#### Machine learning algorithms for inverse problems Methods for image reconstruction and image deblurring Lecture 3

イロト イポト イヨト イヨト

# What is image deblurring

Image deblurring is recovering the original sharp image using a mathematical model.



Original Image



< ロ > < 日 > < 日 > < 日 > < 日 > < 日 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > < 1 > <

#### Figure: left: exact matrix X, right: approximated matrix B

- Assume linear blurring operator A which is applied to the exact image X. As a result we will have blurred image B.
- **X**:  $m \times n$  matrix representing the exact image.
- **B**:  $m \times n$  matrix representing the blurred image.
- The forward problem is well-posed: by knowing **X** and linear operator A compute the blurred image:

$$A(\mathbf{X}) = \mathbf{B}.$$
 (1)

However, the inverse problem: by knowing the blurred image B compute the original image X, is already an ill-posed problem.

#### The point-spread function

 Usually, the blurring model is described by a Fredholm integral equation of the first kind which takes the form

$$\int_{\Omega} K(x, y) \mathbf{X}(x) dx = B(y), y \in \Omega.$$
(2)

Here,  $\Omega$  is a closed bounded set representing our image domain in  $\mathbb{R}^n$ , n = 2, 3. It is assumed that the kernel  $K(x, y) \in C^k(\Omega \times \Omega)$  is the absolutely integrable known function. Then the equation (2) can be represented in the operator form (1) with the operator A defined as

$$A(\mathbf{X}) := \int_{\Omega} K(x, y) \mathbf{X}(x) dx.$$
(3)

- Information from one part of the image "spills over" to another part
- The process is modeled by a point-spread function (PSF) which is modeled by a Kernel in a linear operator A.
- PSF usually assumed to be space-invariant

•

#### Examples of PSF

• *Motion blur*: The point source is smeared into a line. Different choices are available in the literature, see for example,

M. Tico, M. Vehvilainen, Estimation of motion blur PSF from differently exposed image frames, EUSIPCO 2006, Florence, Italy, September 4-8, 2006.

イロト イポト イヨト イヨト

#### Examples of PSF

• *Motion blur*: The point source is smeared into a line. Different choices are available in the literature, see for example,

M. Tico, M. Vehvilainen, Estimation of motion blur PSF from differently exposed image frames, EUSIPCO 2006, Florence, Italy, September 4-8, 2006.

Out-of-focus blur PSF is given by

$$p_{ij} = \begin{cases} 1/(\pi r^2) & \text{if } (i-k)^2 + (j-l)^2 \le r^2, \\ 0 & \text{elsewhere,} \end{cases}$$

イロト イポト イヨト イヨト

where (k, l) is the center of PSF, and r is the radius of the blur.

#### Examples of PSF

• *Motion blur*: The point source is smeared into a line. Different choices are available in the literature, see for example,

M. Tico, M. Vehvilainen, Estimation of motion blur PSF from differently exposed image frames, EUSIPCO 2006, Florence, Italy, September 4-8, 2006.

Out-of-focus blur PSF is given by

$$p_{ij} = \begin{cases} 1/(\pi r^2) & \text{if } (i-k)^2 + (j-l)^2 \le r^2, \\ 0 & \text{elsewhere,} \end{cases}$$

where (k, l) is the center of PSF, and r is the radius of the blur.

 Atmospheric turbulence described as a two-dimensional Gaussian PSF function at every point p<sub>ii</sub> of PSF:

$$p_{ij} = \exp\left(-\frac{1}{2}\begin{bmatrix}i-k\\j-l\end{bmatrix}^T\begin{bmatrix}s_1^2&\rho^2\\\rho^2&s_2^2\end{bmatrix}^{-1}\begin{bmatrix}i-k\\j-l\end{bmatrix}\right),$$

where the parameters  $s_1$ ,  $s_2$  and  $\rho$  determine the width and orientation of the PSF, centered at (k, l).

#### Motion blur

Motion Blurred Image





< ロト < 回 > < 回 > < 回 > < 回 >

æ

#### Out-of-focus blur

Blurred Image



ヘロト ヘロト ヘヨト ヘヨト

æ

Larisa Beilina, http://www.math.chalmers.se/~larisa/

Lecture 3

#### Gaussian blur

Gaussian Blurred Image





◆□ ▶ ◆□ ▶ ◆ □ ▶ ◆ □ ▶ ●

æ



#### Image deblurring: solution of an inverse problem

Mathematically, the problem of image deblurring consist in the solution of a Fredholm integral equation of the first kind which is an ill-posed problem. Let *H* be the Hilbert space  $H^1$  and let  $\Omega \subset \mathbb{R}^m$ , m = 2, 3, be a convex bounded domain. Our goal is to solve a Fredholm integral equation of the first kind for  $x \in \Omega$ 

$$\int_{\Omega} K(x-y)z(x)dx = u(y), \qquad (4)$$

where  $u(y) \in L_2(\overline{\Omega})$ ,  $z(x) \in H$ ,  $K(x - y) \in C^k(\overline{\Omega})$ ,  $k \ge 0$  be the kernel of the integral equation.

Let us rewrite (4) in an operator form as

$$A(z) = u \tag{5}$$

・ロト ・ 日 ・ ・ ヨ ・ ・

with an convolution operator  $A : H \to L_2(\overline{\Omega})$  defined as

$$A(z) := \int_{\Omega} K(x - y) z(x) dx.$$
 (6)

**Image restoration problem** Let the function  $z(x) \in H^1$  of the equation

$$\int_{\Omega} K(x-y)z(x)dx = u(y),$$

be unknown in the domain  $\Omega$ . Determine the function z(x) for  $x \in \Omega$  assuming the functions  $K(x - y) \in C^k(\overline{\Omega})$ ,  $k \ge 0$  and  $u(x) \in L_2(\Omega)$  are known.

The image restoration problem is described by a Fredholm integral equation of the first kind with a linear compact operator  $A : H \to L_2(\bar{\Omega})$ . which is an ill-posed problem.

Thus, regularization should be used.

# Ill-posed problem: the case of a general compact operator

Let  $H_1$  and  $H_2$  be two Hilbert spaces with dim  $H_1 = \dim H_2 = \infty$ . Remind that a sphere in an infinitely dimensional Hilbert

space is not a compact set. The Theorem below (see Theorem 1.2 from [BK]) says that the solution of a Fredholm integral

equation of the first kind is an ill-posed problem.

**Theorem [BK]** Let  $G = \{ ||x||_{H_1} \le 1 \} \subset H_1$ , is not a compact set. Let  $A : G \to H_2$  be a compact operator and let R(A) := A(G) be its range. Consider an arbitrary point  $y_0 \in R(A)$ . Let  $\varepsilon > 0$  be a number and  $U_{\varepsilon}(y_0) = \{ y \in H_2 : ||y - y_0||_{H_2} < \varepsilon \}$ . Then there exists a point  $y \in U_{\varepsilon}(y_0) \setminus R(A)$ . If, in addition, the operator A is one-to-one, then the inverse operator  $A^{-1} : R(A) \to G$  is not continuous. Hence, the problem of the solution of the equation

$$A(x) = z, x \in G, z \in R(A)$$
(7)

イロン イヨン イヨン イヨン

is unstable, i.e. this is an ill-posed problem.

L. Beilina, M. Klibanov, Approximate global convergence and adaptivity for coefficient inverse problems, Springer, 2012.

#### When the problem is an ill-posed problem



**Theorem [BK]** Let  $G = \{ \|x\|_{H_1} \le 1 \} \subset H_1$ , is not a compact set. Let  $A : G \to H_2$  be a compact operator and let R(A) := A(G) be its range. Consider an arbitrary point  $y_0 \in R(A)$ . Let  $\varepsilon > 0$  be a number and  $U_{\varepsilon}(y_0) = \{ y \in H_2 : \|y - y_0\|_{H_2} < \varepsilon \}$ . Then there exists a point  $y \in U_{\varepsilon}(y_0) \setminus R(A)$ . If, in addition, the operator A is one-to-one, then the inverse operator  $A^{-1} : R(A) \to G$  is not continuous. Hence, the problem of the solution of the equation  $A(x) = z, x \in G, z \in R(A)$  is unstable, i.e. this is an ill-posed problem.

**Proof.** First, we prove the existence of a point  $y \in U_{\varepsilon}(y_0) \setminus R(A)$ . Assume to the contrary, i.e. assume that  $U_{\varepsilon}(y_0) \subset R(A)$ . Let  $\{y_n\}_{n=1}^{\infty} \subset H_2$  be an orthonormal basis in  $H_2$ . Then the sequence

$$\left\{y_0+\frac{\varepsilon}{2}y_n\right\}_{n=1}^{\infty}:=\left\{z_n\right\}_{n=1}^{\infty}\subset\left\{\left\|y-y_0\right\|=\frac{\varepsilon}{2}\right\}\subset U_{\varepsilon}\left(y_0\right).$$

- I. We have that ||z<sub>n</sub> − z<sub>m</sub>||<sub>H₂</sub> = ε/√2. Hence, the sequence {z<sub>n</sub>}<sub>n=1</sub><sup>∞</sup> does not contain a fundamental subsequence.
- 2. Therefore,  $U_{\varepsilon}(y_0)$  is not a precompact set in  $H_2$ .
- 3. On the other hand, since G is a closed bounded set and A is a compact operator, then R (A) is a compact set.
- 4. Hence,  $U_{\varepsilon}(y_0)$  is a precompact set. However, by item 2)  $U_{\varepsilon}(y_0)$  is not a precompact set in  $H_2$  contradiction.

We now prove the second assertion. Assume to the contrary: that the operator  $A^{-1} : R(A) \to G$  is continuous. By the definition of the operator A we have  $A^{-1}(R(A)) = G$ . Thus, since R(A) is a compact set in  $H_2$ , then the continuity of  $A^{-1}$  implies that G is a compact set in  $H_1$ , which is impossible.  $\Box$ 

#### Conclusions about solution of an ill-posed problem

- By the theorem above (see Theorem 1.2 from [BK]) the set R(A) is not dense everywhere. Therefore, the question about the existence of the solution of equation A(x) = z does not make an applied sense. Indeed, since the set R(A) is not dense everywhere, it is very hard to describe a set of values belonging to this set.
- An example, consider the case when the kernel K (x, y) ∈ C ([a, b] × [0, 1]) in equation

$$\int_{0}^{1} K(x, y) f(y) dy = g(x), x \in (0, 1),$$
(8)

is an analytic function of the real variable  $x \in (a, b)$ . Then the right hand side g(x) of equation (8) should also be analytic with respect to  $x \in (a, b)$ . However, in applications the function g(x) is a result of measurements, it is given only at a number of discrete points and it definitely contains noise. Clearly it is impossible to determine whether the resulting function is analytic or not. **Conclusion.** Assuming that conditions of Theorem above (Theorem 1.2 from [BK]) are satisfied, the problem of solving equation A(x) = z is an ill-posed problem since:

イロト イヨト イヨト

- The proof of an existence theorem makes no applied sense.
- Small fluctuations of the right hand side can lead to large fluctuations of the solution *x*, i.e. the problem is unstable.

Let  $\delta > 0$  be the error in the right-hand side of the equation (4):

$$A(z^{*}) = u^{*}, \quad ||u - u^{*}||_{L_{2}(\sigma)} \le \delta.$$
(9)

where  $u^*$  is the exact right-hand side corresponding to the exact solution  $z^*$ .

To find the approximate solution of the equation (4) corresponding to applications in image restoration usually is minimized the functional

$$M_{\alpha}(z) = \|Az - u\|_{L_{2}(\Omega)}^{2} + \alpha \|z\|_{H_{1}(\Omega)}^{2}, \qquad (10)$$
$$M_{\alpha} : H^{1} \to \mathbb{R},$$

イロン イヨン イヨン イヨン

where  $\alpha = \alpha(\delta) > 0$  is the small regularization parameter.

#### **Convolution equation**

Let us consider an important class of Fredholm integral equations of the first kind - the convolution equation. These equations can be presented in the form (5) with the convolution operator  $A : H^1 \to L_2(\Omega)$  defined by

$$A(z) := \int_{\Omega} \rho(y - x) z(x) dx, \qquad (11)$$

where  $\rho(y - x) \in C^k(\Omega \times \Omega)$ ,  $k \ge 0$ ,  $z(x) \in H^1$ . Then using the convolution theorem and properties of the Fourier transform we obtain the minimum  $z(x) \in H^1$  of the functional (10) given by

$$z(x) = F^{-1} \left( \frac{\hat{u}(\omega)\hat{\rho}^*(\omega)}{|\hat{\rho}(\omega)|^2 + \alpha(1+\omega^2)} \right), \tag{12}$$

where  $\hat{f}(\omega)$  denotes the Fourier transform  $F(f)(\omega)$  of the function  $f(\omega)$  defined by

$$\hat{f}(\omega) := F(f)(\omega) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(x) e^{-i\omega x} dx.$$
(13)

・ロト ・日ト ・ヨト ・ヨト

We consider now more general form of the Tikhonov functional (10). Let  $W_1, W_2, Q$  be three Hilbert spaces,  $Q \subseteq W_1$  as a set, the norm in Q is stronger than the norm in  $W_1$  and  $\overline{Q} = W_1$ , where the closure is understood in the norm of  $W_1$ . We denote scalar products and norms in these spaces as

 $(\cdot, \cdot)$ ,  $\|\cdot\|$  for  $W_1$ ,  $(\cdot, \cdot)_2$ ,  $\|\cdot\|_2$  for  $W_2$ and  $[\cdot, \cdot]$ ,  $[\cdot]$  for Q.

Let  $A : W_1 \to W_2$  be a bounded linear operator. Our goal is to find the function  $z(x) \in Q$  which minimizes the Tikhonov functional

$$E_{\alpha}(z): Q \to \mathbb{R}, \tag{14}$$

$$E_{\alpha}(z) = \frac{1}{2} \|Az - u\|_{2}^{2} + \frac{\alpha}{2} [z - z_{0}]^{2}, u \in W_{2}; z, z_{0} \in Q,$$
(15)

where  $\alpha \in (0, 1)$  is the regularization parameter. To do that we search for a stationary point of the above functional with respect to *z* satisfying  $\forall b \in Q$ 

$$E'_{\alpha}(z)(b) = 0. \tag{16}$$

イロト イヨト イヨト イヨト

The following lemma is well known [BKK] for the case  $W_1 = W_2 = L_2$ . **Lemma 1.** Let  $A : L_2 \rightarrow L_2$  be a bounded linear operator. Then the Fréchet derivative of the functional (10) is

$$E'_{\alpha}(z)(b) = (A^*Az - A^*u, b) + \alpha [z - z_0, b], \forall b \in Q.$$
(17)

In particular, for the integral operator (4) we have

$$E'_{\alpha}(z)(b) = \int_{\Omega} b(s) \left[ \int_{\Omega} z(y) \left( \int_{\Omega} K(x-y)K(x-s)dx \right) dy - \int_{\Omega} K(x-s)u(x) dx \right] ds + \alpha [z-z_0, b], \forall b \in Q.$$
(18)

イロン イヨン イヨン イヨン

[BKK] A. B. Bakushinsky, M. Y. Kokurin, A. Smirnova, Iterative methods for ill-posed problems, Walter de Gruyter GmbH&Co., 2011. Lemma 2 is also well known, since  $A : W_1 \to W_2$  is a bounded linear operator. We formulate this lemma only for our specific case. **Lemma 2.** Let the operator  $A : W_1 \to W_2$  satisfies conditions of Lemma 1. Then the functional  $E_{\alpha}(z)$  is strongly convex on the space Q with the convexity parameter  $\kappa$  such that

$$(E'_{\alpha}(x) - E'_{\alpha}(z), x - z) \ge \kappa [x - z]^2, \forall x, z \in Q.$$
(19)

It is known from the theory of convex optimization that Lemma 2 implies existence and uniqueness of the global minimizer  $z_{\alpha} \in Q$  of the functional  $J_{\alpha}$  defined in (14) such that

$$J_{\alpha}(z_{\alpha}) = \inf_{z \in Q} J_{\alpha}(z).$$

It is well known that the operator F = Az - u is Lipschitz continuous

$$\|F(z_1) - F(z_2)\| \le \|A\| \cdot \|z_1 - z_2\| \quad \forall z_1, z_2 \in H.$$

We also introduce new constant  $D = D(||A||, \alpha) = const. > 0$  such that

$$\left\|J_{\alpha}'(z_{1}) - J_{\alpha}'(z_{2})\right\| \le D \left\|z_{1} - z_{2}\right\|, \forall z_{1}, z_{2} \in H.$$
(20)

Similarly, the functional  $M_{\alpha}(z)$  is also strongly convex on the Sobolev space  $H_1$ :

$$(M'_{\alpha}(x) - M'_{\alpha}(z), x - z)_{H_{1}} \ge \kappa ||x - z||_{H_{1}}^{2}, \forall x, z \in H_{1},$$
(21)

・ロト ・ 日本 ・ 日本 ・ 日本

To find minimum of (15) or (10) (difference in reg.terms), we can use any gradient-like method. For example, perform usual gradient update to find minimum of (15):

$$z^{k+1} = z^k + \beta E'_{\alpha}(z^k)(b).$$

or to find minimum of (10):

$$z^{k+1} = z^k + \beta M'_{\alpha}(z^k)(b).$$

until  $||z^{k+1} - z^k||$  converges.

## Bayesian approach

- Assume that measured data *y* is random because of noise, also assume that the reconstruction *x* is random. Then it can be used posterior distribution denoted by P(**x** = *x*|**y** = *y*) := P(*x*|*y*). Posterior distribution P(*x*|*y*) is the probability (density) if the random variable **x** is equal to *x* given that a random variable **y** is equal to *y*.
- Bayes theorem

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}$$

Here, P(y|x) is data likelihood - the probability of a measured data y given a reconstruction x, it is defined by the physical model problem. Example: for the Radon transform with Poisson noise the data likelihood is approximated using normal distribution

$$P(y|x) \propto e^{-\frac{1}{2}\|Ax-y\|_{\Lambda^{-1/2}}^2}$$

with a covariance approximated as  $\Lambda = (y_0 e^{-Ax})$ ,  $y_0$  is the number of photons to heat every pixel in homo media.

・ロト ・ 日 ト ・ 日 ト ・ 日 ト ・

#### Bayesian approach

- P(x) is the prior model, or probability of the how the reconstruction should look like. It should include all information which we know about the reconstruction.
- As soon as P(y|x) and P(x) are defined, the estimator of data likelihood P(x|y) should be chosen. The most popular one is the Maximum a posterior estimate MAP(y) which is the solution of the problem

$$MAP(y) = \max_{x \in X} P(x|y).$$

Taking the log of the above equation, we get

$$\max_{x \in X} \log P(x|y) = \max_{x \in X} [\log P(y|x) + \log P(x) - \log P(y)]$$
  
= 
$$\min_{x \in X} - [\log P(y|x) + \log P(x)]$$
 (22)

#### Bayesian approach: an example

Let us consider an example of a Gaussian prior P(x) with mean  $\mu$  and covariance  $\Sigma$ . Then the prior likelihood is:

$$P(x) \propto e^{-\frac{1}{2}||x-\mu||_{\Sigma^{-1/2}}^2},$$
  
$$\ln P(x) = -\frac{1}{2}||x-\mu||_{\Sigma^{-1/2}}^2 + Const.$$
 (23)

If we assume the Gaussian noise with mean zero, then the data has mean A(x) and covariance  $\Lambda$ , and data likelihood P(y|x) can be modeled as

$$P(y|x) \propto e^{-\frac{1}{2}||Ax-y||^{2}_{A^{-1/2}}},$$
  

$$\ln P(y|x) = -\frac{1}{2}||Ax-y||^{2}_{A^{-1/2}} + Const.$$
(24)

Combining models for P(x) (23) and for P(y|x) (24) into the maximum a posteriori estimate MAP(y) (22), we get

$$\max_{x \in X} \ln P(x|y) = \min_{x \in X} -[\ln P(y|x) + \ln P(x)]$$

$$= \min_{x \in X} [\frac{1}{2} ||Ax - y||_{\Lambda^{-1/2}}^{2} + \frac{1}{2} ||x - \mu||_{\Sigma^{-1/2}}^{2}].$$
(25)

#### Bayesian approach: an example

We can approximate the operator A with a matrix  $\mathcal{A}$ , take the gradient of (25) we get:

$$(\log P(x|y))'_{x} = (\frac{1}{2} ||\mathcal{A}x - y||^{2}_{\Lambda^{-1/2}} + \frac{1}{2} ||x - \mu||^{2}_{\Sigma^{-1/2}})'_{x}$$
  
=  $\mathcal{A}^{T} \Lambda^{-1} (\mathcal{A}x - y) + \Sigma^{-1} (x - \mu)$   
=  $\mathcal{A}^{T} \Lambda^{-1} \mathcal{A}x - \mathcal{A}^{T} \Lambda^{-1} y + \Sigma^{-1} x - \Sigma^{-1} \mu$   
=  $(\mathcal{A}^{T} \Lambda^{-1} \mathcal{A} + \Sigma^{-1}) x - (\mathcal{A}^{T} \Lambda^{-1} y + \Sigma^{-1} \mu).$  (26)

Thus, we obtain system of linear equations considering  $(\log P(x|y))'_{x} = 0$ :

$$(\mathcal{A}^{T}\Lambda^{-1}\mathcal{A} + \Sigma^{-1})x = \mathcal{A}^{T}\Lambda^{-1}y + \Sigma^{-1}\mu$$
(27)

The system above can be solved as

$$x = ((\mathcal{A}^T \Lambda^{-1} \mathcal{A} + \Sigma^{-1}))^{-1} (\mathcal{A}^T \Lambda^{-1} y + \Sigma^{-1} \mu).$$
(28)

・ロト ・ 日 ・ ・ ヨ ・ ・ ヨ ・ ・

Another type of prior which can be taken into consideration, is sparsity inducing prior. This prior is considered when the obtained solution is very smooth.

The sparsity inducing prior the prior likelihood is:

$$-\log P(x) = \lambda ||Dx||_1 + Const,$$
(29)

イロト イヨト イヨト イヨト

where D is an operator such that Dx is sparse.

Further reading about Bayesian approach in inverse problems:

J. Kaipio, E. Somersalo, Statistical and computational inverse problems, Springer, 160, 2006.

A. M. Stuart, Inverse Problems: a Bayesian perspective, Acta Numerica, 19, 451-559, 2010.

#### The finite element method

- We discretize the bounded domain  $\Omega \subset \mathbb{R}^n$ , n = 2, 3 by an mesh T using non-overlapping elements K.  $\mathbb{R}^3$  are tetrahedrons, in  $\mathbb{R}^2$  the elements K are triangles such that  $T = K_1, ..., K_l$ , where l is the total number of elements in  $\Omega$ .
- Let the mesh function h = h(x) is a piecewise-constant function such that

$$h(x) = h_K \quad \forall K \in T,$$

where  $h_K$  is the diameter of K which we define as the longest side of K.

We introduce now the finite element space V<sub>h</sub> as

$$V_{h} = \{v(x) \in V : v \in C(\Omega), v|_{K} \in P_{1}(K) \forall K \in T\},$$
(30)

where  $P_1(K)$  denotes the set of piecewise-linear functions on K with

$$V = \{v(x) : v(x) \in H^1(\Omega)\}.$$

The finite dimensional finite element space  $V_h$  is constructed such that  $V_h \subset V$ .

#### The finite element method

Let us consider the Tikhonov functional

$$M_{\alpha}(z) = \frac{1}{2} \|Az - u\|_{L_{2}(\Omega_{k})}^{2} + \frac{\alpha}{2} \||\nabla z\||_{L^{2}(\Omega)}^{2}.$$
 (31)

For (31) the CG(1) finite element method reads: find z<sub>h</sub> ∈ V<sub>h</sub> such that

$$M'_{\alpha}(z_h)(b) = 0 \quad \forall b \in V_h. \tag{32}$$

 Then the Fréchet derivative of the functional (31) is derived in Lemma 2 of [BGN]

**Lemma 2.** Let  $A : H^1(\Omega) \to L_2(\Omega_k)$  be a bounded linear operator. Then the Fréchet derivative of the functional (31) is

$$M'_{\alpha}(z)(b) = (A^*Az - A^*u, b) + \alpha(|\nabla z|, |\nabla b|), \ \forall b \in H^1(\Omega), \quad (33)$$

with a convex growth factor *b*, i.e.,  $|\nabla b| < b$ .

[BGN] Beilina L., Guillot G., Niinimäki K. The Finite Element Method and Balancing Principle for Magnetic Resonance Imaging. In Mathematical and Numerical Approaches for Multi-Wave Inverse Problems. CIRM 2019. Springer Proceedings in Mathematics Statistics, vol 328. Springer, Cham (2020) https://doi.org/10.1007/978-3-030-48634-1\_9

#### A posteriori ananlysis in a finite element method

- Let  $P_h : V \to M$  for  $\forall M \subset V$ , be the operator of the orthogonal projection of V on M. Let the function  $f \in H^1(\Omega) \cap C(\Omega)$  and  $\partial_{x_i} f_{x_i} \in L_{\infty}(\Omega)$ . We define by  $f_k^I$  the standard interpolant [EEJ] on triangles/tetrahedra of the function  $f \in H$ .
- Then by one of properties of the orthogonal projection

$$\|f - P_h f\|_{L_2(\Omega)} \le \|f - f_k^I\|_{L_2(\Omega)}.$$
(34)

It follows from formula 76.3 of [EEJ] that

$$\|f - P_h f\|_{L_2(\Omega)} \le C_I \|h \nabla f\|_{L_2(\Omega)}, \forall f \in V.$$
(35)

イロト イポト イヨト イヨト

where  $C_l = C_l(\Omega)$  is positive constant depending only on the domain  $\Omega$ .

[EEJ] K. Eriksson, D. Estep and C. Johnson, Calculus in Several Dimensions, Springer, Berlin, 2004.

### A posteriori analysis in a finite element method

Let us present a general framework for a posteriori error estimate for:

• For the error  $|M_{\alpha}(z_{\alpha}) - M_{\alpha}(z_{h})|$  in the Tikhonov functional (31).

• For the error  $|z_{\alpha} - z_{h}|$  in the regularized solution  $z_{\alpha}$  of this functional. Note that

$$M_{\alpha}(z_{\alpha}) - M_{\alpha}(z_h) = M_{\alpha}'(z_h)(z_{\alpha} - z_h) + R(z_{\alpha}, z_h), \qquad (36)$$

where  $R(z_{\alpha}, z_{h})$  is the second order remainder term. We assume that  $z_{h}$  is located in the small neighborhood of the regularized solution  $z_{\alpha}$ . Thus, the term  $R(z_{\alpha}, z_{h})$  is small and we can neglect it. We now use the Galerkin orthogonality principle

$$M'_{\alpha}(z_h)(b) = 0 \quad \forall b \in V_h \tag{37}$$

together with the splitting

$$z_{\alpha} - z_{h} = (z_{\alpha} - z_{\alpha}^{\prime}) + (z_{\alpha}^{\prime} - z_{h})$$
(38)

where  $z_{\alpha}^{l} \in V_{h}$  is the interpolant of  $z_{\alpha}$ , and get the following error representation:

$$M_{\alpha}(z_{\alpha}) - M_{\alpha}(z_{h}) \approx M_{\alpha}'(z_{h})(z_{\alpha} - z_{\alpha}'). \tag{39}$$

#### Theorem 1 [KB]

Let  $z_h \in V_h$  be a finite element approximation of the regularized solution  $z_{\alpha} \in H^2(\Omega)$  on the finite element mesh T with the mesh function h. Then there exists a constant D defined by (20) such that the following a posteriori error estimate for the regularized solution  $z_{\alpha}$  holds

$$||z_h-z_\alpha||_{H^1(\Omega)}\leq \frac{D}{\alpha}C_l||hz_h||_{L_2(\Omega)}.$$

[KB] N. Koshev and L. Beilina, An adaptive finite element method for Fredholm integral equations of the first kind and its verification on experimental data, in the Topical Issue "Numerical Methods for Large Scale Scientific Computing" of CEJM, 11(8), 1489-1509, 2013.

#### Theorem 2 [KB]

Suppose that there exists minimizer  $z_{\alpha} \in H^{2}(\Omega)$  of the functional  $M_{\alpha}$  on the set V and mesh T. Suppose also that there exists finite element approximation  $z_{h}$  of a minimizer  $z_{\alpha}$  of  $M_{\alpha}$  on the set  $V_{h}$  and mesh T with the mesh function h. Then the following approximate a posteriori error estimate for the error  $e = |M_{\alpha}(z_{\alpha}) - M_{\alpha}(z_{h})|$  in the Tikhonov functional holds

$$e = |M_{\alpha}(z_{\alpha}) - M_{\alpha}(z_{h})| \le C_{I} \left\| M_{\alpha}'(z_{h}) \right\|_{H^{1}(\Omega)} \|hz_{h}\|_{L_{2}(\Omega)}.$$
(40)

イロト イヨト イヨト

[KB] N. Koshev and L. Beilina, An adaptive finite element method for Fredholm integral equations of the first kind and its verification on experimental data, in the Topical Issue "Numerical Methods for Large Scale Scientific Computing" of CEJM, 11(8), 1489-1509, 2013.

# AFEM algorithm

- Step 0. Choose an initial mesh  $T_0$  in  $\Omega$  and obtain the numerical solution  $z_0$  on  $T_0$  using the convolution theorem. Compute the sequence  $z_k$ , k > 0, on a refined meshes  $T_k$  via following steps:
- Step 1. Obtain the FEM solution  $z_k := z_{hk}$  using the finite element method

$$M'_{lpha}(z_h)(b) = 0 \ \forall b \in V_h.$$

Step 2. Refine the mesh  $T_k$  using the Theorem 1 at all points where

$$\left|z_{k}\left(x\right)\right| \geq \widetilde{\varkappa_{k}} \max_{\Omega} \left|z_{k}\left(x\right)\right|,\tag{41}$$

or using the Theorem 2 at all points where

$$|M_h(z_k)| \ge \beta_k \max_{\Omega} |M_h(z_k)|.$$
(42)

Here, the tolerances  $\tilde{\varkappa_k}, \beta_k \in (0, 1)$  are chosen by the user.

Step 3. Construct a new mesh  $T_{k+1}$  in  $\Omega$  and perform steps 1-3 on the new mesh. Stop mesh refinements when  $||z_k - z_{k-1}|| < \epsilon$  or  $||M_h(z_k)|| < \epsilon$ , or image deteriortaed.

#### Image deblurring: example



a)



ヘロト ヘヨト ヘヨト

# Figure: a) Image of the defect in the planar chip. b) result of reconstruction with bounded total variation functions. Source: [1].

[1] Koshev N.A., Orlikovsky N.A., Rau E.I., Yagola A.G. Solution of the inverse problem of restoring the signals from an electronic microscope in the backscattered electron mode on the class of bounded variation functions, Numerical Methods and Programming, 2011, V.11, pp. 362-367 (in Russian).

# Image deblurring: an adaptive refinement, example



Figure: Reconstruction from the experimental backscattering data obtained by the microtomograph [KB].

・ロト ・ 日 ・ ・ ヨ ト ・ ヨ ト

## Image deblurring: an adaptive refinement, example



Figure: Reconstruction from the experimental backscattering data obtained by the microtomograph [KB].

[KB] N. Koshev and L. Beilina, An adaptive finite element method for Fredholm integral equations of the first kind and its verification on experimental data, in the Topical Issue "Numerical Methods for Large Scale Scientific Computing" of CEJM, 11(8), 1489-1509, 2013.

Larisa Beilina, http://www.math.chalmers.se/~larisa/