x_{\ell(b_{\epsilon_k}, \|\epsilon_k\|)}(b_{\epsilon_k}) \rightarrow x;$$

- ▶ Drawback:
  - ▶ Impossible if  $\|\epsilon\|$  is not available (is it ever available?);
  - ▶ Says nothing when  $\epsilon$  is not small.

# Frequentist Regularization Parameter Choice

- ▶ A stochastic model for the problem may be available if it is possible to estimate  $E\epsilon\epsilon^T$  and we assume  $E\epsilon = \mathbf{0}$ ;

# Frequentist Regularization Parameter Choice

- ▶ A stochastic model for the problem may be available if it is possible to estimate  $E\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T$  and we assume  $E\boldsymbol{\epsilon} = \mathbf{0}$ ;
- ▶ For concreteness we use the model

$$E\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T = \text{diag}\{A\mathbf{x}\} + \sigma^2 I,$$

for a known  $\sigma$ ;

# Frequentist Regularization Parameter Choice

- ▶ A stochastic model for the problem may be available if it is possible to estimate  $E\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T$  and we assume  $E\boldsymbol{\epsilon} = \mathbf{0}$ ;
- ▶ For concreteness we use the model

$$E\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T = \text{diag}\{Ax\} + \sigma^2 I,$$

for a known  $\sigma$ ;

- ▶ That is, a mixed Gaussian-Poissonian model where

$$E\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T = E \text{diag}\{\mathbf{b}_\epsilon\} + \sigma^2 I.$$

# Noise in Imaging Technologies

- ▶ Imaging technologies often work by counting photon arrival;

# Noise in Imaging Technologies

- ▶ Imaging technologies often work by counting photon arrival;
- ▶ Examples: tomographic scanners, charge-coupled device (CCD) sensors;

# Noise in Imaging Technologies

- ▶ Imaging technologies often work by counting photon arrival;
- ▶ Examples: tomographic scanners, charge-coupled device (CCD) sensors;
- ▶ Photon emission is a Poisson process and the underlying electronic circuitry adds Gaussian noise;

# Noise in Imaging Technologies

- ▶ Imaging technologies often work by counting photon arrival;
- ▶ Examples: tomographic scanners, charge-coupled device (CCD) sensors;
- ▶ Photon emission is a Poisson process and the underlying electronic circuitry adds Gaussian noise;
- ▶ Tomographic image reconstruction and image deblurring are ill-posed linear inverse problems.

# Regularization under Poissonian Noise

[Bardsley and Goldes(2009)] advocate to choose  $\gamma$  such that

$$\|\text{diag}(A\mathbf{x}_\gamma)^{-1/2}(A\mathbf{x}_\gamma - \mathbf{b})\|_2^2 = m,$$

because  $E\|(A\mathbf{x})_i - b_i\|_2^2 = (A\mathbf{x})_i$ .

# Regularization under Poissonian Noise

[Bertero *et al.*(2010)] propose choosing  $\gamma$  such that

$$KL(\mathbf{b}, A\mathbf{x}_\gamma) = \frac{m}{2},$$

because

$$E[KL(\mathbf{b}, A\mathbf{x})] = \frac{m}{2} + \sum_{i=1}^m O((A\mathbf{x})_i^{-1}),$$

where  $KL$  is the Kullback-Leibler entropic divergence:

$$KL(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i \log \frac{x_i}{y_i} - x_i + y_i.$$

# Regularization under Poissonian Noise

[Santos and De Pierro(2009)], instead, propose to minimize

$$E(KL(A\mathbf{x}, A\mathbf{x}_\gamma));$$

# Regularization under Poissonian Noise

[Santos and De Pierro(2009)], instead, propose to minimize

$$E(KL(A\mathbf{x}, A\mathbf{x}_\gamma));$$

To this end, they provide an approximate estimator:

$$\begin{aligned} E[KL(A\mathbf{x}, A\mathbf{x}_\gamma)] &\approx E \left[ \sum_{i=1}^m (A\mathbf{x}_\gamma)_i \right] - E \left[ \mathbf{b}^T \log(A\mathbf{x}_\gamma) \right] \\ &\quad - E \left[ \frac{m\boldsymbol{\omega}^T \text{diag}(\mathbf{b}) \{\log(A\mathbf{x}_\gamma^+) - \log(A\mathbf{x}_\gamma^-)\}}{\delta \|\boldsymbol{\omega}\|_2^2} \right] + K, \end{aligned}$$

where  $\boldsymbol{\omega} \sim \mathcal{N}(\mathbf{0}, I)$ ,  $\delta > 0$ , and  $\mathbf{x}_\gamma^\pm$  is the solution obtained using  $b \pm \delta\boldsymbol{\omega}$  as data.

# Regularization with Mixed Noise

Assume  $d_{\mathbf{b}}(\mathbf{x}, \mathbf{x}_\gamma) = U_{\mathbf{b}}(\mathbf{x}) + V_{\mathbf{b}}(A\mathbf{x})^T W_{\mathbf{b}}(\mathbf{x}_\gamma) + X_{\mathbf{b}}(\mathbf{x}_\gamma)$

$$\begin{aligned} E\left(d_{\mathbf{b}}(\mathcal{A}\mathbf{x}, \mathcal{A}\mathbf{x}_\gamma)\right) &\approx E\left(X_{\mathbf{b}}(\mathbf{x}_\gamma)\right) + E\left(V_{\mathbf{b}}(\mathbf{b})^T W_{\mathbf{b}}(\mathbf{x}_\gamma)\right) + K \\ &\quad - mE\left(\frac{\boldsymbol{\omega}^T \text{diag}(\mathbf{b} + \sigma^2 \mathbb{I}) [F_\gamma(\mathbf{b} + \delta\boldsymbol{\omega}) - F_\gamma(\mathbf{b} - \delta\boldsymbol{\omega})]}{2\delta\|\boldsymbol{\omega}\|^2}\right), \end{aligned}$$

where  $F_\gamma(\mathbf{b}) = V'_{\mathbf{b}}(\mathbf{b})^T W_{\mathbf{b}}(\mathbf{x}_\gamma)$ ,  $\boldsymbol{\omega} \sim \mathcal{N}(0, I)$  and  $K$  is a constant which does not depend on  $\gamma$ .

## Special Case: Minimal Expected Squared Norm (MES)

If we use  $d_{\mathbf{b}}(\mathbf{x}, \mathbf{x}_\gamma) = \|A\mathbf{x} - A\mathbf{x}_\gamma\|$  in the above result, we obtain:

$$\begin{aligned} E [\|A\mathbf{x} - A\mathbf{x}_\gamma\|_2^2] &\approx K + E [\|A\mathbf{x}_\gamma\|_2^2] - 2E[\mathbf{b}^T A\mathbf{x}_\gamma] \\ &\quad + mE \left[ \frac{\boldsymbol{\omega}^T \text{diag}(\mathbf{b} + \sigma^2 \mathbb{I})(A\mathbf{x}_\gamma^+ - A\mathbf{x}_\gamma^-)}{\delta \|\boldsymbol{\omega}\|_2^2} \right]. \end{aligned}$$

## Special Case: MES with Linear Regularization

If the regularization method is linear, the estimate is exact!

$$\begin{aligned} E \left[ \|A\mathbf{x} - A\mathbf{x}_\gamma\|_2^2 \right] &= K + E \left[ \|A\mathbf{x}_\gamma\|_2^2 \right] - 2E[\mathbf{b}^T A\mathbf{x}_\gamma] \\ &\quad + mE \left[ \frac{\boldsymbol{\omega}^T \text{diag}(\mathbf{b} + \sigma^2 \mathbb{I})(A\mathbf{x}_\gamma^+ - A\mathbf{x}_\gamma^-)}{\delta \|\boldsymbol{\omega}\|_2^2} \right]. \end{aligned}$$

# Tomography

- ▶ Techniques for non-destructive visualization of cross-sectional images;

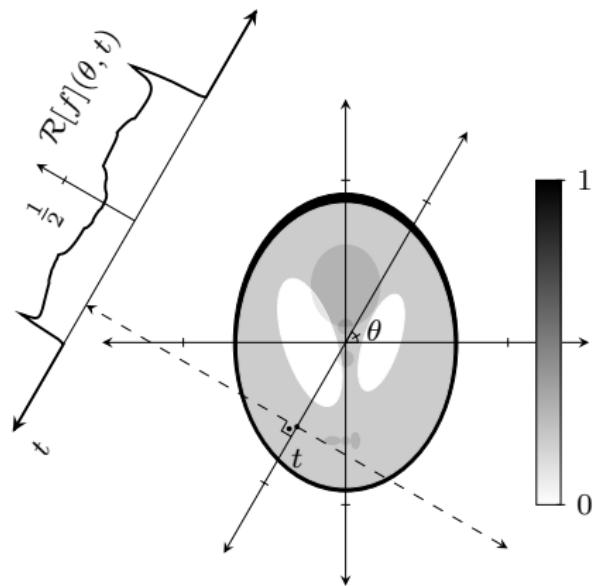
# Tomography

- ▶ Techniques for non-destructive visualization of cross-sectional images;
- ▶ Applications in medicine; materials evaluation; botanics, etc;

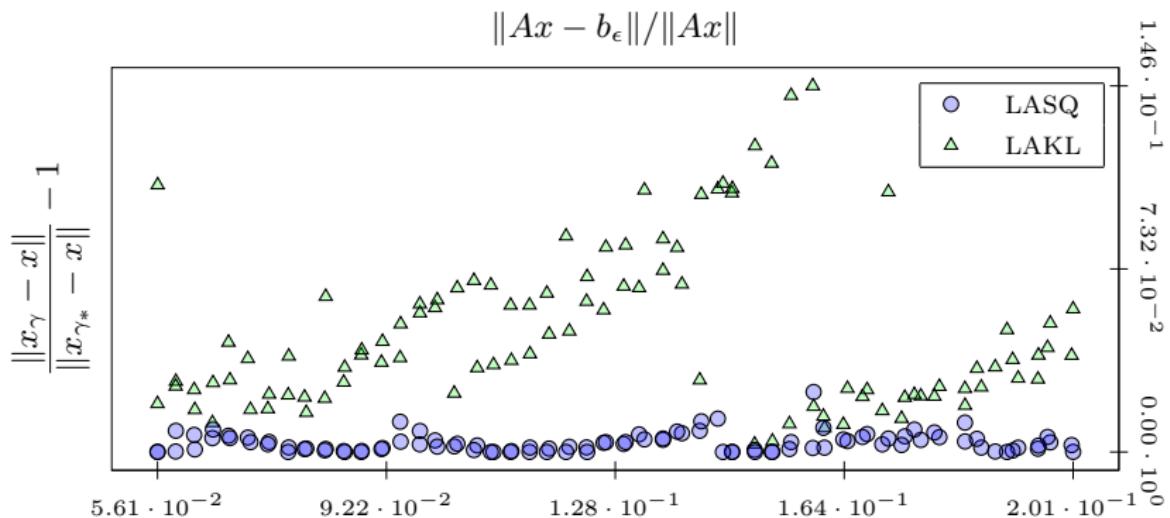
# Tomography

- ▶ Techniques for non-destructive visualization of cross-sectional images;
- ▶ Applications in medicine; materials evaluation; botanics, etc;
- ▶ Measured data are approximate line integrals.

# Tomography



## Applications in Tomography

Figura: Results for FBP regularization. Poissonian noise only ( $\sigma = 0$ ).

## Applications in Tomography

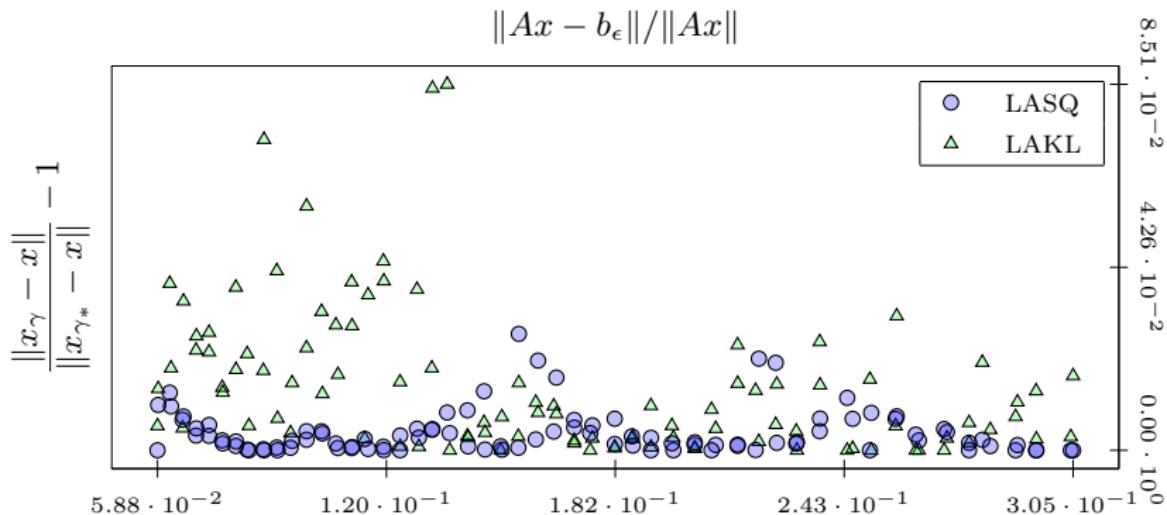


Figura: Results for FBP regularization. Mixed noise ( $\sigma = 5$ ).

# Conclusions and Future Research

- ▶ Very good methodology for parameter selection for linear regularization;

# Conclusions and Future Research

- ▶ Very good methodology for parameter selection for linear regularization;
- ▶ Previous experiments indicates good performance for nonlinear regularization methods as well;

# Conclusions and Future Research

- ▶ Very good methodology for parameter selection for linear regularization;
- ▶ Previous experiments indicates good performance for nonlinear regularization methods as well;
- ▶ Further experimentation required with other divergence measures.

# References

-  JOHNATHAN M. BARDSLEY AND JOHN GOLDES.  
Regularization parameter selection methods for ill-posed Poisson maximum likelihood estimation.  
*Inverse Problems*, 25(9):095005, 2009.  
doi: 10.1088/0266-5611/25/9/095005.
-  MARIO BERTERO, PATRIZIA BOCCACCI, GIORGIO TALENTI, RICCARDO ZANELLA AND LUCA ZANNI.  
A discrepancy principle for Poisson data.  
*Inverse Problems*, 26(10):105004–105023, 2010.  
doi: 10.1088/0266-5611/26/10/105004.
-  REGINALDO J. SANTOS AND ÁLVARO RODOLFO DE PIERRO.  
A new parameters choice method for ill-posed problems with Poisson data and its application to emission tomographic imaging.  
*International Journal of Tomography & Statistics*, 11(s09):33–52, 2009.  
URL <http://www.ceser.res.in/ceserp/index.php/ijts/issue/view/19>.