Mathematisches Forschungsinstitut Oberwolfach

Report No. 51/2012

DOI: 10.4171/OWR/2012/51

Computational Inverse Problems

Organised by
Habib Ammari, Paris
Liliana Borcea, Houston
Thorsten Hohage, Göttingen
Barbara Kaltenbacher, Klagenfurt

21st October – 27th October 2012

Abstract. Inverse problem typically deal with the identification of unknown quantities from indirect measurements and appear in many areas in technology, medicine, biology, finance, and econometrics. The computational solution of such problems is a very active, interdisciplinary field with close connections to optimization, control theory, differential equations, asymptotic analysis, statistics, and probability. The focus of this workshop was on hybrid methods, model reduction, regularization in Banach spaces, and statistical approaches.

Mathematics Subject Classification (2000): 35R25, 35R30, 45Q05, 65J20, 65J22, 65M32, 65N21.

Introduction by the Organisers

The workshop was well attended by 49 participants from four continents, among them 12 females. 14 participants came from outside of Europe, and 8 were PhD students or young postdocs (less than one year after PhD). The scientific program consisted of 27 full and 6 short talks. On Wednesday night, after an excursion to St Roman, an informal discussion on data assimilation and inverse problems in weather prediction was organized by Roland Potthast.
The talks reflected a number of exciting new developments in the field of computational inverse problems. We highlight a few general trends:

- *hybrid methods:* A number of speakers reported on these new techniques in biomedical imaging which combine different physical phenomena such as light and sound. Hybrid methods promise to combine the advantages of more traditional imaging techniques and lead a number of new challenging

mathematical problems. Many classical inverse problems in partial differential equations such as Calderon's conductivity problem assume only boundary measurements of the solution of the differential equation. This leads to severe ill-posedness and low resolution even though one has high contrast in the unknown coefficient. Hybrid methods in principle allow for distributed measurements which would lead to a much better resolution, but additional equations have to be taken into account. A number of clever approaches for the algorithmic treatment, modeling and the analysis of such problems have been presented.

- *model reduction:* Regularization methods typically require numerous solutions of the forward problem, and often one forward problem involves the solution of a three-dimensional differential equation for many right-hand sides. For many interesting problems recent advances in the application of reduced order models allow a drastic reduction of computational complexity without essential loss of accuracy. Other talks reported on progress towards making large scale Bayesian inverse problems computationally feasible using low rank approximations and allowing a quantification of uncertainty.
- *regularization in Banach spaces:* Whereas for many decades regularization methods were studied almost exclusively in a Hilbert space setting using spectral methods, a number of talks reviewed the recent significant progress in formulating and analyzing regularization methods in a Banach space setting using variational methods. Of particular interest are l^1 penalty terms, which promote sparsity of the solution with respect to a given basis, and total (generalized) variation penalties. However, in particular for these most interesting spaces many questions still remain open and are subject of current research.
- *statistical approaches:* Although it is well-known that reconstructions can be significantly improved if knowledge on the distribution of the noise is incorporated in the inversion method, such approaches could be analyzed only recently in the context of nonlinear inverse problems as they appear in differential equations. A different type of error is caused by random perturbations of coefficients or boundaries in wave equations. Here a remarkable body of theory has been developed analyzing the effects of such random perturbations. Finally, despite some seemingly discouraging negative results progress was reported on the construction of confidence bands and credibility sets in inverse problems.

Many talks initiated lively discussions during and after the talks. The long lunch breaks and evening were used to continue existing and start new collaborations among the participants.

The organizers would like to thank the Oberwolfach institute for their great hospitality and for the opportunity to arrange this stimulating workshop.

Workshop: Computational Inverse Problems

Table of Contents

Abstracts

Photoacoustic Sectional Imaging and Reconstruction Formulas for a Single Scattering Model

Otmar Scherzer
(joint work with Peter Elbau, Rainer Schulze)

We consider photoacoustic sectional imaging experiments. Opposed to standard photoacoustic imaging (see e.g. [9, 8, 6, 10] for some mathematical and physical review papers), where the detectors record sets of two-dimensional projection data over time, from which the three-dimensional imaging data can be reconstructed, in sectional imaging, a single scan procedure is implemented to be able to reconstruct a set of two-dimensional slice imaging data. The advantages of the latter approach are a considerable increase in measurement efficiency and the possibility to perform selective plane imaging. However, the disadvantage is a decreased out-of-plane resolution (i.e. the direction orthogonal to the focusing plane). Experimentally, one can realize photoacoustic sectional imaging with pulsed laser illuminations focusing to a single plane and with focusing detectors for the measurement of the pressure wave [7].

Analogously to standard quantitative photoacoustic imaging we observe two decoupled reconstruction problems:

(1) The inverse acoustic problem of recovering the initial pressure in the illuminated slice from the two-dimensional measurements of the pressure wave and
(2) the inverse optical problem of reconstructing the absorption coefficient from this initial pressure data, which is assumed to be proportional to the absorption coefficient and to the light fluence of the laser pulse.

For the acoustic problem, we model the propagation of the pressure wave by the linear three-dimensional wave equation and derive explicit reconstruction formulas for the initial pressure distribution for various detector geometries (including point, line, and plane shaped detectors placed around the object). In particular for line detectors in the illumination plane and vertical plane detectors, exact reconstruction formulas for an arbitrary placement of these detectors around the object are available [4]. Moreover, these photoacoustic sectional measurements allow for a simultaneous reconstruction of an unknown speed of sound if measurements for all possible slices through the object are performed [5].

For the optical problem, we assume (in accordance with the focused illumination of only one slice) that scattering effects are sufficiently weak so that a single scattering model for the light propagation of the illuminating laser beam is practible. This is different to recent approaches in standard quantitative photoacoustic imaging [1, 2] where the scattering is typically assumed to be so large that a diffusion approximation model for the light propagation can be used. In this sectional imaging approach, however, the scattering in the object should be rather small, since a

localisation of the illumination is otherwise not possible. Nevertheless, our reconstruction formulas for the single scattering approach rely on a similar strategy as in the diffusion approximation and are based on deriving equations for quotients of independent measurement data obtained from measurements for two different laser illuminations (e.g. from two opposing directions) [3].

Acknowledgements: The work has been supported by the Austrian Science Fund (FWF) within the national research network Photoacoustic Imaging in Biology and Medicine, project S10505-N20.

References

[1] G. Bal, *Hybrid inverse problems and internal functionals*, arXiv:1110.4733v1, 2011, http://arxiv.org/abs/1110.4733v1.

[2] G. Bal, *Explicit reconstructions in QPAT, QTAT, TE, and MRE*, arXiv:1202.3117v1, 2012, http://arxiv.org/abs/1202.3117v1.

[3] P. Elbau and O. Scherzer, *Reconstruction Formulas for a Single Scattering Model in Photoacoustic Imaging and Applications to Sectional Imaging*, arXiv:1206.6989, 2012, http://arxiv.org/abs/1206.6989.

[4] P. Elbau, O. Scherzer, and R. Schulze, *Reconstruction formulas for photoacoustic sectional imaging*, Inverse Probl. 28 (4):045004, 2012, ISSN 0266-5611, http://dx.doi.org/10.1088/0266-5611/28/4/045004.

[5] A. Kirsch and O. Scherzer, *Simultaneous reconstructions of absorption density and wave speed with photoacoustic measurements*, SIAM J. Appl. Math. 72 (5): 1508-1523, 2012, http://dx.doi.org/10.1137/110849055.

[6] P. Kuchment and L. Kunyansky, *Mathematics of thermoacoustic tomography*, European J. Appl. Math. 19:191–224, 2008.

[7] R. Nuster, S. Gratt, K. Passler, G. Paltauf, and D. Meyer, *Photoacoustic section imaging using an elliptical acoustic mirror and optical detection*, J. Biomed. Opt. 17:030503, 2012, ISSN 1083-3668, http://dx.doi.org/10.1117/1.JBO.17.3.030503.

[8] C. Li and L. V. Wang, *Photoacoustic tomography and sensing in biomedicine*, Phys. Med. Biol. 54:R59–R97, 2009.

[9] M. Xu and L. V. Wang, *Photoacoustic imaging in biomedicine*, Rev. Sci. Instruments 77 (4):1–22, 2006.

[10] L. V. Wang, *Prospects of photoacoustic tomography*, Med. Phys. 35 (12):5758–5767, 2008.

An acousto-optic imaging model for the reconstruction of the optical absorption parameter

Laurent Seppecher

The idea is to increase the resolution of the Near Infa-Red light tomography using acoustic perturbations of the medium. We use a classical approach to describe the light diffusion which is diffusion absorption equation

$$-\nabla D \nabla \Phi + a\Phi = 0. \tag{1}$$

We perturb the PDE solution by some spherical short pulses which displace the inside of the medium and we measure the variations of Φ on the boundary due to the traveling of the acoustic pulse. Our unknown is the absorption term in the

equation (1) and our measurements fead though a spherical means transform to the knowledge of Ψ in the Helmholtz decomposition of

$$\Phi^2 \nabla a = \nabla \Psi + \nabla \times A.$$

Then we write the solution a as a solution of

$$\begin{aligned} \nabla \left(\Phi^2 \nabla a\right) &= \triangle \Psi && \text{in } \Omega \\ a &= a_0 && \text{on } \partial\Omega. \end{aligned}$$

And we finally solve the coupled system that we have obtained using the fixed point algorithm.

Reconstruction of dynamic objects in computerized tomography

Bernadette Hahn

In computerized tomography, the rotation of the x-ray source around the specimen is the time consuming part of the scanning process since beams from only one source position can be emitted at the same time. In 2D parallel scanning geometry, this position is characterized by the direction θ of the emitted x-rays. Thus, each beam direction θ can be uniquely identified by a time $t \in \mathbb{R}_T \subset \mathbb{R}$ via the correlation

$$\theta = \theta(t) = \begin{pmatrix} \cos(t\,\phi) \\ \sin(t\,\phi) \end{pmatrix},$$

where ϕ denotes the constant angular velocity of the radiation source. The symbol t_{in} denotes the starting time of the scanning.

One basic assumption in this context is that the object does not change during the data acquisition. In many applications, however, this supposition does not hold, for example in medical imaging due to respiratory and cardiac motion or in the imaging of driven liquid fronts in field oil extraction. The temporal changes of the object lead to inconsistent data and the application of standard reconstruction methods results in serious motion artefacts in the images. Hence, algorithms should take into account the dynamics of the investigated object.

Let $\Omega \subset \mathbb{R}^2$ and $f \in L_2(\Omega \times \mathbb{R}_T)$ be a dynamic function, such that the static function $f_{\theta(t)}(x) := f(x,t)$ denotes the x-ray attenuation coefficient of the object at time t. The mathematical model of dynamic computerized tomography is given by

$$\mathcal{A}f = g,$$

with

$$\mathcal{A}f(\theta(t), s) = \int_\Omega f(x,t)\,\delta(s - x^T\theta(t))\,\mathrm{d}x, \qquad \theta(t) \in S^1,\ s \in \mathbb{R}.$$

For a fixed unit vector θ, it holds

$$\mathcal{A}f(\theta, s) = \mathcal{R}f_\theta(\theta, s)$$

with the Radon transform $\mathcal{R}$. Hence, some additional information of the motion is required to obtain an adequate reconstruction of the functions f_θ, [1]. With this in

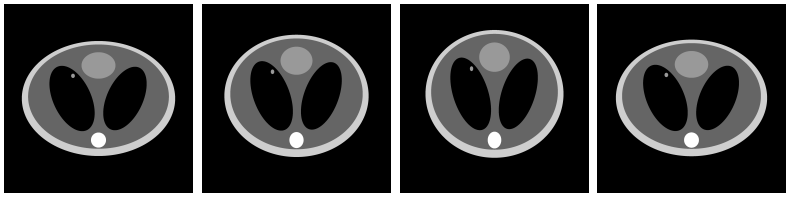


FIGURE 1. Movements of the phantom during the scanning.

mind, we consider a dynamic behaviour that is described by bijective, sufficiently smooth motion functions

$$\Gamma_t : \mathbb{R}^2 \longrightarrow \mathbb{R}^2.$$

With this motion model, the dynamic object is represented by a static reference function $f^{\mathrm{ref}} \in L_2(\Omega)$,

$$f(x,t) = f^{\mathrm{ref}}(\Gamma_t x).$$

Without loss of generality, $\Gamma_{t_{in}}$ equals the identity and it holds

$$f^{\mathrm{ref}}(x) = f(x, t_{in}).$$

Hence, f^{ref} can be determined by reconstructing $f_{t_{in}}$ from equation $\mathcal{A}f = g$. Therefore, the method of the approximate inverse [2] is applied as regularization scheme due to the ill-posedness of the problem: Instead of $f_{t_{in}}$, we approximate

$$f^{\gamma}_{t_{in}} = f^{\gamma}(x, t_{in}) := \langle f, \delta^{\gamma}_{x,t_{in}} \rangle$$

with a mollifier $\delta^{\gamma}_{x,t_{in}} \in L_2(\Omega \times \mathbb{R}_T)$. With a solution $\psi^{\gamma}_{x,t_{in}}$ of the auxiliary problem

$$\mathcal{A}^* \psi^{\gamma}_{x,t_{in}} = \delta^{\gamma}_{x,t_{in}},$$

the function $f^{\gamma}_{t_{in}}$ can be computed from the data via

$$f^{\gamma}_{t_{in}}(x) = \langle g, \psi^{\gamma}_{x,t_{in}} \rangle.$$

Moreover, the information about the dynamic behaviour needs to be included in order to obtain adequate reconstructions. This is done by choosing an appropriate mollifier. If e^{γ}_x denotes a mollifier for the static function f^{ref}, then

$$\delta^{\gamma}_{x,t_{in}}(y,v) = \Big(\int_{\mathbb{R}_T} \big| \det D\Gamma^{-1}_{\tau}(\Gamma_v y) \big| \, \mathrm{d}\tau \Big)^{-1} e^{\gamma}_x(\Gamma_v y)$$

is a suitable mollifier for $f_{t_{in}}$ considering the motion model.

The corresponding reconstruction kernel $\psi^{\gamma}_{x,t_{in}}$ can be computed by minimizing the defect $\|\mathcal{A}\,\psi^{\gamma}_{x,t_{in}} - \delta^{\gamma}_{x,t_{in}}\|^2$. In the static case, this minimizer makes an inadequate approximation to the static reconstruction kernel that can be computed exactly. Hence, the kernels $\psi^{\gamma}_{x,t_{in}}$ are rather calculated as a weighted sum of the defect's minimizer and the kernel in the static case, that corresponds to the mollifier e^{γ}_x.

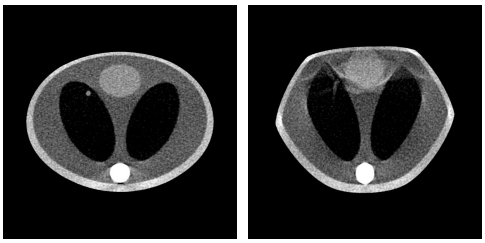


FIGURE 2. Dynamic (left) and static (right) reconstruction.

The sequence of pictures in Figure 1 illustrates the respiratory movement of a chest phantom during one breathing cycle. The involved motion functions represent affine deformations. The analytically computed data are corrupted by 2% noise. Including the dynamical information actually avoids the motion artefacts that occur within the static filtered backprojection approach, Figure 2.

REFERENCES

[1] F. Natterer, *The Mathematics of Computerized Tomography*, John Wiley & Sons, Chichester, 1986.
[2] A. K. Louis, *Approximate inverse for linear and some nonlinear problems*, Inverse Problems **12** (1996), 175–190.

Mathematical modeling of fluorescence diffuse optical imaging of cell membrane potential changes

LAURE GIOVANGIGLI

1. GOVERNING MODEL

We consider a cell, that we want to image. We inject fluorescent indicators, which stick only on the cell membrane. These agents are chosen so that their concentration responds linearly to the potential jump across the membrane, when the cell is immersed in an external electric field. We apply such en external electric field at the boundary of our domain and use fluorescence optical diffuse tomography to reconstruct the position and shape of the membrane, and image changes in the membrane potential.

1.1. **Coupled diffusion equations.** A sinusoidally modulated near infrared monochromatic light source g, located at the boundary $\partial\Omega$ of the examined domain Ω, launches an excitation light fluence $\phi_{\rm exc} = \Phi_{\rm exc}(x,\omega)\, e^{i\omega t}$, into Ω. After it undergoes multiple scattering and absorption, this light wave reaches the fluorescent markers, which are accumulated on ∂C, the membrane of the cell C. The excited fluorophores emit a wave $\phi_{\rm emt} = \Phi_{\rm emt}(x,\omega)\, e^{i\omega t}$. The intensity of the emitted wave

is proportional to the intensity of the excitation wave, when it reaches the fluorescent molecule. The emitted waves pass through the absorbing and scattering domain and are detected at the boundary $\partial\Omega$.
After some simplifying assumptions, our model can be described by the following coupled diffusion equations completed by Robin boundary conditions :

$$\begin{cases} -D\Delta\Phi^g_{\text{exc}}(x,\omega) + \left(\mu + \dfrac{i\omega}{c}\right)\Phi^g_{\text{exc}}(x,\omega) = 0 & \text{in}\,\Omega, \\ \ell\dfrac{\partial\Phi^g_{\text{exc}}}{\partial\nu}(x,\omega) + \Phi^g_{\text{exc}}(x,\omega) = g(x) & \text{on}\,\partial\Omega, \end{cases}$$

$$\begin{cases} -D\Delta\Phi^g_{\text{emt}}(x,\omega) + \left(\mu + \dfrac{i\omega}{c}\right)\Phi^g_{\text{emt}}(x,\omega) = \tilde{\gamma}(\omega)\, c_{\text{flr}}(x)\,\Phi^g_{\text{exc}}(x,\omega) & \text{in}\,\Omega \\ \ell\dfrac{\partial\Phi^g_{\text{emt}}}{\partial\nu}(x,\omega) + \Phi^g_{\text{emt}}(x,\omega) = 0 & \text{on}\,\partial\Omega, \end{cases}$$

where the source g is in $L^2(\partial\Omega)$.

1.2. **Electrical model of the cell.** We couple this model with an electric model of the cell, which gives us a modelisation of the fluorophores concentration.

We apply at the boundary of our domain an electric field $g_{\text{ele}} \in L^2(\partial\Omega)$. We consider that $\Omega\setminus\overline{C}$ and C are homogeneous and isotropic media with conductivity 1. The thickness ϵ of the cell membrane is supposed to be small. We denote by σ the conductivity of the cell membrane. We assume that $\sigma \ll 1$ and $\beta > 0$ to be given by $\beta = \sigma^{-1}\epsilon$.
We can approximate the voltage potential u within our medium by the solution to the following problem :

$$\text{(1)} \qquad \begin{cases} \Delta u = 0 & \text{in}\, C \cup \Omega\setminus\overline{C}, \\ \left.\dfrac{\partial u}{\partial\nu}\right|_+ - \left.\dfrac{\partial u}{\partial\nu}\right|_- = 0 & \text{on}\,\partial C, \\ u\,|_+ - u\,|_- = \beta\dfrac{\partial u}{\partial\nu} & \text{on}\,\partial C, \\ \left.\dfrac{\partial u}{\partial\nu}\right|_{\partial\Omega} = g_{\text{ele}}, \qquad \displaystyle\int_{\partial\Omega} u = 0. \end{cases}$$

Since we have chosen the fluorescent indicators of the cell membrane such that they respond linearly to the potential jump across the membrane, we can express their concentration as

$$\text{(2)} \qquad c_{\text{flr}} = \delta\,[u]\big|_{\partial C},$$

where δ is a constant.

2. Forward problem

Given a cell C, optical parameters of our domain, a light source g and an electric field g_{ele}, we solve the forward problem. We give in particular an explicit expression of the excitation wave in the case of circular domain and cell.

3. Reconstruction of the cell membrane and its potential changes : case of a perturbed disk

We study then the inverse problem in the case of a circular domain Ω of radius 1 and a cell C_ϵ, whose shape is a perturbed disk :

$$\partial C_\epsilon = \{\tilde{x}; \tilde{x}(\theta) = (R + \epsilon h(\theta))e_r, \theta \in [0, 2\pi]\},$$

with $h \in \mathcal{C}^2([0, 2\pi])$.
Given the intensity of the exciting light, the boundary potential and optical parameters, we reconstruct the Fourier coefficients of the cell deformation h. We provide explicit formulas for the resolving power of the algorithm in the presence of measurement noise.
All results and references cited in my presentation can be found in the article [1].

References

[1] H. Ammari, J. Garnier, L. Giovangigli, *Mathematical modeling of fluorescence diffuse optical imaging of cell membrane potential changes*, Quarterly of Applied Mathematics, to appear.

Efficient Reconstruction Algorithms for Inverse Problems in Quantitative Photoacoustic Imaging

Kui Ren
(joint work with Hao Gao, Alexander V. Mamonov, Hongkai Zhao)

Photoacoustic tomography (PAT) is a hybrid imaging modality that combines the high-resolution ultrasound imaging with the high-contrast optical tomography to take the advantages of both modalities. Reconstruction in PAT is a two-step process. In the first step, one reconstructs the initial pressure field generated from the photoacoustic effect using measured acoustic signal on medium surface. This is a well-known inverse problem that has been thoroughly studied; see for instance [2, 9, 10, 11, 13, 14, 15, 17, 18, 19] and references therein.

This work is concerned with the second step, call quantitative PAT (QPAT). The objective is to reconstruct the optical absorption and the scattering coefficients as well as the photoacoustic efficiency of the medium from the result of the first step, i.e., the initial pressure field data. This step has recently attracted significant attention as well [1, 3, 4, 5, 6, 7, 8].

We consider the problem in two different regimes: the weakly scattering transport regime and the strongly scattering diffusive regime. In the first regime, we

model the light propagation with the radiative transport equation:

(1)
$$\begin{array}{rcll} v\cdot\nabla u+\big(\sigma_a(x)+\sigma_s(x)\big)u &=& \sigma_s\int_{\mathbb{S}^{d-1}}\mathcal{K}(v,v')u(x,v')dv' & \text{in } \Omega\times\mathbb{S}^{d-1}\\ u(x,v) &=& g(x,v) & \text{on } \Gamma_-, \end{array}$$

where $u(x,v)$ is the density of photons at $x\in\Omega\in\mathbb{R}^d$ ($d\geq 2$) traveling in direction $v\in\mathbb{S}^{d-1}$, $\Gamma_-=\{(x,v):(x,v)\in\partial\Omega\times\mathbb{S}^{d-1}\text{ s.t. }-\mathbf{n}(x)\cdot v>0\}$, and $g(x,v)$ is the incoming illumination source. The functions $\sigma_a(x)$ and $\sigma_s(x)$ are the absorption and the scattering coefficients respectively. The scattering kernel $\mathcal{K}(v,v')$ satisfies the normalization condition $\int_{\mathbb{S}^{d-1}}\mathcal{K}(v,v')dv'=1,\ \forall\ v\in\mathbb{S}^{d-1}$. The initial pressure data is given as

$$H(x)=\gamma(x)\sigma_a(x)\int_{\mathbb{S}^{d-1}}u(x,v)dv. \tag{2}$$

We showed in [12] that when the medium to be probed is non-scattering ($\sigma_s=0$), explicit reconstruction schemes can be derived to reconstruct γ and σ_a uniquely and stably from two well-selected data sets. When data at multiple wavelengths are utilized, we can reconstruct simultaneously γ, σ_a and σ_s. We presented some numerical simulations to validate the reconstruction methods developed.

In the strongly scattering diffusive regime, we replace the radiative tranport model with the diffusion model:

(3)
$$\begin{array}{rcll} -\nabla\cdot D\nabla U+\sigma_a U &=& 0, & \text{in } \Omega\\ U+\eta\mathbf{n}\cdot D\nabla U &=& f(x), & \text{on } \partial\Omega \end{array}$$

where $U(x)=\int_{\mathbb{S}^{d-1}}u(x,v)dv$ is the density of photons at x, $D(x)$ is the diffusion coefficient that is determined by σ_a and σ_s, and $\eta>0$ is the rescaled extrapolation length.

In this diffusive regime, we proposed in [16] a hybrid numerical reconstruction procedure that uses both the initail pressure data and boundary current data. We showed that these data allow the unique reconstruction of the boundary and interior values D and σ_a. We developed an efficient reconstruction algorithm for the numerical reconstruction.

References

[1] H. Ammari, E. Bossy, V. Jugnon, and H. Kang, *Reconstruction of the optical absorption coefficient of a small absorber from the absorbed energy density*, SIAM J. Appl. Math., 71 (2011), pp. 676–693.

[2] H. Ammari, E. Bretin, J. Garnier, and A. Wahab, *Time reversal in attenuating acoustic media*, Contemporary Mathematics, 548 (2011), pp. 151–163.

[3] G. Bal, A. Jollivet, and V. Jugnon, *Inverse transport theory of photoacoustics*, Inverse Problems, 26 (2010). 025011.

[4] G. Bal and K. Ren, *Multi-source quantitative PAT in diffusive regime*, Inverse Problems, 27 (2011). 075003.

[5] G. Bal and K. Ren, *On multi-spectral quantitative photoacoustic tomography in diffusive regime*, Inverse Problems, 28 (2012). 025010.

[6] G. Bal and G. Uhlmann, *Inverse diffusion theory of photoacoustics*, Inverse Problems, 26 (2010). 085010.

[7] B. T. Cox, S. R. Arridge, and P. C. Beard, *Estimating chromophore distributions from multiwavelength photoacoustic images*, J. Opt. Soc. Am. A, 26 (2009), pp. 443–455.
[8] H. Gao, S. Osher, and H. Zhao, *Quantitative photoacoustic tomography*, in Mathematical Modeling in Biomedical Imaging II: Optical, Ultrasound, and Opto-Acoustic Tomographies, H. Ammari, ed., Lecture Notes in Mathematics, Springer, 2012.
[9] M. Haltmeier, T. Schuster, and O. Scherzer, *Filtered backprojection for thermoacoustic computed tomography in spherical geometry*, Math. Methods Appl. Sci., 28 (2005), pp. 1919–1937.
[10] Y. Hristova, *Time reversal in thermoacoustic tomography - an error estimate*, Inverse Problems, 25 (2009). 055008.
[11] P. Kuchment and L. Kunyansky, *Mathematics of thermoacoustic tomography*, Euro. J. Appl. Math., 19 (2008), pp. 191–224.
[12] A. V. Mamonov and K. Ren, *Quantitative photoacoustic imaging in radiative transport regime*, Comm. Math. Sci., (2012).
[13] L. V. Nguyen, *A family of inversion formulas in thermoacoustic tomography*, Inverse Probl. Imaging, 3 (2009), pp. 649–675.
[14] S. K. Patch and O. Scherzer, *Photo- and thermo- acoustic imaging*, Inverse Problems, 23 (2007), pp. S1–S10.
[15] J. Qian, P. Stefanov, G. Uhlmann, and H. Zhao, *An efficient Neumann-series based algorithm for thermoacoustic and photoacoustic tomography with variable sound speed*, SIAM J. Imaging Sci., 4 (2011), pp. 850–883.
[16] K. Ren, H. Gao, and H. Zhao, *A hybrid reconstruction method for quantitative photoacoustic imaging*, SIAM J. Imag. Sci., (2012).
[17] P. Stefanov and G. Uhlmann, *Thermoacoustic tomography with variable sound speed*, Inverse Problems, 25 (2009). 075011.
[18] P. Stefanov and G. Uhlmann, *Thermoacoustic tomography arising in brain imaging*, Inverse Problems, 27 (2011). 045004.
[19] M. Xu and L. Wang, *Universal back-projection algorithm for photoacoustic computed tomography*, Phys. Rev. E, 71 (2005). 016706.

Inverse Problem of Acousto-Optic Imaging

John Schotland

The acousto-optic effect is a phenomenon in which the optical properties of a material medium are modified due to interaction with acoustic radiation. Brillouin scattering from density fluctuations in a fluid and the ultrasonic modulation of multiply scattered light in a random medium are familiar examples of this effect. We propose a tomographic method to reconstruct the optical properties of a highly scattering medium from acousto-optic measurements. The method is based on the solution to an inverse problem for the diffusion equation and makes use of the principle of interior control of boundary measurements by an external wave field. We prove local nonlinear injectivity and stability. We also propose a reconstruction method and give an estimate for the approximation error.

The inverse Born series in optical tomography and related inverse problems.

Shari Moskow
(joint work with Simon Arridge, Kimberly Kilgore, and John Schotland)

For optical waves in highly scattering media, such as clouds or breast tissue, the diffusion equation is an approximate model for the radiative transport equation.

$$-\Delta u(x) + k^2(1+\eta(x))u(x) = 0 \quad x \in \Omega \tag{1}$$

$$u(x) + l\nu(x)\cdot\nabla u(x) = 0 \quad x \in \partial\Omega \tag{2}$$

The forward problem is to determine the energy density $u(x)$ for a given change in absorption η of compact support. One can express the problem in integral equation form for $x \in \Omega$,

$$u(x) = u_{\mathbf{i}}(x) - k^2 \int_\Omega G(x,y)u(y)\eta(y)dy \tag{3}$$

where G is the Green's function for the given domain Ω with homogeneous Robin boundary conditions and $u_{\mathbf{i}}$ is the solution to the known background problem. Assume the medium is illuminated by a point source at some $x_1 \in \partial\Omega$, which the resulting intensity response is read at another point $x \in \partial\Omega$. The inverse problem is to determine $\eta(x)$ from the data $\phi(x_1, x) = u(x) - u_{\mathbf{i}}(x)$. By inserting $u \approx u_{\mathbf{i}}$ into the right hand side of (3) above, and iterating in a fixed point fashion one obtains the well known Born series for the data ϕ. This series can be expressed as a tensor power series in η :

$$\phi = K_1\eta + K_2\eta\otimes\eta + K_2\eta\otimes\eta\otimes\eta + \cdots \tag{4}$$

which is the data written as a series in the unknown. So, let us assume formally that we could write the unknown coefficient in a series in the data:

$$\eta = \mathcal{K}_1\phi + \mathcal{K}_2\phi\otimes\phi + \mathcal{K}_2\phi\otimes\phi\otimes\phi + \cdots . \tag{5}$$

If we substitute the forward series for ϕ into the formal series for η and equate like tensor powers, we find formulas for the inverse series operators

$$\begin{aligned}
\mathcal{K}_1 &= K_1^+, \\
\mathcal{K}_2 &= -\mathcal{K}_1K_2\mathcal{K}_1\otimes\mathcal{K}_1, \\
\mathcal{K}_3 &= -(\mathcal{K}_2K_1\otimes K_2 + \mathcal{K}_2K_2\otimes K_1 + \mathcal{K}_1K_3)\mathcal{K}_1\otimes\mathcal{K}_1\otimes\mathcal{K}_1, \\
\mathcal{K}_j &= -\left(\sum_{m=1}^{j-1}\mathcal{K}_m \sum_{i_1+\ldots+i_m=j} K_{i_1}\otimes\ldots\otimes K_{i_m}\right)\mathcal{K}_1\otimes\ldots\otimes\mathcal{K}_1. \qquad (6)
\end{aligned}$$

Here, since K_1 does not have a bounded inverse, its inversion is ill-posed. We assume K_1^+ is some regularized pseudo-inverse, and this is the only inversion required to compute the terms of the the inverse series. Using the above formulas, this yields an inversion technique distinctly different than Newton type iterations, since no forward solutions are computed (aside from those of the background). The use of this series for optical tomography was first proposed in [4]. In [5],[6]

and [2] we show that if in some appropriately chosen norms, the forward operators satisfy

$$\|K_j\| \leq \nu\mu^{j-1},$$

the inverse series converges (in the related norm) provided $\|K_1^+\| < \frac{1}{\mu+\nu}$ and

$$\|K_1^+\phi\| < \frac{1}{\mu+\nu}.$$

If a is the radius of a region known to contain the support of η, we find that this 'radius' of the inverse series stays bounded above away from zero as $ka \to \infty$. For the related problem of electrical impedance tomography (or equivalently diffusion-only optical tomography), one can write a similar forward series with operators of quite different properties. In [1] we find that the forward series converges if the contrast $\frac{\delta D}{D_0} < 1$ is less than one, and an analogous result for convergence of the inverse series. We test the approach numerically on a heart and lungs phantom. The higher order terms in the inverse series yield improvements qualitatively like those of Newton iterations, but without the necessity of computing forward solutions. We also studied the use of this series on propagating scalar waves [3], where we see that the radius of convergence in this case does indeed go to zero as $ka \to \infty$, in contrast to the situation for diffuse optical tomography. We see some numerical examples that demonstrate that the inverse Born series works, just not quite as well as in the optical tomography and EIT cases. For Maxwell equations, we have some preliminary theoretical results that suggest the convergence behavior of the inverse series is similar to that for propagating scalar waves, however, we suspect that reconstructions might be better for measurements in the near field due to the presence of evanescent modes.

References

[1] S. Arridge, S. Moskow, J.C. Schotland, *Inverse Born series for the Calderon problem*, Inverse Problems **28**, 035003 (2012).

[2] Kilgore, K., Moskow, S. and Schotland, J. C., "Inverse Born series for diffuse waves." *Imaging microstructures*, 113-122, Contemp. Math., 494, *Amer. Math. Soc., Providence, RI*, (2009).

[3] Kilgore, K., Moskow, S. and Schotland, J. C., *Inverse Born series for scalar waves*, Journal of Computational Mathematics **30** , 601-614, (2012).

[4] Markel, V., O'Sullivan, J. and Schotland, J. C., *Inverse problem in optical diffusion tomography. IV. Nonlinear inversion formulas.*, J. Opt. Soc. Am. A **30**, pp. 903-912 (2003).

[5] Moskow, S. and Schotland, J.C., *Numerical studies of the inverse born series for diffuse waves*, Inverse Problems **25** 095007 (2009).

[6] Moskow, S. and Schotland, J.C., *Convergence and stability of the inverse scattering series for diffuse waves*, Inverse Problems **24** 065005-1065005-16 (2008).

Geometric Reconstruction in Bioluminescence Imaging

Andreas Rieder
(joint work with Tim Kreutzmann)

Bioluminescence imaging is a promising technique to study cancer/tumor growth in a small animal. To this end the target cells are manipulated to emit photons under stimulation, see e.g. [1]. From the photon flux over the animal's surface one has to recover location and intensity of the photon source q.

Photon propagation is modeled by the radiative transfer (Boltzmann transport) equation. In this talk, however, we consider a much simpler model, the diffusion approximation, in which the photon density $u\colon \Omega \to \mathbb{R}$ is governed by the boundary value problem

$$\begin{aligned} -\nabla\cdot(D\nabla u) + \mu u &= q \quad \text{in } \Omega,\\ u + 2D\partial_{\mathbf{n}} u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

where $\Omega \subset \mathbb{R}^d$ is the small animal ($d = 3$) or a cross section of it ($d = 2$). Further, $D\colon \Omega \to]D_0, \infty[$ and $\mu\colon \Omega \to]\mu_0, \infty[$ are the *diffusion* and *absorption* coefficients, respectively. Further, the Robin boundary value is set to zero implementing our assumption that no photons penetrate the animal.

The (linear) forward operator

$$A\colon L^2(\Omega) \to L^2(\partial\Omega), \quad q \mapsto D\partial_{\mathbf{n}} u,$$

maps the photon source to the measurements. Bioluminescence imaging entails the inverse problem: Given $g \in \mathcal{R}(A)$ find a source $q \in L^2(\Omega)$ satisfying $Aq = g$. Due to the compactness of A we have to deal with an ill-posed problem which is, moreover, not uniquely solvable. Indeed, A has a rather huge null space, see [3]. Even, if we model the source as a hot spot we cannot restore uniqueness. A hot spot is a source of type $q = \lambda\chi_S$ where $\lambda > 0$ is the intensity and the measurable domain S supports the source. Roughly speaking we cannot differ between a hot spot with a large intensity and a small support and a hot spot with a small intensity and a large support knowing only the photon flux.

As briefly explained above the bioluminescence sources are clusters of manipulated and marked cells. Therefore, depending on the cell type and used markers a sound guess of the expected intensity of the source is at hand: $\lambda \in [\underline{\lambda}, \overline{\lambda}] = \Lambda$ for a known Λ. Under this a priori knowledge the restriction to hot spots as searched-for sources is reasonable. Consequently, we define the (nonlinear) forward operator

$$F\colon \Lambda \times \mathcal{L} \to L^2(\partial\Omega), \quad F(\lambda, S) = \lambda A\chi_S,$$

where $\mathcal{L}$ is the set of all measurable subsets of Ω.

The corresponding inverse problem (solve $F(\lambda, S) = g$ for given g) remains ill-posed and needs regularization. As regularizer we accept any minimizer of the Tikhonov-like functional

$$J_\alpha(\lambda, S) = \frac{1}{2}\|F(\lambda, S) - g\|_{L^2}^2 + \alpha \mathrm{Per}(S) \text{ over } \Lambda \times \mathcal{L}$$

where $\alpha > 0$ is the regularization parameter and $\mathrm{Per}(S)$ is the perimeter of S:

$$\mathrm{Per}(S) = |\mathrm{D}(\chi_S)|$$

with $|\mathrm{D}(\cdot)|$ denoting the BV-semi-norm.

Our approach is well-defined as the following theorems show.

Theorem (Existence of a minimizer) *For any $\alpha > 0$ and any $g \in L^2(\partial\Omega)$ there exists a minimizer $(\lambda^*, G^*) \in \Lambda \times \mathcal{L}$ of J_α:*

$$J_\alpha(\lambda^*, G^*) \leq J_\alpha(\lambda, G) \qquad \textit{for all } (\lambda, G) \in \Lambda \times \mathcal{L}.$$

Theorem (Stability) *Let $g_n \to g$ in L^2 as $n \to \infty$ and let (λ^n, S^n) minimize*

$$J_\alpha^n(\lambda, S) = \frac{1}{2}\|F(\lambda, S) - g_n\|_{L^2}^2 + \alpha\mathrm{Per}(S) \ \textit{ over } \Lambda \times \mathcal{L}.$$

Then there exists a subsequence $\{(\lambda^{n_k}, S^{n_k})\}_k$ converging to a minimizer $(\lambda^*, S^*) \in \Lambda \times \mathcal{L}$ of J_α in the sense that

$$\|\lambda^{n_k}\chi_{S^{n_k}} - \lambda^*\chi_{S^*}\|_{L^2} \to 0 \quad \textit{as } k \to \infty.$$

Furthermore, every convergent subsequence of $\{(\lambda^n, S^n)\}_n$ converges to a minimizer of J_α.

Theorem (Regularization property) *Let g be in* range(F) *and let $\delta \mapsto \alpha(\delta)$ where*

$$\alpha(\delta) \to 0 \quad \textit{and} \quad \frac{\delta^2}{\alpha(\delta)} \to 0 \quad \textit{as } \delta \to 0.$$

In addition, let $\{\delta_n\}_n$ be a positive null sequence and $\{g_n\}_n$ such that

$$\|g_n - g\|_{L^2} \leq \delta_n.$$

Then, the sequence $\{(\lambda^n, S^n)\}$ of minimizers of $J^n_{\alpha(\delta_n)}$ possesses a subsequence converging to (λ^+, S^+) where

$$S^+ = \arg\min\{\mathrm{Per}(S) : S \in \mathcal{L} \textit{ s.t. } \exists\lambda \in \Lambda \textit{ with } F(\lambda, S) = g\},$$
$$\lambda^+ \in \{\lambda \in \Lambda : F(\lambda, S^+) = g\}.$$

Furthermore, every convergent subsequence of $\{(\lambda^n, S^n)\}_n$ converges to a pair $(\lambda^\dagger, S^\dagger)$ with above property.

For approximating a minimizer of J_α by a steepest descent-like scheme we need to calculate the derivative of J_α with respect to the domain. To avoid technical difficulties we suppose throughout that the coefficients D, μ are continuously differentiable, Ω is an open C^2-domain, and $S \in \mathcal{S} = \{\Gamma \subset \Omega : \partial\Gamma \in C^2\}$.

Theorem (Derivative of J_α) *for $k \in \mathbb{R}$, $h \in C_0^2(\Omega, \mathbb{R}^3)$ we have that*

$$J'_\alpha(\lambda, S)(k, h) = \big\langle F(\lambda, S) - g, F(k, S) + u' \big\rangle_{L^2(\partial\Omega)} + \alpha \int_{\partial S} \mathrm{H}_{\partial S}(h \cdot \mathbf{n})\, \mathrm{d}s$$

where $\mathrm{H}_{\partial S}$ *denotes the mean curvature of* ∂S *and* $u' \in H^1(\Omega\backslash\partial S)$ *solves the transmission boundary value problem*

$$\begin{aligned} -\nabla\cdot(D\nabla u') + \mu u' &= 0 && \text{in } \Omega\backslash\partial S,\\ 2D\partial_{\mathbf{n}} u' + u' &= 0 && \text{on } \partial\Omega,\\ [u']_{\pm} &= 0 && \text{on } \partial S,\\ [D\partial_{\mathbf{n}} u']_{\pm} &= -\lambda h\cdot\mathbf{n} && \text{on } \partial S. \end{aligned}$$

Next we formulate an approximate variational principle which guarantees existence of smooth almost critical points of J_α near to any of its minimizers. This is a crucial property from a numerical point of view as it allows to approximate the searched-for domain by smooth domains.

Theorem (Approximate variational principle) *Let* $(\lambda^*, S^*) \in \Lambda\times\mathcal{L}$ *be a minimizer of* J_α *where* λ^* *is an inner point of* Λ*. Then, for any* $\varepsilon > 0$ *sufficiently small there is a* $(\lambda^\varepsilon, S^\varepsilon) \in \Lambda\times\mathcal{S}$ *with*

$$J_\alpha(\lambda^\varepsilon, S^\varepsilon) - J_\alpha(\lambda^*, S^*) \le \varepsilon,\ \ \|\lambda^\varepsilon\chi_{S^\varepsilon} - \lambda^*\chi_{S^*}\|_{L^1} \le \varepsilon,\ \ \|J'_\alpha(\lambda^\varepsilon, S^\varepsilon)\|_{\mathbb{R}\times C^2\to\mathbb{R}} \le \varepsilon.$$

For a numerical realization of our approach in 2D we consider star-shaped domains only

$$S = \{x\in\mathbb{R}^2 : x = m + t\,\theta(\vartheta)r(\vartheta),\ 0\le t\le 1,\ 0\le\vartheta\le 2\pi\}$$

where $\theta(\vartheta) = (\cos\vartheta, \sin\vartheta)^\top$, m is the center, and $r\colon [0,2\pi]\to[0,\infty[$ parameterizes the boundary of S. All previous results hold in this setting as well if we work in a space of smooth parameterizations, say, $r \in H^3_{\mathrm{p}}(0,2\pi) \subset C^2_{\mathrm{p}}(0,2\pi)$. We have implemented star-shaped domains using trigonometric polynomials. Our numerical experiments and results are reported in [2].

References

[1] Christopher H. Contag and Brian D. Ross, *It's not just about anatomy: In vivo bioluminescence imaging as an eyepiece into biology*, Journal of Magnetic Resonance Imaging **16** (2002), no. 4, 378–387.

[2] Tim Kreutzmann and Andreas Rieder, *Geometric reconstruction in bioluminescence tomography*, IWRMM preprint 12/06, Dept. of Mathematics, Karlsruhe Institute of Technology, Germany. Download under
`www.math.kit.edu/iwrmm/seite/preprints/media/preprint\%20nr.\%2012-06.pdf`

[3] Ge Wang, Yi Li, and Ming Jiang, *Uniqueness theorems in bioluminescence tomography*, Medical Physics **31** (2004), 2289–2299.

Blind Deconvolution and Nonnegative Matrix Factorization

Christine De Mol
(joint work with Loïc Lecharlier)

We consider linear inverse problems in a discrete setting and in the case where nonnegativity constraints apply. This amounts to solving the equation $Kx = y$, where $y \in \mathbb{R}^n$ is the vector containing the data, $x \in \mathbb{R}^p$ is the vector of nonnegative coefficients describing the unknown object and K is the $n \times p$ matrix with nonnegative elements modelling the link between the two. Since observed data are noisy, the problem is reformulated as the minimization of a cost function (also called contrast or discrepancy) reflecting the statistical properties of the noise.

For the case of Poisson noise, one minimizes the (log-likelihood) cost function (i.e. the Kullback-Leibler divergence between y and Kx)

$$F(x) = \sum_{i=1}^{n} \left[y_i \ln \left(\frac{y_i}{(Kx)_i} \right) - y_i + (Kx)_i \right]. \tag{1}$$

To do so, a classical and popular iterative algorithm is referred to as Richardson-Lucy's in astronomy [5, 4] and EMML (for Expectation-Maximization Maximum Likelihood) in medical imaging. This algorithm is of multiplicative type and the successive iterates ($k = 0, 1, \dots$) are given by $x^{(k+1)} = \dfrac{x^{(k)}}{K^T \mathbf{1}} \circ K^T \dfrac{y}{Kx^{(k)}}$, where $\mathbf{1}$ is a vector of ones, K^T is the transpose of K and we use the Hadamard (entrywise) product $\circ$ and division. When initialized with $x^{(0)} > 0$, positivity is automatically preserved throughout the iteration. To avoid instabilities due to ill-conditioning, regularization by early stopping is usually applied.

A similar multiplicative algorithm for the case of Gaussian noise, i.e. to minimize the least-squares cost function $\|Kx - y\|_2^2$ (where $\|y\|_2^2 = \sum_i |y_i|^2$ denotes the squared ℓ^2-norm), is the so-called Image Space Reconstruction Algorithm (ISRA): $x^{(k+1)} = x^{(k)} \circ \dfrac{K^T y}{K^T K x^{(k)}}$ [1].

Motivated by several applications, we address the blind inversion problem where both x and the linear operator K are unknown and should be recovered from the data. The minimization of the cost function for both unknowns leads to a non-convex optimization problem even if it is convex with respect to x or K separately. The usual strategy, advocated by many authors, is then an alternate minimization on x (with K fixed) and K (with x fixed).

We notice that the problem can be easily generalized to include multiple inputs/unknowns (x becomes a $p \times m$ matrix X) and multiple outputs/measurements (y becomes a $n \times m$ matrix Y). The resulting problem of solving the equation $KX = Y$ is then often referred to as "Nonnegative Matrix Factorization" (NMF), for which alternating ISRA or EMML algorithms have been popularized by Lee and Seung [3]. An important special case is "blind deconvolution" under positivity constraints, which hold e.g. in incoherent optical imaging or for probability densities.

To regularize the problem we use additional penalties on both unknowns, either quadratic (i.e. proportional their squared ℓ^2-norm) or sparsity-enforcing (i.e. proportional to their ℓ^1-norm), or else a combination of both. For example, in the case of Gaussian noise, we minimize, for K and X nonnegative (assuming Y nonnegative as well), the following cost function

$$F(K,X) = \frac{1}{2}\,\|Y - KX\|_F^2 + \frac{\mu}{2}\,\|K\|_F^2 + \lambda\,\|X\|_1 + \frac{\nu}{2}\,\|X\|_F^2 \tag{2}$$

where λ, μ, ν denote the positive regularization parameters, $\|K\|_F^2 = \sum_{i,j} K_{i,j}^2$ is the squared Frobenius norm and $\|X\|_1 = \sum_{i,j} |X_{i,j}|$ the ℓ^1-norm. We notice that the minimization can be done column by column on X and line by line on K and we derive the following multiplicative alternating minimization algorithm

$$K^{(k+1)} = K^{(k)} \circ \frac{Y(X^{(k)})^T}{K^{(k)}X^{(k)}(X^{(k)})^T + \mu K^{(k)}} \tag{3}$$

$$X^{(k+1)} = X^{(k)} \circ \frac{(K^{(k+1)})^T Y}{(K^{(k+1)})^T K^{(k+1)} X^{(k)} + \nu X^{(k)} + \lambda O} \tag{4}$$

to be initialized with arbitrary but strictly positive $K^{(0)}$ and $X^{(0)}$ (O is a $p \times m$ matrix of ones). This algorithm reduces for $\mu = \nu = 0$ to a blind algorithm proposed in [2] and to ISRA for K fixed and $\lambda = \mu = \nu = 0$. Since it can be derived as a Majorization-Minimization scheme through the use of surrogate cost functions, a monotonic decrease of the cost function at each iteration is ensured. Moreover, building upon Zangwill's theory [6], we are able to analyze the convergence properties of this algorithm and in particular to prove convergence of the sequence $(K^{(k)}, X^{(k)})$ of iterates to a stationary point of (2).

A similar algorithm and similar convergence results can be derived for the case of Poisson noise, i.e. when in (2) the least-squares discrepancy between Y and KX is replaced by a Kullback-Leibler divergence. Moreover, the previous framework can also accommodate a penalty on the (smoothed) total variation of X. Finally, we report on work in progress concerning the numerical validation of these iterative schemes as well as possible strategies to accelerate these algorithms.

References

[1] M. E. Daube-Witherspoon and G. Muehllehner, *An Iterative Image Space Reconstruction Algorthm Suitable for Volume ECT*, IEEE Trans. Med. Imag. **5** (1986), 61–66.

[2] P. O. Hoyer, *Non-negative Matrix Factorization with Sparseness Constraints*, J. Mach. Learn. Res. **5** (2004), 1457–1469.

[3] D. D. Lee and H. S. Seung, *Learning the parts of objects by non-negative matrix factorization*, Nature **401** (1999), 788–791.

[4] L. Lucy, *An iteration technique for the rectification of observed distributions*, The Astronomical Journal **79** (1974), 745–754.

[5] W. Richardson, *Bayesian-Based Iterative Method of Image Restoration*, J. Opt. Soc. Amer. **62** (1972), 55–59.

[6] W. I. Zangwill, *Nonlinear Programming: A Unified Approach*, Prentice Hall, 1969.

Detection of qualitative features in statistical inverse problems

Johannes Schmidt-Hieber
(joint work with Axel Munk and Lutz Dümbgen)

In the theory of deterministic inverse problems the main focus lies on stable reconstruction methods. However, by assuming additionally that measurement errors are random we can say more: Besides point estimators/reconstruction methods, tests for statistical hypotheses and confidence statements can be considered as well.

In applied inverse problems one often faces the problem that the shape of the reconstruction may be highly dependent on the regularization parameter. This dependence increases with the ill-posedness of the problem. For example it might happen that some regularization parameter selection rules add an additional mode to the reconstruction. Then, the question is the interpretability. Can we state that there is a mode? In this talk we outline how to assign confidence values to qualitative features such as modes or local increases and decreases for the (one dimensional) nonparametric density deconvolution model.

The application for real data sets is evident. Given a point estimate together with the constructed confidence statements we can distinguish which qualitative features of the reconstruction are significant and which are likely to be artefacts. To state is differently, we construct a tool that allows to substantiate visual impressions from an estimator by confidence values.

In nonparametric settings there are delicate issues involved with confidence sets if the smoothness of the true function is not known a-priori (cf. Low [3]). Nevertheless, confidence statements for qualitative features can be obtained (uniformly over function classes) since they do not involve bias estimates (for nonparametric regression cf. [1], for density estimation, cf. [2], and for density deconvolution, cf. [4]). In order to construct these confidence statements, convergence of a multiscale statistic to a distribution-free limit has to be proved (this is explained in more detail in [5]). The multiscale statistic combines tests computed over all scales. This has a number of advantages. Firstly, construction of the confidence statements does not require the choice of a bandwidth. It is only necessary to select a global confidence level. Secondly, since all scales are taken into account the method is able to detect features of different size. This can be viewed as adaptation property. Thirdly, by construction of the multiscale test, the derived confidence statements hold simultaneously with probability at least the confidence level.

In Figure 1, we give a numerical illustration of our result, based on Laplace deconvolution with sample size 2.000 (we consider the same setting as [4], Section 6). In the upper plot the true density and the (very smooth) convolved density are displayed. As reconstruction method we take a kernel density estimator. In the middle panel of Figure 1, reconstructions for three different bandwidth choices are given. Two of them find a mode around .05, whereas the reconstruction corresponding to the largest bandwidth increases on $[0, .24]$. Can we state that the additional mode is significant and therefore reject the smooth reconstruction? In

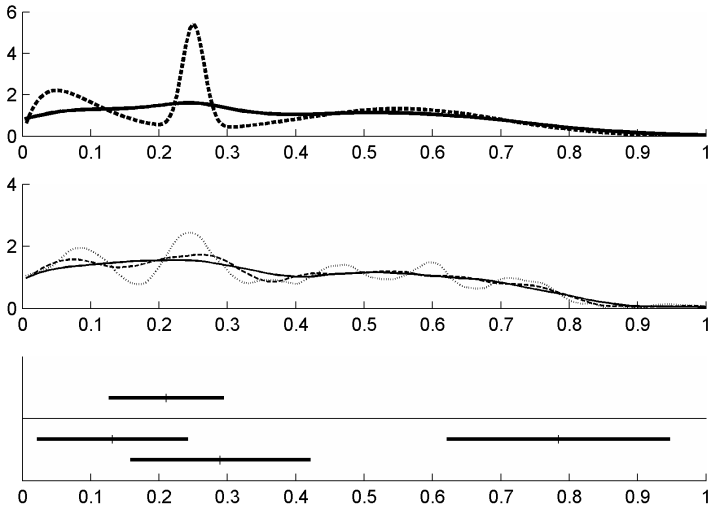


FIGURE 1. Numerical example for Laplace density deconvolution. *Upper panel:* True (dashed) and and observed density (solid). *Middle panel:* Reconstructions for different bandwidth choices. *Lower panel:* Confidence statement.

the lower plot, the confidence statement is displayed. Before interpreting the result, let us shortly explain how to read the plot. Pick one of the horizontal lines. Then, with 90% confidence, we can conclude that *somewhere* on this interval there is an increase or decrease of the true function, depending whether the line is plotted above or below the thin line. By looking at the leftmost horizontal line, we find that with confidence 90% there has to be a decrease somewhere on $[.02, .24]$. Therefore, we can conclude that the reconstruction that increases monotonically on $[0, .24]$ does not have the right shape behavior. On the other hand the additional modes occuring for the oscillating reconstruction cannot be substantiated by the confidence statements and are thus classified as possible artefacts.

REFERENCES

[1] L. Dümbgen and V. Spokoiny (2001), *Multiscale testing of qualitative hypothesis*, Ann. Statist. **29**, 124–152.

[2] L. Dümbgen and G. Walther (2008), *Multiscale inference about a density*, Ann. Statist. **26**, 1758–1785.

[3] M. Low (1997), *On nonparametric confidence intervals*, Ann. Statist. **25**, 2547–2554.

[4] J. Schmidt-Hieber, A. Munk, and L. Dümbgen (2012), *Multiscale methods for shape constraints in deconvolution: confidence statements for qualitative features*, Preprint arxiv.org/abs/1107.1404.

[5] J. Schmidt-Hieber (2012), *Obtaining qualitative statements in deconvolution models*, Oberwolfach Reports **9**, 859–861.

Solution of large scale PDE inverse problems in model reduction framework

Vladimir Druskin

(joint work with Liliana Borcea, Alexander Mamonov, Valeria Simoncini and Mikhail Zaslavsky)

Many transient and steady-state inverse problems of remote sensing, such as electromagnetic and seismic geophysical exploration, electrical impedance tomography (EIT), etc., can be conveniently formulated as multi-input/multi-output (MIMO) problems of control theory. Let us assume, that our measured data are obtained with the help of m_1 transmitters (inputs) and m_2 receivers (outputs), and they can be presented via matrix valued impedance function $Y(s) \in \mathbb{C}^{m_1 \times m_2}$, a.k.a. Weyl function or MIMO transfer function, given by

$$Y(s) = C^* \left[A(\sigma) + s\mathbf{I}\right]^{-1} B. \tag{1}$$

Here $s \in \mathbb{C}$ is the excitation frequency, $A(\sigma)$ is a PDE operator with coefficient distribution σ, e. g., diffusion operator $A(\sigma)u = \nabla \cdot (\sigma \nabla u)$ in $\mathbb{R}^3$ with space-variable diffusion coefficient σ (the unknown in the inverse problem), assuming regular enough function u. Operators B and C have respectively transmitter and receiver density distributions as columns, e. g., for the above mentioned diffusion problem $B^*u = (\int_{\mathbb{R}^3} b_1 u,\ \ldots,\ \int_{\mathbb{R}^3} b_{m_1} u)^*$, $C^*u = (\int_{\mathbb{R}^3} c_1 u,\ \ldots,\ \int_{\mathbb{R}^3} c_{m_2} u)^*$. Let data $D(s)$ (including possible measurement error) be available for some set of frequencies $S \subset \mathbb{C}$ with probability measure dS. We assume, that the solution of the inverse problem σ_{opt} can be uniquely obtained by minimization of the misfit functional on a compact set Σ of admissible solutions σ for $s \in S$, i.e.,

$$\sigma_{min} = \arg\min_{\sigma \in \Sigma} \int \left\| D(s) - C^* \left[A(\sigma) + s\mathbf{I}\right]^{-1} B \right\|_w^2 dS, \tag{2}$$

where $\|.\|_w$ is weighted matrix norm with weights dependent on s. If necessary, a regularization penalty functional (e.g., Tikhonov's) can be added to the r.h.s. of (2). For example, such a formulation includes the *EIT* problem, in which case $S = \{0\}$ and $A(\sigma)$ is the diffusion operator with σ being variable electrical conductivity. If S is the entire imaginary axis, then problem (1) corresponds to the transient inverse problems. Well known examples of such problems are: *transient control source electromagnetic method (tCSEM)* with Maxwell's operator $A(\sigma) = \nabla \times \sigma \nabla \times$; the *inverse dynamic seismic problem*, in which case $A(\sigma)$ is the first order elasticity operator.

The regularized Gauss-Newton (GN) algorithm is the method of choice for the solution of (2). However, it requires multiple calls of forward solver accurately enough approximating $C^* \left[A(\sigma) + s\mathbf{I}\right]^{-1} B$. That can prohibitively expensive, especially for large scale 3D problems. This situation becomes aggravated in the

transient problems, because in addition to very expensive forward solution it requires handling extremely large Jacobian matrices.

Our objective is to reduce inversion complexity (without sacrificing its quality) by reducing dimensionality of the forward solver. For that we consider a so-called reduced order model (ROM) $Y_n(s) \in \mathbb{C}^{m_1 \times m_2}$ of $Y(s)$, given by

$$Y_n(s) = C_n^* \left[A_n + s\mathbf{I}\right]^{-1} B_n, \tag{3}$$

where $C_n \in \mathbb{C}^{n\times m_2}$, $B_n \in \mathbb{C}^{n\times m_1}$, $A_n \in \mathbb{C}^{n\times n}$. In the standard control theory setting $A(s)$ is usually a large finite-dimensional operator, so $Y(s)$ is a high order rational function and $Y_n(s)$ can be viewed as its low order rational approximation (assuming that order n is much smaller then A's dimension), hence the name. Usually Y_n constructed as a rational interpolants (a.k.a. multipole Padé) of $Y(s)$, or using rational Krylov subspaces via Padé-Lanczos connections. The frequency interpolation is not available for the EIT problem, so instead we use the DtN interpolation based on network approximations. As result, the ROMs have exponential convergence similar (or even superior) to spectral methods and sparsity structure similar to the conventional second order finite-volume discretizations, so the ROMS are vastly superior compared to conventional forward solvers from the computational perspective. Original concept of such ROMs became known at the end of the 19th century in context of network synthesis and filter design. The goal of those techniques was to construct analogous electrical network matching a given transfer function as accurately as possible in a targeted frequency range.

Projection operators B_n, C_n and class of discrete ROM operators $\tilde{A}_n$ (not necessary compact) should satisfy the following requirements: A_n should be uniquely recovered as

$$A_{n,opt} = \arg \min_{A_n \in \tilde{A}_n} \int \|D(s) - C_n^* \left[A_n + s\mathbf{I}\right]^{-1} B_n\|_w^2 dS, \tag{4}$$

preferably using a direct layer stripping approach; A_n should mimic a matrix of a dicretized $A(s)$ via finite-volume or Galerkin method and B_n should mimic corresponding discretization of B, etc. for C_n. Order n can play role of regularization parameter in (4).

We interpret the synthetic operator A_n as a discretization of A, so the desired estimate is obtained as cell average. It turned out, that such an estimate strongly depends on the discretization grid, i.,e., $\sigma_{opt,n}$ even diverges or converges to a wrong limit, if the grid is not chosen properly. This property highlights fundamental difference between conventional inversion (2) (or its variants using Tikhonov regularization) and ROM inversion (4). The former explicitly impose restrictions on the class of admissible solutions Σ, such that problem becomes stable to data perturbation, including errors in the forward solution. In contrary, optimization (4) is performed in terms of A_n from (possibly noncompact) admissible set $\tilde{A}_n$, which can not be always translated to an appropriate constrain on σ. That may lead to spurious oscillation of the latter.

Another research direction is to construct A_n via already mentioned Padé-Galerkin connection using the RKS for the transient problems. We found (and

proved), that the ROM approximation error $Y_n - Y$ lays in the null-space of the ROMS Jacobian for a special class of H_2 (Hardy space) -optimal approximations.

Model reduction method for a parabolic inverse resistivity problem

Alexander V. Mamonov
(joint work with Liliana Borcea, Vladimir Druskin and Mikhail Zaslavsky)

We consider a problem of recovery of a resistivity coefficient $r(x) > 0$ in a one-dimensional parabolic equation

$$\partial_x[r(x)\partial_x u(x,t)] = \partial_t u(x,t), \quad x \in [0,1], \quad t > 0, \tag{1}$$

$$u(x,0) = \delta(x), \quad \partial_x u(0,t) = u(1,t) = 0, \tag{2}$$

from the measurement of the time-domain data $d(t) = u(0,t)$. This particular formulation is motivated by the controlled source electromagnetic exploration in geophysics. This highly non-linear inverse problem is ill-posed and thus its numerical solution is challenging. To mitigate the ill-posedness of the problem we put it into the model reduction framework.

We treat the semi-discrete analogue of (1)–(2) as a dynamical system with a transfer function

$$G(s) = \mathbf{b}^T(sI - A(\mathbf{r}))^{-1}\mathbf{b}, \quad A(\mathbf{r}) \in \mathbb{R}^{N\times N}, \quad \mathbf{b} \in \mathbb{R}^N, \tag{3}$$

where $A(\mathbf{r})$ is a fine grid discretization of the operator in (1) that depends linearly on the vector of discrete resistivities $\mathbf{r} \in \mathbb{R}^N_+$. Using the rational Krylov subspace model reduction techniques we obtain a reduced model of (3) that has the form

$$G_m(s) = \mathbf{b}_m^T(sI_m - A_m)^{-1}\mathbf{b}_m,\ A_m = V^T A(\mathbf{r})V \in \mathbb{R}^{m\times m},\ \mathbf{b}_m = V^T\mathbf{b} \in \mathbb{R}^m, \tag{4}$$

with V being an orthonormal basis for a rational Krylov subspace

$$\mathcal{K}_m(\sigma) = \text{span}\{(\sigma_1 I - A(\mathbf{r}))^{-1}\mathbf{b}, \ldots, (\sigma_m I - A(\mathbf{r}))^{-1}\mathbf{b}\}. \tag{5}$$

This corresponds to a rational interpolation of $G(s)$ by $G_m(s)$ at nodes σ_j, the choice of which is discussed in detail in [1].

Since the reduced order transfer function (4) is a rational function of s, one may consider its Stieltjes continued fraction expansion with positive coefficients κ_k, $k = 1, \ldots, 2m$, which can be obtained with Lanczos algorithm. It turns out that these coefficients have a physical meaning of resistivities, as they themselves correspond to coefficients of a discretization of (1)–(2) on a three point stencil [2].

We formulate the inverse problem of finding an estimate $\mathbf{r}^\star$ of the resistivity as a minimization of a functional

$$\mathbf{r}^\star = \arg\min_{\mathbf{r}\in\mathbb{R}^N_+} \frac{1}{2}\left\|\mathcal{Q}(d(t)) - \mathcal{Q}(y(t;\mathbf{r}))\right\|_2^2, \tag{6}$$

where $y(t;\mathbf{r}) = \mathbf{b}^T e^{A(\mathbf{r})t}\mathbf{b}$ is the time-domain response of the dynamical system. The *non-linear preconditioner* $\mathcal{Q}$ maps the time domain response to the vector of the continued fraction coefficients κ_k of the reduced model. If viewed as a function of the fine grid discrete resistivities $\mathbf{r}$, then $\mathcal{Q}(y(\,\cdot\,,\mathbf{r}))$ is an approximate identity,

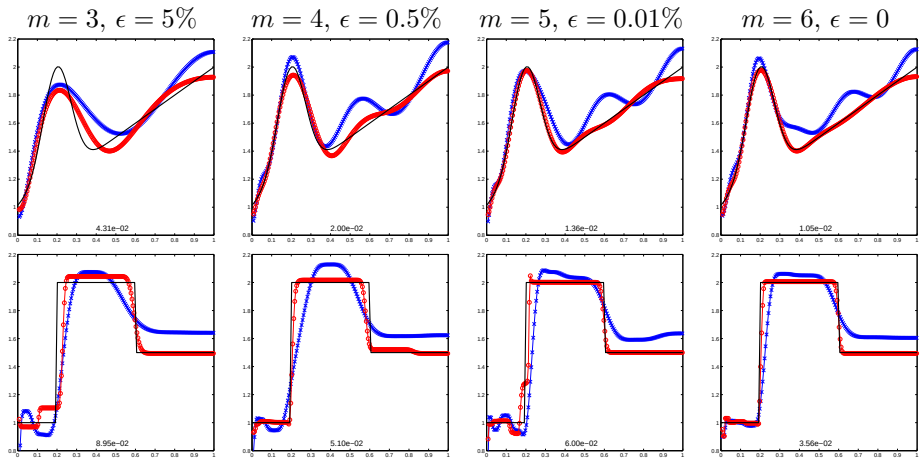


FIGURE 1. Reconstructions of resistivities $r(x)$ (solid black line): smooth (top row), piecewise constant (bottom row). Reconstruction after the first Gauss-Newton iteration is blue $\times$, after five iterations is red $\circ$. Relative ℓ_2 error is printed at the bottom of the plots. Noise level is ϵ, reduced model dimension is m.

since the continued fraction coefficients correspond to the resistivities on a coarser grid with $2m$ nodes. Thus, if we apply the Gauss-Newton iteration to (6) it will converge quickly. It is also much less likely to get stuck in a local minimum, because the functional in (6) is close to convex. These are the significant advantages of our method over the more traditional regularized output least squares approaches, that are prone to slow convergence and local minima. Another advantage comes from the fact that the mapping from $\mathbf{r}$ to κ_k can be computed directly, without the need to calculate the time domain solution $y(t, \mathbf{r})$. This makes our method computationally inexpensive.

The reconstructions from noisy data obtained in [1] with the inversion procedure outlined above are given in Figure 1. We observe that the method performs well for both the smooth and the piecewise constant resistivities. The model reduction approach discussed here can be applied to the inverse problems in higher spatial dimensions. Depending on the setting, such extension may require the study of tangential rational interpolation and block Lanczos algorithms instead of their scalar counterparts used here.

REFERENCES

[1] L. Borcea, V. Druskin, A.V. Mamonov and M. Zaslavsky, *A model reduction approach to numerical inversion for a parabolic partial differential equation*, Preprint: arXiv:1210.1257 [math.NA] (2012).

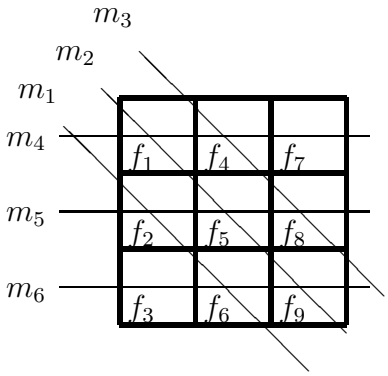
[2] V. Druskin and L. Knizhnerman, *Gaussian spectral rules for second order finite-difference schemes*, Numerical Algorithms **25**:1 (2000), 139–159.

Sparsity-based choice of regularization parameter

Samuli Siltanen
(joint work with Keijo Hämäläinen, Aki Kallonen, Ville Kolehmainen, Matti Lassas and Kati Niinimäki)

We study the so-called *S-curve method*, introduced in [7], in the context of two-dimensional X-ray tomography. The S-curve method can be seen as a generalization of the ideas in [6].

Consider the discrete tomographic measurement model $m = Af + \varepsilon$, where $f \in \mathbb{R}^n$ is a piecewise constant discretization of the unknown attenuation coefficient, $m \in \mathbb{R}^k$ is the vector of measurements, A is the system matrix and ε models measurement noise. We assume that ε is white noise with standard deviation $\sigma > 0$. Here is a simple geometric illustration of such a model:



$$A = \begin{bmatrix} 0 & \sqrt{2} & 0 & 0 & 0 & \sqrt{2} & 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 & 0 & \sqrt{2} & 0 & 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 & \sqrt{2} & 0 & 0 & 0 & \sqrt{2} & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

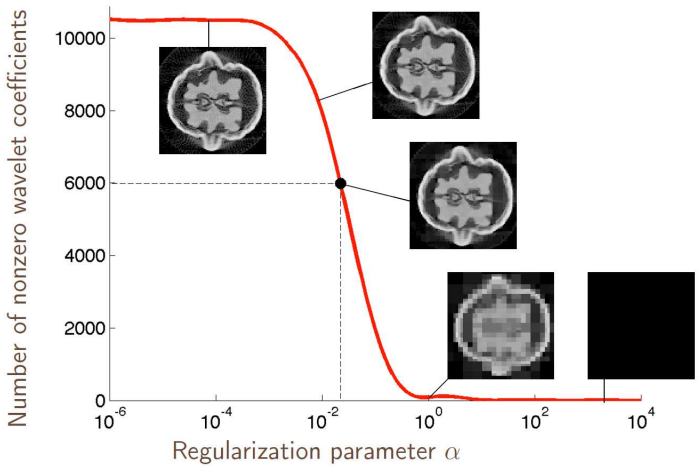
We are interested in computing sparse reconstructions by minimizing the functional

$$\arg\min_{f \in \mathbb{R}^n_+} \arg\min_{f \in \mathbb{R}^n_+} \left\{ \frac{1}{2\sigma^2} \|Af - m\|^2_{\ell^2} + \alpha \|f\|_{B^1_{11}} \right\}, \tag{1}$$

where $0 < \alpha < \infty$ is the regularization parameter and the notation $f \in \mathbb{R}^n_+$ means that we minimize over vectors f having non-negative elements. The Besov space norm in (1) can be written in terms of the wavelet coefficients of f:

$$\|f\|_{B^1_{11}} = \sum_{\vec{k}} |c_{J_0 \vec{k}}| + \sum_{j,\ell,\vec{k}} |w_{j\vec{k}\ell}|.$$

The analysis in [1, 2, 3, 4] shows that the minimizer of (1) is *sparse*, or has only finitely many nonzero wavelet coefficients. Now of course the formulation above is finite in the first place, so it may seem like the sparsity has no meaning. However, one can refine the discretization by taking n larger above; in that case the number of nonzero coefficients will stabilize at some point and remain constant for all $n > n_0$.

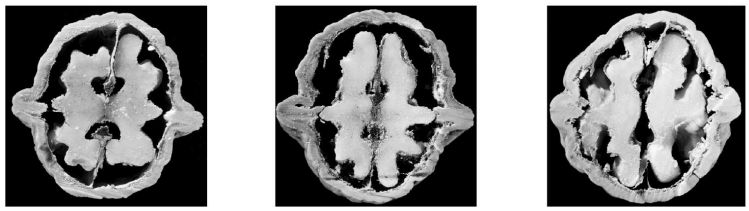


It turns out that the minimization problem (1) can be transformed into the standard form

$$\arg\min_{z\in\mathbb{R}^{5n}} \left\{ \frac{1}{2} z^T Q z + c^T z \right\}, \quad z \geq 0, \quad Ez = b, \tag{2}$$

where Q is symmetric and E implements equality constraints. A large-scale primal-dual interior point quadratic programming (QP) method was developed for (2) in [7, 5].

The idea of the S-curve method is to use *a priori* knowledge of the sparsity of the unknown function to be reconstructed. To demonstrate this idea, we X-ray imaged a walnut from different directions to produce tomographic data. Then we took three other walnuts, sawed them in half, and took photos of the cross-sections:



We then computed the wavelet transform of the above photographs at resolution 128×128 and determined the number of nonzero wavelet coefficients. Next we computed the minimizers of (1) with various choices of α and determined the number of nonzero wavelet coefficients in each case. Now with large α the reconstruction approaches to zero, and in the limit there are no nonzero wavelet coefficients. On the other hand, with small α the reconstructions are very erratic due to too weak regualrization. In such cases almost all of the wavelet coefficients are nonzero. In

between these extreme cases there is some value of alpha giving the right sparsity, and that optimal α can be revealed by the S-curve. See the picture above.

References

[1] E. J. Candès, J. K. Romberg, and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math., 59 (2006), pp. 1207–1223.

[2] I. Daubechies, M. Defrise, and C. De Mol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, Communications on pure and applied mathematics, 57 (2004), pp. 1413–1457.

[3] M. Grasmair, M. Haltmeier, and O. Scherzer., *Sparse regularization with l^q penalty term*, Inverse Problems, 24 (2008), p. 055020 (13pp).

[4] M. Grasmair, M. Haltmeier, and O. Scherzer, *Necessary and sufficient conditions for linear convergence of l^1-regularization*, Communications on Pure and Applied Mathematics, 64 (2011), pp. 161–182.

[5] K. Hämäläinen, A. Kallonen, V. Kolehmainen, M. Lassas, K. Niinimäki and S. Siltanen, *Sparse tomography*, submitted manuscript.

[6] V.K. Ivanov and T.I. Koroljuk, *The estimation of errors in the solution of linear ill-posed problems*, Ž. Vyčisl. Mat. i Mat. Fiz. **9** (1969), 30–41.

[7] V. Kolehmainen, M. Lassas, K. Niinimäki and S. Siltanen, *Sparsity-promoting Bayesian inversion*, Inverse Problems **28** (2012), 02005.

Regularization in Variable RKHSs with application to the Blood Glucose Reading

Valeriya Naumova

(joint work with Sergei V. Pereverzyev, Sivananthan Sampath)

This is the first from two presentations on regularization methods for Diabetes Technology. Here we consider the problem of a reconstruction of a real-valued function $f : X \to \mathbb{R}, X \subset \mathbb{R}^d$, from a given data set

$$\mathbf{z} = \{(x_i, y_i)\}_{i=1}^n \subset X \times \mathbb{R},$$

where it is assumed that $y_i = f(x_i) + \xi_i$, and $\xi_i = \{\xi_i\}_{i=1}^n$ is a noise vector. At this point it should be noted that the reconstruction problem can be considered in two aspects. One aspect is to evaluate the value of a function $f(x)$ for $x \in \overline{co\{x_i\}}$, where $\overline{co\{x_i\}}$ is the closed convex hull of data points $\{x_i\}$. It is sometimes called as interpolation. The other aspect is to predict the value of $f(x)$ for $x \notin \overline{co\{x_i\}}$, which is known as extrapolation.

Due to ill-posedness of the reconstruction problem, a regularization mechanism is required [2]. A well-known technique to stabilize the ill-posed problems with noisy data is by the Tikhonov regularization, i.e., by minimizing the functional

$$T_{\lambda,r}(f) = \frac{1}{|\mathbf{z}|}\sum_{i=1}^{|\mathbf{z}|}(y_i - f(x_i))^2 + \lambda||f||^2_{W_2^r}, \tag{1}$$

where $|\mathbf{z}|$ is the cardinality of the set $\mathbf{z}$, i.e., $|\mathbf{z}| = n$, and λ is a regularization parameter, which trades off data error with smoothness measured in terms of a Sobolev space W_2^r [7].

The Tikhonov method, even in its simplest form (1), raises two main issues that should be clarified before use of this scheme. One of them is how to choose a regularization parameter λ. This problem has been extensively discussed. A few selected references from the literature are [1, 2, 3, 6].

The second issue is concerned with the choice of a space, in which a regularization is performed and whose norm is used for a penalization. It is important to stress that this choice is problem-dependent and can make a significant difference in practice. Despite its significance, the second issue is much less studied and, in general, the question about the proper choice of a regularization space is open until now. At the same time, keeping in mind that a Sobolev space W_2^r used in (1) is a particular example of a *Reproducing Kernel Hilbert Space (RKHS),* the issue about the choice of a proper regularization space is, in fact, about the choice of a kernel for an RKHS. At the same time, as it has been mentioned, for example, by Micchelli, Pontil [4], even for the classical RKHS setting, a challenging and central problem is the choice of the kernel itself, and the choice is tied to the problem of choosing the basis for the approximation of the unknown function.

From the above discussion, it should be clear that despite the great success of the regularization theory, the choice of the suitable regularization space still remains an issue. The present talk aims to shed light on this important but as of yet under-researched problem.

With this in mind, we describe a novel approach, so-called Kernel Adaptive Regularized (KAR) algorithm, where the kernel and the regularization parameter are adaptively chosen within regularization procedure [5].

The construction of such fully adaptive regularization algorithm is motivated by the problem of reading the blood glucose concentration of diabetic patients. To this end, we present an extensive collection of the results of the numerical experiments with real clinical data which confirm the theoretical achievements. Efficiency and superiority of the proposed approach is demonstrated by comparing the performance of the constructed blood glucose estimators with the performance of the commercially available systems for an estimation of the current blood glucose, and, by doing so, we observe and discuss several attractive features of the constructed estimators, which are of interest to both medical researchers and mathematicians.

References

[1] E. De Vito, S. Pereverzyev, L. Rosasco, *Adaptive kernel methods using the balancing principle*, Found. Comput. Math. **10** (2010), 455-479.

[2] H. Engl, M. Hanke, A. Neubauer, *Regularization of Inverse Problems*, volume 375 of Mathematics and Its Applications, Kluwer Academic Publishers, Dordrecht, Boston, London, 1996.

[3] S. Kindermann, A. Neubauer, *On the convergence of the quasi-optimality criterion for (iterated) Tikhonov regularization*, Inverse Problems and Imaging **2** (2008), 291–299.

[4] C. A. Micchelli, M. Pontil, *Learning the Kernel Function via Regularization*, J. Mach. Learn. Res. **6** (2005), 1099-1125.

[5] V. Naumova, S. V. Pereverzyev, S. Sampath, *Extrapolation in variable RKHSs with application to the blood glucose reading*, Inverse Problems **27** (2011), 075010, 13 pp.

[6] A. N. Tikhonov, V. B. Glasko, *Use of the regularization methods in non-linear problems*, USSR Comput. Math. Phys. **5** (1965), 93–107.
[7] G. Wahba, *Spline Models for Observational Data*, volume 59 of Series in Applied Mathematics, CBMS-NSF Regional Conf., SIAM, 1990.

A Meta-Learning Approach to the Adaptive Regularization – Case Study: Blood Glucose Prediction

Sergei V. Pereverzyev
(joint work with Valeriya Naumova, Sivananthan Sampath)

This is the second from two presentations on regularization methods for Diabetes Technology. The positive theoretical and numerical results of the developed KAR-algorithm [3] encouraged us to analyze the problem of the adaptive regularization but this time in the framework of a *meta-learning approach*. The concept of meta-learning presupposes that the kernel and the regularization parameter are selected on the base of previous experience with similar learning tasks. Therefore, selection rules developed in this way are intrinsically problem-oriented. Moreover, meta-learning is very much dependent on the quality of data extracted from previous experience. In the literature [2] it is usually difficult obtaining good results since such data (meta-examples, meta-features) are, in general, very noisy. This gives a good reason for using regularization methods [1] in meta-learning, because these methods are aimed for treating noisy data. Despite the naturalness of this approach, the idea of a combination of meta-learning and regularization seems to be new.

In this way we have developed the *Fully Adaptive Regularized Learning (FARL) approach* [4], which allows a dynamic adjustment of the regularization space and parameter to each particular input. The efficient applicability of the FARL-algorithm is demonstrated on the problem of prediction of the blood glucose (BG) levels of diabetic patients. The developed approach allows the construction of blood glucose predictors which, as it has been demonstrated in the extensive clinical trials, outperform the state of the art. Moreover, we also show how the results of the application of the KAR-algorithm can be used as the input to the FARL-algorithm for achieving more accurate clinical results.

Finally, we discuss the versatility and effectiveness of the proposed approach for other applications from diabetes therapy management.

The material is patent pending, the patent application [5] has been filed jointly by Austrian Academy of Sciences and Novo Nordisk A/S (Denmark).

References

[1] H. Engl, M. Hanke, A. Neubauer, *Regularization of Inverse Problems*, volume 375 of Mathematics and Its Applications, Kluwer Academic Publishers, Dordrecht, Boston, London, 1996.
[2] T.A.F. Gomes, R.B.C. Prudencio, C. Soares, A.L.D. Rossi, A. Carvalho, *Combining meta-learning and search techniques to select parameters for support vector machines*, Neurocomputing **75** (2012), 3 –13.

[3] V. Naumova, S. V. Pereverzyev, S. Sampath, *Extrapolation in variable RKHSs with application to the blood glucose reading*, Inverse Problems **27** (2011), 075010, 13 pp.
[4] V. Naumova, S. V. Pereverzyev, S. Sampath, *A meta-learning approach to the regularized learning – Case study: Blood glucose prediction*, Neural Networks **33** (2012), 181–193.
[5] S. Pereverzyev, S. Sivananthan, J. Randløv, S. McKennoch, *Glucose predictor based on regularization networks with adaptively chosen kernels and regularization parameters*, Patent application EP 11163219.6, filing date April 20, 2011.

On the regularity of fixed-point iterations and practical convergence results

Russell Luke

We consider optimization problems of the form

$$(\mathcal{P}_\alpha) \qquad \underset{x\in C}{\text{minimize}}\ \mathcal{F}_b \circ \mathcal{A}(x) + \alpha\varphi(x)$$

where X and Y are Hilbert spaces and $C \subset X$ is closed. For example: $C = X_+$, $\varphi = TV, \|\cdot\|_p^p$, etc. and either $\mathcal{F}_b(y) := D_\phi(y, b)$ where

$$D_\phi(y, b) := \int_\Omega \phi(y(t)) - \phi(b(t)) - \langle \phi'(b(t)),\ (y(t) - b(t))\rangle\, d(\mu(t)),$$

the Bregman divergence parameterized by ϕ, or

$$\mathcal{F}_b(y) := \iota_{\text{lev}\ \leq\beta(D_\phi(\cdot,b))}(y)$$

where

$$\iota_B(y) := \begin{cases} +\infty & \text{if } y \notin B \\ 0 & \text{else.} \end{cases}$$

Define $\overline{x}_\alpha$ to be a solution to $(\mathcal{P}_\alpha)$ and $\overline{x}_0$ a solution to $(\mathcal{P}_0)$. Let the regularization φ be such that $\overline{x}_\alpha \to \overline{x}_0$ linearly as $\alpha \to 0$. Suppose also that $\overline{x}_\alpha$ is computed by an iterative method, that is, you construct a sequence of iterates $(x_\alpha^k)_{k\in\mathbb{N}}$ with $x_\alpha^k \to \overline{x}_\alpha$. If x_α^k converges to $\overline{x}_\alpha$ arbitrarily slowly for a fixed $\alpha > 0$, the complexity estimates for the regularization scheme are moot. Convergence of the iterative scheme for fixed α is not enough. On the other hand, you will never compute $\overline{x}_\alpha$ exactly *anyway*, so how close is close enough, and how long will it take you to get there?

What could go wrong? Even for nice geometries, one can have bad behavior. In $(\mathcal{P}_\alpha)$ let $C = X = \ell_2$, $\mathcal{A} = Id$, $\mathcal{F}_b(y) := \iota_{\Omega_1}(y)$ and $\varphi(x) = \iota_{\Omega_2}$ where $\Omega_1 = \{e_1\}^\perp$ and Ω_2 is a closed convex cone with $\sup\langle e_1,\ \Omega_2\rangle = 0$ and $\Omega_1 \cap \Omega_2 = \{0\}$. For any $\alpha \geq 0$ problem $(\mathcal{P}_\alpha)$ in this instance is a convex feasibility problem with $\overline{x}_\alpha = \overline{x}$: Given the sets Ω_1 and Ω_2

$$\text{Find} \qquad \overline{x} \in \Omega_1 \cap \Omega_2.$$

We attempt to solve this problem via the method of alternating projections (MAP):

$$x^{2k+1} = P_{\Omega_1} x^{2k} \quad , x^{2k} = P_{\Omega_2} x^{2k-1}$$

where for a general set $\Omega \subset X$,

$$P_\Omega x := \text{argmin}_{y\in\Omega}\|x - y\|.$$

Hundal [5] showed that there is a convex cone Ω_2 for which the sequence of MAP iterates for this problem converge weakly to 0, but not strongly. Indeed, in [2] it was shown for the case of *subspaces* of a Hilbert space X the MAP algorithm for finding $\overline{x} \in \Omega_1 \cap \Omega_2$ converges either uniformly linearly or arbitrarily slowly. The problem is not specific to the MAP algorithm. Indeed, by the well known connection between the projection and gradients of the squared distance function [3] this dichotomy also afflicts the method of steepest descents without step-length optimization applied to the problem of minimizing the sum of squared distances to the subspaces. □

We take a closer look at the regularity assumptions of linear convergence results for simple algorithms such as MAP in an effort to establish linear convergence guarantees for nonconvex instances of this algorithm. In particular, we show the weakest assumptions known at this time to guarantee local linear convergence of the MAP iterates [6, 1]. Using the same notions of regularity we are also able to prove local linear convergence of the more sophisticated Douglas Rachford algorithm in the case where one of the sets is a subspace and the other set is what we call (ϵ, δ)-regular [4]. Following [7] the results are extended to finite precision implementations and demonstrated on the problem of phase retrieval in diffractive X-ray imaging.

References

[1] H. H. Bauschke, D. R. Luke, H. M. Phan, and X. Wang. Restricted normal cones and the method of alternating projections. arXiv:1205.0318, 2012.

[2] Heinz H. Bauschke, Frank Deutsch, and Hein Hundal. Characterizing arbitrarily slow convergence in the method of alternating projections. *Int. Trans. Oper. Res.*, 16(4):413–425, 2009.

[3] F.H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley, New York, 1983. Republished as Vol. 5, Classics in Applied Mathematics, SIAM, 1990.

[4] R. Hesse and D. R. Luke. Nonmonotone fix point mappings and convergence of first-order algorithms. Preprint, 2012.

[5] H. Hundal. An alternating projection that does not converge in norm. *Nonlinear Anal.*, 57:35–61, 2004.

[6] A. S. Lewis, D. R. Luke, and J. Malick. Local linear convergence of alternating and averaged projections. *Found. Comput. Math.*, 9(4):485–513, 2009.

[7] D. R. Luke. Local linear convergence of approximate projections onto regularized sets. *Nonlinear Anal.*, 75:1531–1546, 2012.

Local analysis of the inverse boundary value problem for the Helmholtz equation and iterative reconstruction

Maarten V. de Hoop
(joint work with E. Beretta, L. Qiu, O. Scherzer, X.S. Li. S. Wang, J. Xia)

We consider time-harmonic seismic waves, described by the Helmholtz equation, and view the Dirichlet-to-Neumann map on the earth's surface as the data. We establish conditional Lipschitz stability for the inverse boundary value problem. The stability is obtained for models of the form of linear combinations of piecewise constant functions, naturally admitting the presence of certain conormal singularities. The dimension is determined by the number of linearly independent combinations. The stability constant grows exponentially with the dimension of this model space. We also include attenuation.

We consider then the nonlinear Landweber iteration and obtain a convergence result, assuming Lipschitz stability. Specifically, we obtain a radius of convergence, which depends on the above mentioned stability constant, and a convergence rate. Essentially, the radius reduces rapidly with the dimension of the model space, and hence compressive approximations become an important component of the procedure. We finally introduce a convergent nonlinear projected steepest descent iteration for the case of conditional Lipschitz stability.

To mitigate the growth of the stability constant with dimension on the one hand, and the approximations by sparse model representations with associated errors on the other hand, we introduce a multi-level approach with an associated condition on stability constants and on the approximation errors between neighboring levels to guarantee convergence.

We briefly summarize the computational techniques behind our massively parallel structured direct Helmholtz solver. We make use of a nested dissection based domain decomposition, with separators of variable thickness, and introduce an approximate multifrontal solver by developing a parallel Hierarchically SemiSeparable (HSS) matrix compression, factorization, and solution approach.

Wave propagation in waveguides with random boundaries

Ricardo Alonso
(joint work with Liliana Borcea, Josselin Garnier)

We consider acoustic waves propagating in a waveguide with axis along the range direction z. In general, the waveguide effect may be due to boundaries or the variation of the wave speed with cross-range, as described for example in [7, 5]. We consider here only the case of waves trapped by boundaries, and take for simplicity the case of two dimensional waveguides with cross-section $\mathcal{D}$ given by a bounded interval of the cross-range x. The results extend to three dimensional waveguides with bounded, simply connected cross-section $\mathcal{D} \subset \mathbb{R}^2$.

The pressure field $p(t, x, z)$ satisfies the wave equation

$$\left[\partial_z^2 + \partial_x^2 - \frac{1}{c^2(x)}\partial_t^2\right] p(t, x, z) = F(t, x, z)\,, \tag{1}$$

with wave speed $c(x)$ and source excitation modeled by $F(t, x, z)$. Since the equation is linear, it suffices to consider a point-like source located at $(x_0, z = 0)$ and emitting a pulse signal $f(t)$,

$$F(t, x, z) = f(t)\delta(x - x_0)\delta(z)\,. \tag{2}$$

Solutions for distributed sources are easily obtained by superposing the wave fields computed here. The boundaries of the waveguide are rough in the sense that they have small variations around the values $x = 0$ and $x = X$, on a length scale comparable to the wavelength. Explicitly, we let

$$B(z) \le x \le T(z)\,, \quad \text{where } |B(z)| \ll X,\ |T(z) - X| \ll X, \tag{3}$$

and take either Dirichlet boundary conditions

$$p(t, x, z) = 0\,, \quad \text{for } x = B(z) \text{ and } x = T(z), \tag{4}$$

or mixed, Dirichlet and Neumann conditions

$$p(t, x = B(z), z) = 0\,, \quad \frac{\partial}{\partial n} p(t, x = T(z), z) = 0\,, \tag{5}$$

where n is the unit normal to the boundary $x = T(z)$.

The goal of this work is to quantify the long range effect of scattering at the rough boundaries. More explicitly, to characterize in detail the statistics of the random field $p(t, x, z)$. This is useful in sensor array imaging, for designing robust source or target localization methods, as shown recently in [1] in waveguides with internal inhomogeneities. Examples of other applications are in long range secure communications and time reversal in shallow water or in tunnels [4, 8].

Our approach based on changes of coordinates that straighten the boundary leads to a transformed problem that is similar from the mathematical point of view to that in waveguides with interior inhomogeneities, so we can use the techniques from [7, 6, 2, 4, 3] to obtain the long range statistical characterization of the wave field. However, the cumulative scattering effects of rough boundaries are different from those of internal inhomogeneities. We quantify these effects by estimating in a high frequency regime three important, mode dependent length scales: the scattering mean free path, which is the distance over which the modes lose coherence, the transport mean free path, which is the distance over which the waves forget the initial direction, and the equipartition distance, over which the energy is uniformly distributed among the modes, independently of the initial conditions at the source. We show that the random boundaries affect most strongly the high order modes, which lose coherence rapidly, that is they have a short scattering mean free path. Furthermore, these modes do not exchange efficiently energy with the other modes, so they have a longer transport mean free path. The

lower order modes can travel much longer distances before they lose their coherence and remarkably, their scattering mean free path is similar to the transport mean free path and to the equipartition distance. That is to say, in waveguides with random boundaries, when the waves travel distances that exceed the scattering mean free path of the low order modes, not only all the modes are incoherent, but also the energy is uniformly distributed among them. At such distances the wave field has lost all information about the cross-range location of the source in the waveguide. These results can be contrasted with the situation with waveguides with interior random inhomogeneities, in which the main mechanism for the loss of coherence of the fields is the exchange of energy between neighboring modes [7, 6, 2, 4, 3], so the scattering mean free paths and the transport mean free paths are similar for all the modes. The low order modes lose coherence much faster than in waveguides with random boundaries, and the equipartition distance is longer than the scattering mean free path of these modes.

References

[1] L. Borcea, L. Issa, and C. Tsogka, Source localization in random acoustic waveguides, SIAM Multiscale Modeling Simulations, **8** (2010), 1981-2022.

[2] L. B. Dozier and F. D. Tappert, Statistics of normal mode amplitudes in a random ocean, J. Acoust. Soc. Am. **63** (1978), 353-365; J. Acoust. Soc. Am. **63** (1978), 533-547.

[3] J.-P. Fouque, J. Garnier, G. Papanicolaou, and K. Sølna, *Wave propagation and time reversal in randomly layered media*, Springer, New York, 2007.

[4] J. Garnier and G. Papanicolaou, Pulse propagation and time reversal in random waveguides, SIAM J. Appl. Math. **67** (2007), 1718-1739.

[5] C. Gomez, Wave propagation in shallow-water acoustic random waveguides, Commun. Math. Sci., **9** (2011), 81-125.

[6] W. Kohler, Power reflection at the input of a randomly perturbed rectangular waveguide, SIAM J. Appl. Math. **32** (1977), 521-533.

[7] W. Kohler and G. Papanicolaou, Wave propagation in randomly inhomogeneous ocean, in Lecture Notes in Physics, Vol. 70, J. B. Keller and J. S. Papadakis, eds., Wave Propagation and Underwater Acoustics, Springer Verlag, Berlin, 1977.

[8] W.A. Kuperman, W. S. Hodkiss, H. C. Song, T. Akal, C. Ferla, and D.R. Jackson, Experimental demonstration of an acoustic time-reversal mirror, Journal of the Acoustical Society of America, **103** (1998), 25-40.

Large-Scale Solution of Bayesian Inverse Problems Governed by Wave Propagation

Omar Ghattas

(joint work with Tan Bui-Thanh, Carsten Burstedde, James Martin, Georg Stedler)

Inverse problems governed by wave propagation - in which we search to reconstruct the unknown shape of a scatterer, or the unknown properties of a medium, from observations of waves that are scattered by the shape or medium- play an important role in a number of engineered and natural systems. Our goal is to address the qualification of uncertainty in the solution of the inverse problem by

casting the inverse problem as one in Bayesian inference. This provides a systematic and coherent treatment of uncertainties in all components of the inverse problem, from observations to prior knowledge to the wave propagation model, yielding the uncertainty in the infered medium/shape in a systematic and consistent manner. Unfortunately, state-of-the-art MCMC methods for characterizing the solution of Bayesian inverse problems are prohibitive when the forward problem is expensive and a high dimensional parametrization is employed to describe the unknown medium. We report on recent research aimed at overcoming the mathematical and computational barrier for large-scale Bayesian inverse problems. This includes: parallel adaptive mesh refinement/coarsening algorithms; a high order parallel adaptive *hp*-confirming discontinuous Galerkin method for wave propagation; parallel hyprid algebraic-geometric multigrid for treatment of the regularizing priors; infinite-dimensional formulations of Baysian inverse problems and their consistent finite dimensional discretization; a stochastic Newton MCMC method for solution of the statistical inverse problems; fast low rank randomized SVD approximation of the Hessian based on compactness properties; and applications to synthetic Bayesian inverse wave propagation in whole earth seismology characterized by up to 10^6 earth parameters and 10^9 wave propagation variables, on up to 10^5 processor cores.

Oracle inequalities in inverse problems

Laurent Cavalier

(joint work with Yuri Golubev, Dominique Picard, Alexandre Tsybakov)

Inverse problems appear in many fields of application, from geophysics to medical image processing. The aim is to reconstruct some unknown object (or function) based on noisy indirect observations. From a mathematical point of view this often corresponds to inverting some operator equation : given g, find f such that $g = Af$. Usually A is some linear operator $A : H \to H$, where H is some Hilbert space. The most interesting cases are ill-posed inverse problems where A is not invertible as an operator. These problems are ill-posed in the sense that a small error on g can imply a large one on the "inverse" f.

In this paper, we consider inverse problems with random noise, i.e. that the error on the observation is random. The model is the following

$$Y = Af + \varepsilon\xi,$$

where ξ is a white noise, $0 < \varepsilon < 1$ is a small parameter (the noise level) and Y is the observation.

The study of inverse problems with random noise is a fastly growing field in statistics. One of the most standard (and sometimes implicit) hypothesis is that the operator A is compact. In this case one often uses Singular Value Decomposition (SVD) in order to project the observation Y on some basis appropriate for the operator. An equivalent model is then obtained in the form of a sequence space model in the coefficients space.

Let a finite set Λ of linear estimators be given (Truncated SVD, Tikhonov, Pinsker...). Our aim is to mimic the estimator in Λ that has the smallest risk on the true f. Under general conditions, we show that this can be achieved by simple minimization of unbiased risk estimator (URE), provided the singular values of the operator of the inverse problem decrease as a power law. The main result is a nonasymptotic exact oracle inequality, i.e. that in some sense it mimics the best estimator in a given family. This inequality can be also used to obtain sharp minimax adaptive results. In particular, we apply it to show that minimax adaptation on ellipsoids with polynomial or exponential weights, i.e. on Sobolev or Analytic classes, is possible to obtain, without any loss of efficiency with respect to optimal non-adaptive procedures.

Dealing with compact operators is an interesting and natural assumption in statistics, in the sense that we decompose the observation Y in some basis, and then handle its coefficients. However, one may ask if this fine hypothesis of compactness is really needed, or is just more easy to deal with.

Firstly, SVD is usually used only as a mathematical tool and the method may be computed in a much faster way, from a numerical point of view, for example for Tikhonov regularization, Landweber iteration...

Secondly, compact operators are not necessary, and any continuous linear operator with a known spectral decomposition can be studied, in an analogous way. Instead of using the SVD we use the Spectral Theorem. The eigenvalues of the compact operator A are replaced by the continuous spectrum of the linear operator A.

The main point is to understand that this generalization to non-compact operators is not a theoretical exercise, but that it helps to really understand the problem. The examples deal with the operator of convolution on R and with the estimation of the derivative of some function. The difference from convolution on R with circular convolution, is that the operator is not any more compact and the Fourier basis are not the eigenfunctions. However, from a heuristic point of view, in this framework the Fourier transform is the analog of the Fourier basis. This idea is true and the Fourier coefficients which are the eigenvalues of the circular convolution are replaced here by the Fourier transform on $L^2(R)$ of the convolution kernel r.

Using the unbiased risk estimation approach, in this context, we define our estimator. We prove a precise oracle inequality, i.e. that in some sense it mimics the best estimator in a given family. Thus, we generalize the results to this natural extension.

References

[1] Cavalier, L., Golubev, G.K., Picard, D. and Tsybakov, A.B., *Oracle inequalities for inverse problems*, Annals of Statistics **30**, (2002), 843–874.

[2] Cavalier L., *Inverse problems with non-compact operators*, Journal of Stat. Planning and Inference, **136**, (2006), 390–400.

Nonlinear inverse problems with noisy operators in econometrics

Fabian Dunker
(joint work with Thorsten Hohage, Jean-Pierre Florens, Jan Johannes, Enno Mammen)

We consider a nonlinear ill-posed operator equation

$$\mathcal{F}(\varphi^\dagger) = 0\,,$$

where $\mathcal{F} : \mathfrak{B} \subset \mathbb{X} \to \mathbb{Y}$ maps form a convex set $\mathfrak{B}$ in a Banach space $\mathbb{X}$ to a Hilbert space $\mathbb{Y}$. In addition, we assume that the operator is not known exactly. Only some estimate $\widehat{\mathcal{F}} : \mathfrak{B} \to \widehat{\mathbb{Y}}$ with random noise is given. This setup allows for convex constraints on the solution $\varphi^\dagger$ which can be incorporated in the set $\mathfrak{B}$.

We analyze for this problem the convergence of the iteratively regularized Gauss-Newton method with a convex lower semi-continuous penalty functional $\mathcal{R}$ and quadratic data misfit

$$\widehat{\varphi}_{j+1} := \operatorname*{argmin}_{\varphi\in\mathfrak{B}} \left(\|\widehat{\mathcal{F}}_n'[\widehat{\varphi}_j](\varphi - \widehat{\varphi}_j) + \widehat{\mathcal{F}}_n(\widehat{\varphi}_j)\|^2 + \alpha_j \mathcal{R}(\varphi) \right).$$

Using a variational source condition similar to [7], we derive a convergence rate result, for convergence in probability. The source condition is weak in the sense that it actually holds for exponentially ill-posed problems. The proved rates are known to be order optimal in special cases.

Among other applications, problems like this appear in nonparametric instrumental quantile regression considered in [2] and [6], as well as in nonparametric regression with instrumental variables. In this econometric applications convex constraints like monotonicity or concavity are often imposed on the solution.

Let us focus on the latter regression problems. We propose the new nonparametric instrumental regression model

$$Y = \varphi(Z) + U \qquad \text{while } E[U] = 0 \text{ and } U \perp\!\!\!\perp W \tag{1}$$

which assumes that the instrumental variable W is independent of the error term U. If the joint density f_{YZW} of the explanatory variable Z, the response variable Y, and the instrument W exists, this leads to a nonlinear integral equation

$$\mathcal{F}(\varphi)(u,w) := \begin{pmatrix} \int f_{YZW}(u+\varphi(z),z,w) - f_{YZ}(u+\varphi(z),z) f_W(w)\,dz \\ \int y f_Y(y)\,dy - \int \varphi(z) f_Z(z)\,dz \end{pmatrix}.$$

Here, f_{YZ} is the joint density of Y and Z while f_Y, f_Z, and f_W denote the marginals. Since this densities are not known exactly in applications but have to be estimated from data, the operator is only given with random errors. In addition, as probability densities are typically very smooth, the problem is often severely ill-posed in practice.

We compare the new regression model (1) to the more standard instrumental regression model

$$Y = \varphi(Z) + U \qquad \text{while } E[U|W] = 0 \tag{2}$$

considered in [4], [3], [8], [5], and [1] among others. Here, the independence assumption in (1) is weakened to mean independence. This leads to a linear inverse problem with noisy operator

$$\int \varphi(z) f_{Z,W}(z,y)\,dz = \int f_{Y,W}(y,w)\,dy.$$

We carried out numerical experiments for the case of a binary instrument W and real-valued Z and Y. This provides an example where the nonlinear model (1) yields good results whereas in the linear model (2) the solution is not identifiable. Further more, numerical experiments were made where all three random variables are one-dimensional and both models can identify the solution. In this examples model (1) produced considerably better results.

References

[1] X. Chen and M. Reiss. On rate optimality for ill-posed inverse problems in econometrics. *Econometric Theory*, 27(3):497–521, 2011.

[2] V. Chernozhukov, G. W. Imbens, and W. K. Newey. Instrumental variable estimation of nonseparable models. *J. Econometrics*, 139(1):4–14, 2007.

[3] S. Darolles and Y. Fan and J.P. Florens and E.- Renault Nonparametric Instrumental Regression. *Econometrica*, 79(5):1541–1565, 2011.

[4] J.-P. Florens. Inverse problems and structural economics: The example of instrumental variables. In M. Dewatripont, L. P. Hansen, and S. Turnovsky, editors, *Advances in Economics and Econometrics: Theory and Applications*, pages 284–311. Cambridge Univ. Press, 2003.

[5] P. Hall and J. L. Horowitz. Nonparametric methods for inference in the presence of instrumental variables. *Ann. Statist.*, 33:2904–2929, 2005.

[6] J. L. Horowitz and S. Lee. Nonparametric instrumental variables estimation of a quantile regression model. *Econometrica*, 75(4):1191–1208, 2007.

[7] B. Kaltenbacher and B. Hofmann. Convergence rates for the iteratively regularized Gauss-Newton method in Banach spaces. *Inverse Problems*, 26(3):035007, 21, 2010.

[8] W. K. Newey and J. L. Powell. Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578, 2003.

Convergence Rates for Tikhonov Regularisation on Banach Spaces

Markus Grasmair

The goal of this talk is the presentation of a general approach to the derivation of convergence rates for non-smooth Tikhonov regularisation on Banach spaces. We assume that $F\colon U \to V$ is a, possibly non-linear, operator between two Banach spaces U and V, the first of which is assumed to be reflexive. Given noisy data $v^\delta \in V$, we define an approximate solution of the operator equation $F(u) = v^\delta$ as any minimiser of the *Tikhonov functional*

$$\mathcal{T}(u;\alpha,v^\delta) := \frac{1}{2}\bigl\|F(u) - v^\delta\bigr\|^2 + \alpha\mathcal{R}(u).$$

Here $\mathcal{R}\colon U \to [0,+\infty]$ is a *regularisation term* encoding *a–priori* information about the noise-free solution of the equation, and $\alpha > 0$ is a regularisation parameter determining the emphasis put on regularity versus data explanation. Assume now

that the noise level $\delta := \|v^\delta - v\|$ is known. Then it is known that, under some additional conditions on $\mathcal{R}$ and F, the regularised solution $u_\alpha^\delta \in \arg\min_u \mathcal{T}(u; \alpha, v^\delta)$ converge to the *true solution*

$$u^\dagger := \arg\min\{\mathcal{R}(u) : u \in U \text{ satisfies } F(u) = v\}$$

as $\delta \to 0$ and $\alpha \to 0$ in a suitable way. In this talk we treat the question of the *speed* of this convergence.

A classical result in the context of linear operators between Hilbert spaces states that, if $u^\dagger$ satisfies the *source condition* $u^\dagger \in \mathrm{ran}(F^*F)^\nu$ for some $0 < \nu \leq 1$, then the estimate $\|u_\alpha^\delta - u^\dagger\| = O(\delta^{2\nu/(2\nu+1)})$ holds for a parameter choice $\alpha \sim \delta^{2/(2\nu+1)}$. In the special cases $\nu = 1/2$ and $\nu = 1$ we obtain convergence rates with respect to the norm of order $O(\delta^{1/2})$ and $O(\delta^{2/3})$, respectively. These results can also be extended to a non-linear setting on a Hilbert space by linearising the function F and then applying the results from the linear case. This, however, requires that the operator F is sufficiently smooth and very well described by a linear mapping in a neighbourhood of $u^\dagger$.

In any case, the approach usually taken in Hilbert spaces is not easily made to work in the general setting of Tikhonov regularisation that was introduced above. First, the fact that a general Banach space cannot identified with its dual makes the definition of fractional powers of an operator (in fact already integer powers) difficult if not impossible. In addition, as we allow quite general (lower semi-continuous and coercive) regularisation terms $\mathcal{R}$, there is no reason why we should be able to obtain convergence rates with respect to the norm at all.

As an alternative, we propose to base the derivation of convergence rates on *variational inequalities* of the form

$$\mathcal{D}(u, u^\dagger) \leq \mathcal{R}(u) - \mathcal{R}(u^\dagger) + \gamma \|F(u) - F(u^\dagger)\|^\mu, \tag{1}$$

where $\mathcal{D}\colon U \times U \to [0, +\infty]$ is some distance like measure (for instance a power of the norm or the Bregman distance), $\gamma > 0$, and $0 < \mu \leq 2$. If such an inequality holds for all u sufficiently close to $u^\dagger$ (more precisely: all $u \in U$ that can be minimisers of the Tikhonov functional for small α and δ), then one can show that $\mathcal{D}(u_\alpha^\delta, u^\dagger) = O(\delta^\mu)$ for a parameter choice $\alpha \sim \delta^{2-\mu}$ (see [1, 3]). Some instances where this approach can be successfully applied are the following:

- Assume that U and V are Hilbert spaces, F is bounded linear, and $\mathcal{R}(u) = \|u\|^2$. Then (1) holds with $\mathcal{D}(u, u^\dagger) = \beta\|u - u^\dagger\|^2$ and $\mu = 4\nu/(2\nu + 1)$ if $u^\dagger \in \mathrm{ran}(F^*F)^\nu$ for some $0 < \nu \leq 1/2$. Thus the classical rates of lower order are recoverable with this approach.
- Assume that $U = \ell^2$, F is bounded linear, and $\mathcal{R}(u) = \|u\|_{\ell^1}$. If $u^\dagger$ is sparse, satisfies a restricted injectivity property, and the source condition $\partial\mathcal{R}(u^\dagger) \cap \mathrm{ran}F^* \neq \emptyset$ holds, then (1) holds with $\mathcal{D}(u, u^\dagger) = \beta\|u - u^\dagger\|_{\ell^1}$ and $\mu = 1$. Thus $\|u_\alpha^\delta - u^\dagger\|_{\ell^1} = O(\delta)$ for $\alpha \sim \delta$ (see [5]).
- Assume that $U = \ell^2$, F is locally Lipschitz and weakly continuous, and $\mathcal{R}(u) = \|u\|_{\ell^p}$ with $0 < p < 1$. If $u^\dagger$ is the unique $\mathcal{R}$-minimising solution of $F(u^\dagger) = v$, is sparse, and the restriction of F to the span of the support of $u^\dagger$ has a Lipschitz continuous inverse, then (1) holds with $\mathcal{D}(u, u^\dagger) =$

$\beta\|u - u^\dagger\|_{\ell^1}$ and $\mu = 1$, which implies a linear convergence rate. Thus it is possible to obtain convergence rates even in non-convex cases (see [2, 3] for results in this direction).

- The proposed approach can also be applied to non-differentiable functions. Consider for instance $F\colon L^2_{\geq 0}(S^1) \to L^2(S^1 \times \mathbb{R}_{\geq 0})$, $F(u)(t,z) = 1$ for $u(t) \leq z$ and $F(u)(t,z) = 0$ else. This function, related to the *snake medel* in image processing, is Hölder continuous of degree $1/2$, but nowhere differentiable. Still it is possible to derive variational inequalities of the form (1) with $\mathcal{R}(u) = \|u'\|_2^2$ and $\mathcal{D}(u, u^\dagger) = \beta\|(u - u^\dagger)'\|_2^2$, if $u^\dagger$ satisfies smoothness assumptions of the form $u^\dagger \in W^{s,q}(S^1)$. If, for instance, $u^\dagger \in W^{2,\infty}(S^1)$, then a linear convergence rate with respect to the norm can be derived (see [4]).

References

[1] R. I. Boţ and B. Hofmann. An extension of the variational inequality approach for obtaining convergence rates in regularization of nonlinear ill-posed problems. *J. Integral Equations Appl.*, 22(3):369–392, 2010.

[2] M. Grasmair. Non-convex sparse regularisation. *J. Math. Anal. Appl.*, 365(1):19–28, 2010.

[3] M. Grasmair. Generalized Bregman distances and convergence rates for non-convex regularization methods. *Inverse Probl.*, 26(11):115014, 2010.

[4] M. Grasmair. An application of source inequalities for convergence rates of tikhonov regularization with a non-differentiable operator. arXiv:1209.2246, University of Vienna, Austria, 2012.

[5] M. Grasmair, M. Haltmeier, and O. Scherzer. Sparse regularization with l^q penalty term. *Inverse Probl.*, 24(5):055020, 13, 2008.

Convergence rates for cyclic iterative regularization methods

Antonio Leitão

(joint work with Stefan Kindermann)

This talk is devoted to the convergence analysis of a special family of iterative regularization methods for solving systems of ill–posed operator equations in Hilbert spaces, namely the Kaczmarz type methods.

The analysis is focused on the Landweber-Kaczmarz (LK) explicit iteration an the iterated-Tikhonov-Kaczmarz (iTK) implicit iteration. The corresponding symmetric versions of these iterative methods are also investigated (sLK and siTK).

We prove convergence rates for all four iterative methods above, extending and complementing the convergence analysis established originally in [1, 2, 3, 4].

References

[1] J. Baumeister, A. DeCezaro, A. Leitão, *Modied iterated Tikhonov methods for solving systems of nonlinear ill-posed equations*, Inverse Probl. Imaging **5** (2011), 1–17.

[2] M. Haltmeier, A. Leitão, and O. Scherzer, *Kaczmarz methods for regularizing nonlinear ill-posed equations. I: Convergence analysis*, Inverse Probl. Imaging **1** (2007), 289–298.

[3] M. Haltmeier, *Convergence analysis of a block iterative version of the loping Landweber-Kaczmarz iteration*, Nonlinear Anal. **71** (2009), e2912–e2919.

[4] R. Kowar, O. Scherzer, *Convergence analysis of a Landweber-Kaczmarz method for solving nonlinear ill-posed problems*, Ill posed and inverse problems (book series), **23** (2002), 69–90.

TV-denoising and the evolution of sets - the magnadoodle approach

Christiane Pöschl
(joint work with Vicent Caselles and Matteo Novaga)

The purpose of this work is to compute **explicit solutions** of the total variation denoising problem

$$\text{minimize}_{u\in BV(\mathbb{R}^2)\cap L^2(\mathbb{R}^2)} \int_{\mathbb{R}^2} \frac{1}{2}\left|u-\chi_S\right|^2 dx \; + \lambda TV_{l^2}(u) \tag{1}$$

where $S \in \mathbb{R}^2$ and $TV_{l^2}(u)$ is the total variation of u, defined by

$$TV_{l^2}(u) := \sup\left\{ \int u\nabla\cdot\psi, \psi \in \mathcal{C}_0^1, |\psi(x)|_{l^2} \leq 1, \text{for all } x\in\Omega \right\} .$$

While the solution in [1] is obtained by an explicit computation, we describe it by the use of more generic **geometric arguments**. Our starting point is the observation that u_λ is a solution of (1) if and only if the sets $[u_\lambda \geq s]$ minimize the variational problem

$$\mathcal{F}_{s,\lambda}(X) := P(X) + \frac{s}{\lambda}\left|X\setminus S\right| - \frac{(1-s)}{\lambda}\left|X\cap S\right| \quad s\in[0,1], \lambda>0 \, , \tag{2}$$

where $P(X)$ is the perimeter of X and $|X|$ is the area [3, 1]. Let us point out that for $\lambda > 0$ fixed, the solutions of (2) are monotonously decreasing as s increases and can be then packed together to build up a function which solves (1) [2, 3]. Thus, in order to compute u_λ, the minimizer of (1), we study solutions of (2) for any value of $\lambda > 0$ and $s \in [0, 1]$, which can be constructed by means of geometric arguments.

We introduce the morphological imaging operators: closing $Close_r(X)$ and opening $Open_r(X)$ (see [4]). Iterating these operators, we construct sets $\Gamma_{s,\lambda}(S)$ that have the following properties: $\Gamma_{s,\lambda}(S)$ has smooth boundary that is either equal to the boundary of S or an arc of a circle with radius $\frac{\lambda}{1-s}$ inside S and an arc of a circle with radius $\frac{\lambda}{s}$ outside S. Hence $\Gamma_{s,\lambda}(S)$ satisfies the Euler Lagrange conditions of (2), and thus it is a possible minimizer of (2). In some cases the minimizers of (2) are not unique, so it remains to find all minimizers (there is only a finite number of them) and compare their energies.

We consider the following examples:

- In the case where S is a **convex set**, $\Gamma_{s,\lambda}(S)$ is given by the opening of S with parameter $\frac{\lambda}{1-s}$. Moreover we are in the situation where $\Gamma_{s,\lambda}(S)$ and the emptyset are the only sets that satisfy the Euler-Lagrangeconditions. We observe that the corners of $L^2 - TV$ minimizers are blurred.

- In the case where S is a **star-shaped** set such that $Open_{\frac{\lambda}{1-s}}(S)$ is a connected set, we construct the sets $\Gamma_{s,\lambda}(S)$ by iterating the opening and closing operators, composed with the intersection operator.
- In the case where S is the union of two convex sets S_1, S_2, there are several non-empty sets, that satisfy the Euler-Lagrange conditions, for instance: $Open_{\frac{\lambda}{1-s}}(S_1), Open_{\frac{\lambda}{1-s}}(S_2), Open_{\frac{\lambda}{1-s}}(S_1 \cup S_2), Close_{\frac{\lambda}{s}}(S)$. Additionally we construct a set $\Gamma_{s,\lambda}(S_1, S_2)$, that has non-empty intersection with S_1 and S_2 and and $\Gamma_{s,\lambda}(S) \setminus S \neq \emptyset$. The difficult part is to show that $\Gamma_{s,\lambda}(S_1, S_2)$ is the set with lowest energy concerning all sets with these properties.

  However for simple examples, we manage to do so and in these cases we can observe that $L^2 - TV$-minimization has a smearing effect, that is u_λ can be non-zero outside of S, meaning that the edges between two sets are not preserved by $L^2 - TV$ minimization.

With this deeper inside on explicit solutions of the $L^2 - TV$-problem, we are able to calculate dual TV-semi norms explicitly. The G-space (introduced by Y. Meyer [5]) is said to be a good space to model textures, hence it is of high interest in the imaging-community. This space is defined by

$$G := \left\{ v : v = \nabla \cdot \vec{v}, \vec{v} \in L^\infty(\mathbb{R}^2, \mathbb{R}^2) \right\},$$

with the norm $|u|_G := \inf \{ |\vec{v}|_\infty : v = \nabla \vec{v} \}$. This dual TV-seminorm is connected to our set-minimization problem with $s = 0$. Finding a set X such that $\frac{|X \cap S|}{P(X)}$ is maximal is equivalent to calculating the G-norm of χ_S or finding the smallest λ such that $u_\lambda = 0$. Since we know how to construct the zero-levelsets of u_λ, calculating the $G-$norm of χ_S reduces to maximizing $\rho(S_1, S_2), \rho(S_1), \rho(S_2)$, where $\rho(S_1, S_2)$ satisfies

$$\rho(S_1, S_2) = \frac{\left| \Gamma_{0,\rho(S_1,S_2)}(S_1, S_2) \cap (S_1 \cup S_2) \right|}{P\left(\Gamma_{0,\rho(S_1,S_2)}(S_1 \cup S_2) \right)} .$$

and $\rho(S_i) = \frac{\left| \Gamma_{0,\rho(S_i)}(S_i) \cap S \right|}{P\left(\Gamma_{0,\rho(S_i)}(S_i) \right)}$, $i = 1, 2..$ With this we obtain a simple geometric interpretation of the dual TV-seminorm (G-norm) which allows us to calculate it without solving $\chi_S = \nabla \cdot \vec{v}$.

References

[1] W. Allard. *Total variation for image denoising: III. Examples.* SIAM Journal on Imaging Sciences, vol. 2 (2009).

[2] F. Alter, V. Caselles, A. Chambolle. *A characterization of convex calibrable sets in* $\mathbb{R}^N$. Math. Ann. **332**, 329-366 (2005).

[3] V. Caselles, A. Chambolle, and M. Novaga. *The discontinuity set of solutions of the TV denoising problem and some extensions.* SIAM Mult. Model. Simul. **6**, 879–894, 2008.

[4] P. Soille: *Morphological Image Analysis: Principles & applications.* Springer-Verlag, 2003.

[5] Y. Meyer. *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures. University Lecture Series. American Mathematical Society, 2001.*

Regularization of diffusion weighted MRI-data without blurring the geometrical structure

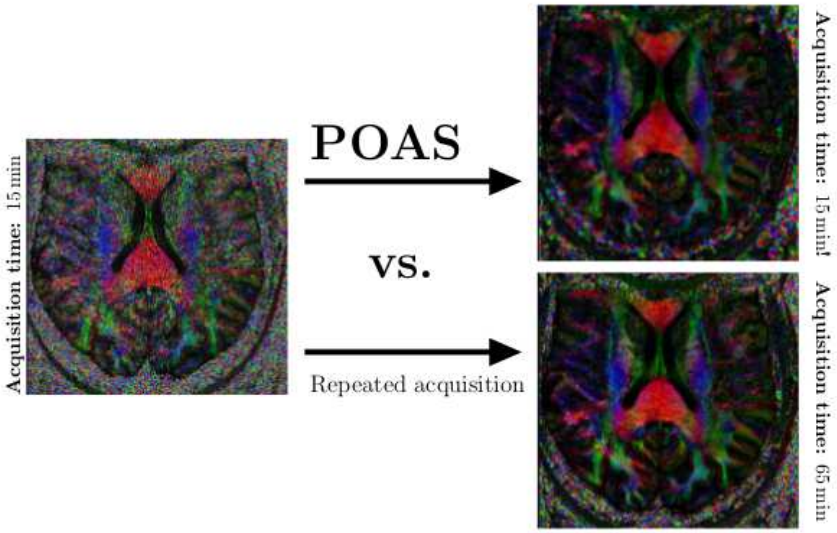
Saskia M. A. Becker

Motivation. Diffusion weighted magnetic resonance imaging (dMRI) is an important tool for in-vivo exploration of micro-structure in the human body [4]. Subsequent evaluation of the data and as a consequence medical decisions are complicated by the significant noise from which dMRI suffers. Simple smoothing methods may blur the structures observed with dMRI. This can be avoided using position-orientation adaptive smoothing (POAS), see [1]. Using the example of dMRI-data, we disuss the importance and consequences of an appropriate handling of the specific geometry when regularizing geometrical structures.

The measurement space. The measurement process of dMRI is based on the pulsed gradient spin echo sequence [7]. Here, data are acquired on a three-dimensional regular grid for varying directions of the diffusion magnetic field gradient. Hence, dMRI data can be described by a function $S : \mathbb{R}^3 \times \mathbb{S}^2 \to \mathbb{R}$. Taking into account the specific geometry of the measurement space we can benefit from the whole information of the data in position and orientation. We follow the approach in [2, 3], where dMRI-data has been interpreted as orientation score and analyzed by embedding $\mathbb{R}^3 \times \mathbb{S}^2$ into the special Euclidean motion group $\mathrm{SE}(3) = \mathbb{R}^3 \rtimes \mathrm{SO}(3)$. This approach suggests to apply solely left-invariant operations on the orientation score S.

Position-orientation adaptive smoothing (POAS) for dMRI. POAS takes the specific geometry of the measurement space into account and reduces noise without blurring the observed structures. The method is left-invariant and well-defined w.r.t. the embedding of $\mathbb{R}^3 \times \mathbb{S}^2$ into $\mathrm{SE}(3)$. It is based on the Propagation-Separation approach [6] which relates to Lepski's method [5]. The Propagation-Separation approach is especially powerful in case of large homogeneous regions and sharp discontinuities as they appear in dMRI-data of the human brain. The pointwise estimator is defined as a weighted mean of the observations. The adaptive weights are calculated as the product of two kernel functions. The adaptation kernel compares the pointwise estimates of the previous iteration step enforcing zero weights in case of significantly distinct values. This yields similar results as non-adaptive smoothing within the homogeneity regions (propagation) and avoids smoothing at structural borders (separation). The location kernel determines the neighborhood under consideration, which is extended during iteration according to a pre-selected sequence of bandwidths. The increasing number of included observations enables a monotone variance reduction during iteration, while the adaptation kernel leads to a decreasing or, in case of model misspecification, bounded estimation bias.

Numerical results. The algorithm POAS has been evaluated on simulated and experimental data. It significantly improves the quality of dMRI-data without blurring the observed structure. In particular, it can be used to reduce acquisition time as illustrated in the following figure showing one representative color coded fractional anisotropy (FA) map generated by estimating the diffusion tensor from



the experimental diffusion weighted data for one subject (Left: Original data, Right: Data after smoothing with POAS, and average image over four subsequent measurements as reference to a kind of ground truth). More details can be found in [1].

Acknowledgements. The author would like to thank Peter Mathé, Jörg Polzehl, and Karsten Tabelow (WIAS Berlin) for helpful discussions.

References

[1] S.M.A. Becker, K. Tabelow, H.U. Voss, A. Anwander, R.M. Heidemann, and J. Polzehl, *Position-orientation adaptive smoothing of diffusion weighted magnetic resonance data (POAS)*, Med. Image Anal. **16(6)** (2012), 1142–1155.

[2] R. Duits and E. Franken, *Left-invariant diffusions on the space of positions and orientations and their application to crossing-preserving smoothing of HARDI images*, International Journal of Computer Vision **92(3)** (2011), 231–264.

[3] E. M. Franken, *Enhancement of crossing elongated structures in images*, PhD thesis, Eindhoven University of Technology (2008).

[4] H. Johansen-Berg and T. E. J. Behrens, *Diffusion MRI: From Quantitative Measurement to In-Vivo Neuroanatomy*, Academic Press (2009).

[5] O. V. Lepskiĭ, *A problem of adaptive estimation in Gaussian white noise*, Teor. Veroyatnost. i Primenen. **35(3)** (1990), 459–470.

[6] J. Polzehl and V. Spokoiny, *Propagation-separation approach for local likelihood estimation*, Probability Theory and Related Fields **135** (2006), 335–362.

[7] E. O. Stejskal and J. E. Tanner, *Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient*, J Chem Phys **42** (1965), 288–292.

Numerical realization of Tikhonov regularization: appropriate norms, implementable stopping criteria, and optimal algorithms

Herbert Egger

A stable approximate solution to linear and nonlinear inverse problems can be obtained by Tikhonov regularization [3]. The nonlinear case can also be reduced to successive minimization of linearized Tikhonov functionals leading to Newton-type regularization methods [7, 10]. For problems of practical interest, the application of the forward operator usually involves the solution of partial differential or integral equations. Therefore, already the minimization of the linear or linearized Tikhonov functionals has to be realized by iterative methods in practice.

We consider linear inverse problems

$$Tx = y^\delta \tag{1}$$

with bounded linear operators $T : X \to Y$ acting on Hilbert spaces X, Y and perturbed data y^δ with bound $\|y - y^\delta\| \le \delta$ on the data perturbation. Here $y = Tx^\dagger$ denotes the correct data for the true solution $x^\dagger$. We apply Krylov subspace methods [6] for the iterative minimization of the Tikhonov functional

$$J_\alpha^\delta(x) := \|Tx - y^\delta\|^2 + \alpha\|x\|^2 \tag{2}$$

with positive regularization parameter α, including reliable stopping criteria that guarantee the optimal error estimates provided by regularization theory. A key ingredient for our analysis is to measure errors in the *energy norm* defined by

$$|||x|||^2 := \|x\|^2_{T^*T+\alpha I} := \|Tx\|^2 + \alpha\|x\|^2. \tag{3}$$

If x_α^δ denotes the minimizer of the Tikhonov functional and $x^\dagger$ satisfies a source condition $x^\dagger = (T^*T)^\mu w$, then regularization theory guarantees that

$$|||x^\dagger - x_\alpha^\delta|||^2 \preceq \alpha^{2\mu+1}\|w\|^2 + \delta^2 =: \rho^2(\mu, \|w\|, \delta). \tag{4}$$

A simple argument shows that $|||x - x^\dagger||| \preceq \rho(\mu, \|w\|, \delta) + \eta$ provided that

$$J_\alpha^\delta(x) - J_\alpha^\delta(x_\alpha^\delta) = |||x - x_\alpha^\delta|||^2 \le \eta^2, \tag{5}$$

hence an optimal approximation for the regularized solution x_α^δ can be found by choosing $\eta \preceq \rho(\mu, \|w\|, \delta)$; see also [4, 5]. Since the regularization parameter can always be bounded from below by $\alpha \succeq \delta^{2/(2\mu+1)}$, a sufficient condition to optimal error estimates for x is to require $\eta \approx \delta$. We will call an iterative method *optimal* (under all iterative methods working on the same Krylov subspace) if (5) can be guaranteed in the minimal number of iterations.

Formally, the conjugate gradient method applied to the optimality system

$$(T^*T + \alpha I)x_\alpha^\delta = T^*y^\delta \tag{6}$$

minimizes the error $|||x - x_\alpha^\delta|||$ on the Krylov space $\mathcal{K}_k(T^*T + \alpha I, T^*y^\delta)$ and thus is the fastest method to reach (5). Since x_α^δ is not known, (5) cannot be used as a stopping rule in practice. Due to $|||x - x_\alpha^\delta||| \le \alpha^{-1/2}\|(T^*T + \alpha I)x - T^*y^\delta\|$, one can however utilize the following criterion

$$\|(T^*T + \alpha I)x - T^*y^\delta\| \preceq \alpha^{1/2}\delta \tag{7}$$

to terminate the iteration and guarantee (5) with $\eta = \delta$; see also [1, 8, 11]. The method reaching this *implementable stopping rule* first is the *minimal residual method*, which will therefore outperform the conjugate gradient method here.

Tikhonov regularization (6) can be formulated equivalently in a *dual* form [3, 9]

$$(TT^* + \alpha I) z_\alpha^\delta = y^\delta, \qquad x_\alpha^\delta = T^* z_\alpha^\delta. \tag{8}$$

The energy error can now be estimated by

$$\|\|x - x_\alpha^\delta\|\| \le \|(TT^* + \alpha I) z_\alpha^\delta - y^\delta\|$$

yielding another implementable stopping rule

$$\|(TT^* + \alpha I) z - y^\delta\| \le \delta \tag{9}$$

for the iterative solution of (8) guaranteeing (5). Again, the minimal residual method applied to (8) will be the fastest method satisfying (9) under all methods working on the Krylov subspaces $\mathcal{K}_k(TT^* + \alpha I, y^\delta)$ for the dual formulation.

Summarizing, we observe that the minimal residual method is the optimal iterative method for solving Tikhonov regularization in its primal (2) or dual (8) form with implementable stopping rules. Since the minimal residual method and the conjugate gradient method can be implemented in one algorithm [6], one can actually utilize the minimal residual method for stopping the iteration, while employing the iterates of the conjugate gradient method as approximations for x_α^δ. Let us finally mention that the arguments also apply to preconditioned Krylov subspace methods [2], which can be considered in a similar manner.

References

[1] H. Egger, *Fast fully iterative Newton-type methods for inverse problems*, J. Inverse Ill-Posed Probl. **15** (2007), 257–275.

[2] H. Egger, *Preconditioning CGNE iteration for inverse problems*, Numer. Linear Algebra Appl. **14** (2007), 183–196.

[3] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.

[4] H. W. Engl, K. Kunisch, and A. Neubauer, *Convergence rates for Tikhonov regularisation of nonlinear ill-posed problems*, Inverse Problems **5** (1989), 523–540.

[5] A. Griesbaum, B. Kaltenbacher, and B. Vexler. *Efficient computation of the Tikhonov regularization parameter by goal-oriented adaptive discretization*, Inverse Problems **24** (2008), 025025.

[6] M. Hanke, *Conjugate gradient type methods for ill-posed problems*, Lonmgman Scientific, Harlow, 1995.

[7] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative regularization methods for nonlinear ill-posed problems*, Walter de Gruyter, Berlin, 2008.

[8] S. Langer and T. Hohage, *Convergence analysis of an inexact iteratively regularized Gauss-Newton method under general source conditions*, J. Inverse Ill-Posed Probl. **15** (2007), 311–327.

[9] A. Neubauer and J. T. King, *A variant of finite-dimensional Tikhonov regularization with a posteriori parameter choice*, Computing **40** (1988), 91–109.

[10] A. Rieder, *On the regularization of nonlinear ill-posed problems via inexact Newton iterations*, Inverse Problems 15 (1999), 309–327.

[11] O. Scherzer, *A modified Landweber iteration for solving parameter estimation problems*, Appl. Math. Optim. **38** (1998), 45–68.

Target Identification in Electrolocation: How Fishes Solve Inverse Problems

Thomas Boulier
(joint work with Habib Ammari, Josselin Garnier)

In turbid waters of South America and Africa, there can be found small fishes that are called *weakly electric*. In 1958, Lissmann and Machin explained the purpose of the electric field they emit [4]: it is a location mechanism. In other words, knowing the distorsion of the field induced by an object, these fishes are able to recognize this latter.

In the last decades, behavioral studies have shown that they are able to recognize the location, the shape, and the electrical parameters of any object located in their vincinity (for a review, see [5]). Hence, in a mathematical point of view, studying this ability - called *active electrolocation* - is a great opportunity for the understanding of inverse problems.

1. Mathematical Model

In $\mathbb{R}^2$, let us denote Ω the body of the fish, and $D \subset \mathbb{R}^d \setminus \overline{\Omega}$ an object to recognize.

Analysis of the electroreceptors have shown that the multi-frequency content of the measurements are very important for these fishes [5]. The typical wavelength is about 3km and the range of location does not exceed 1m. Thus, the *electro-quasistatic* (or EQS) approximation is best suited for this problem. That is, the electric field $\mathbf{E}$ in the Maxwell system is approximately irrotational, so that there exists an electric potential u such that $\nabla u = \mathbf{E}$ and

$$\nabla \cdot (\sigma + i\varepsilon\omega)\nabla u = f, \tag{1}$$

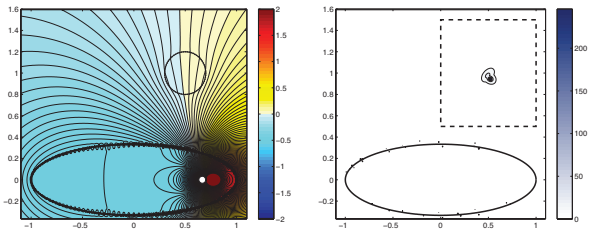
where σ and ε are respectively the conductivity and the permittivity, ω is the frequency, and f is the source of current. Electric parameters will be constant by part: σ_0, ε_0 in the water, σ_b, ε_b in the fish's body, σ_s, ε_s inside the skin, and σ_1, ε_1 in the anomaly D.

The equation (1) with a constant by part complex-valued conductivity function $k := \sigma + i\varepsilon\omega$ is simplified to the Laplace equation with jump relations across the surfaces of discontinuity for k. Thus, it only remains to write the boundary conditions across the skin, which is very thin and very resistive. From a multi-scale analysis of the equations in terms of layer potentials developed in [7, chap. 3], we have shown that on $\partial\Omega$, one has [1]

$$[u] - \xi \left.\frac{\partial u}{\partial \nu}\right|_+ = 0, \quad \text{and} \quad \left.\frac{\partial u}{\partial \nu}\right|_- = 0, \tag{2}$$

where ξ is called the *effective thickness* by Assad [2], who derived formally these equations during his Ph.D. thesis. It can be seen as the ratio between the small conductivity of the skin σ_s and its small thickness.

Finally, for the sake of unicity of the solution, behavior at infinity has to be made precise: the electric potential must goes to 0 at infinity.



(a) (b)

FIGURE 1. Numerical results. (a) Simulation of forward problem. (b) Plot of SF-MUSIC imaging functional.

2. Algorithm of Localization

In this section, an algorithm that locates the anomaly D will be presented. It uses multi-frequency measurements : if $\omega_1, \ldots, \omega_M$ are the frequencies of the emitted signal, and $x_1, \ldots, x_N$ are the locations of the receptors on the skin $\partial\Omega$, the following *space-frequency response matrix* can be built

$$\mathcal{M} = \left[\left(\frac{1}{2} I - \mathcal{K}_\Omega^* - \xi \frac{\partial \mathcal{D}_\Omega}{\partial \nu} \right) \left(\left. \frac{\partial u}{\partial \nu} \right|_{\partial\Omega^+} - \left. \frac{\partial U}{\partial \nu} \right|_{\partial\Omega^+} \right) (x_p, \omega_q) \right]_{1 \le p \le N,\, 1 \le q \le M}, \tag{3}$$

where U is the background electric field, $\mathcal{K}_\Omega^*$ and $\mathcal{D}_\Omega$ are respectively Neumann-Poincaré and double layer potential operators associated to Ω. In the case where $D = z + \delta B$, B being an open set of radius 1, $\operatorname{dist}(z, \partial\Omega) \gg 1$ and $\delta \ll 1$, a dipolar approximation of this measurements gives us [1]

$$\mathcal{M}_{pq} \approx \delta^2 \nabla U(z) \cdot M(k_q, B) \cdot \nabla_z \frac{\partial G}{\partial \nu_x}(z, x_p), \tag{4}$$

where G is the Green's function in $\mathbb{R}^2$ and $M(k_q, B)$ is the first-order polarization tensor (PT) associated to B with the ratio $k_q := (\sigma_1 + i\varepsilon_1\omega_q)/\sigma_0$.

From (4), we deduce that the columns of the matrix $\mathcal{M}$ are linear combination of partial derivatives of G. An algorithm called *Space-Frequency MUSIC* [6] uses this remark to image the anomaly (see Figure 1).

References

[1] H. Ammari, T. Boulier, and J. Garnier. Modeling active electrolocation in weakly electric fish. *to appear in SIAM Journal on Imaging Sciences*, 2012.

[2] C. Assad. *Electric field maps and boundary element simulations of electrolocation in weakly electric fish*. PhD thesis, California Institute of Technology, 1997.

[3] C. Darwin. *On the origins of species by means of natural selection*. John Murray, 1859.

[4] HW Lissmann and KE Machin. The mechanism of object location in gymnarchus niloticus and similar fish. *Journal of Experimental Biology*, 35(2):451–486, 1958.

[5] P. Moller. *Electric fishes: history and behavior*, volume 17. Chapman & Hall London, 1995.

[6] B. Scholz. Towards virtual electrical breast biopsy: space-frequency music for trans-admittance data. *Medical Imaging, IEEE Transactions on*, 21(6):588–595, 2002.

[7] H. Zribi. *La Méthode des Équations Intégrales pour des Analyses de Sensitivité*. PhD thesis, Ecole Polytechnique X, 2005.

Optimization-Based Sampling for Estimation and Uncertainty Quantification in Large-Scale Inverse Problems

Johnathan M. Bardsley

We begin with the linear, Gaussian statistical model

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{e}, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is the ill-conditioned forward model matrix, $\mathbf{b} \in \mathbb{R}^n$ is the data, $\mathbf{x} \in \mathbb{R}^n$ is the unknown image, and $\mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$ is the noise. In all instances, we assume that n is large.

The likelihood function for statistical model (1) has the form

$$p(\mathbf{b}|\mathbf{x}, \lambda) \propto \lambda^{n/2} \exp\left(-\frac{\lambda}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2\right), \tag{2}$$

where $\lambda = 1/\sigma^2$ is the noise precision and '$\propto$' denotes proportionality. If we additionally assume a Gaussian prior of the form

$$p(\mathbf{x}|\delta) \propto \delta^{n/2} \exp\left(-\frac{\delta}{2}\mathbf{x}^T\mathbf{L}\mathbf{x}\right), \tag{3}$$

where δ is the prior precision, then the conditional probability density for $\mathbf{x}$ is given by Bayes' Law:

$$\begin{aligned} p(\mathbf{x}|\mathbf{b}, \lambda, \delta) &\propto p(\mathbf{b}|\mathbf{x}, \lambda)p(\mathbf{x}|\delta) \\ &\propto \exp\left(-\frac{\lambda}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 - \frac{\delta}{2}\mathbf{x}^T\mathbf{L}\mathbf{x}\right). \end{aligned} \tag{4}$$

We can efficiently compute a sample $\mathbf{x}^*$ from (4) by first generating new random data $\hat{\mathbf{b}}$ from $N(\mathbf{b}, \lambda^{-1}\mathbf{I})$ and $\hat{\mathbf{c}}$ from $N(\mathbf{0}, \delta^{-1}\mathbf{I})$ and then by solving the least squares problem

$$\begin{aligned} \mathbf{x}^* &= \arg\min_{\mathbf{x}} \left\{ \left\| \begin{bmatrix} \lambda^{1/2}(\mathbf{A}\mathbf{x} - \hat{\mathbf{b}}) \\ \delta^{1/2}(\mathbf{L}\mathbf{x} - \hat{\mathbf{c}}) \end{bmatrix} \right\|^2 \right\} \\ &= \arg\min_{\mathbf{x}} \left\{ \frac{\lambda}{2}\|\mathbf{A}\mathbf{x} - \hat{\mathbf{b}}\|^2 + \frac{\delta}{2}\|\mathbf{L}^{1/2}\mathbf{x} - \hat{\mathbf{c}}\|^2 \right\}. \end{aligned} \tag{5}$$

Note that in this approach, we compute a sample $\mathbf{x}^*$ by first randomizing the "data" $\begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}$ by adding a noise realization from the correct distribution, and then optimizing to obtain the sample $\mathbf{x}^*$. Thus we call the approach, randomize then optimize (RTO). With a collection of samples in hand, we can both estimate the unknown $\mathbf{x}$ via, e.g., the sample mean, and quantify uncertainty via, e.g., the sample standard deviation, or we can create a movie from the samples in order to visualize the uncertainty in $\mathbf{x}$.

In certain instances, samples from (5) can be computed directly, e.g. in the case of deconvolution with periodic or Neumann boundary conditions. However in many cases, the optimization problem must be solved iteratively. This is the case, for example, for deconvolution with zero boundary conditions, computed

tomography (CT), and also for nonlinear models in which the linear mapping $\mathbf{x} \mapsto \mathbf{Ax}$ is replaced by a nonlinear mapping $\mathbf{x} \mapsto \mathbf{A}(\mathbf{x})$. For some of these cases see [1].

In many applications (e.g., astronomical imaging and positron emission tomography (PET)), the data $\mathbf{b}$ arises from a Poisson distribution. Thus, instead of (1), we assume

$$\mathbf{b} = \text{Poiss}(\mathbf{Ax} + \mathbf{g}), \tag{6}$$

where $\mathbf{g}$ is the $N \times 1$ vector of background counts and is assumed to be known. In this case, the probability density function for the data is given by

$$p(\mathbf{b}|\mathbf{x}) = \prod_{i=1}^{N} \frac{([\mathbf{Ax}]_i + g_i)^{b_i} \exp[-([\mathbf{Ax}]_i + g_i)]}{b_i!}, \quad \mathbf{x} \geq \mathbf{0}. \tag{7}$$

Assuming the prior $p(\mathbf{x}|\delta)$ defined by (3), the conditional probability density $p(\mathbf{x}|\mathbf{b}, \delta)$ then has the form

$$p(\mathbf{x}|\mathbf{b}, \delta) \propto \exp\left(-\sum_{i=1}^{n} \{[\mathbf{Ax}]_i + g_i - b_i \ln([\mathbf{Ax}]_i + g_i)\} + \frac{\delta}{2}\mathbf{x}^T\mathbf{Lx} \right). \tag{8}$$

The implementation of RTO in the Poisson noise case is very similar to the Gaussian case. To compute a sample $\mathbf{x}^*$ from $p(\mathbf{x}|\mathbf{b}, \delta)$, we first generate new data $\hat{\mathbf{b}}$ from Poiss($\mathbf{b}$) and $\hat{\mathbf{c}}$ from $N(\mathbf{0}, \delta^{-1}\mathbf{I})$, and then solve the optimization problem

$$\mathbf{x}^* = \arg\min_{\mathbf{x} \geq \mathbf{0}} \left\{ \sum_{i=1}^{n} \{[\mathbf{Ax}]_i + g_i - \hat{b}_i \ln([\mathbf{Ax}]_i + g_i)\} + \frac{\delta}{2}\|\mathbf{L}^{1/2}\mathbf{x} - \hat{\mathbf{c}}\|^2 \right\}. \tag{9}$$

Note the presence of the nonnegativity constraint. The main difficulty in this approach is that the optimization problem (9) is non-trivial. Moreover, it is not obvious that RTO samples are exact samples from $p(\mathbf{x}|\mathbf{b}, \delta)$.

In my talk, I will discuss the RTO approach for computing samples of $\mathbf{x}$ from $p(\mathbf{x}|\mathbf{b}, \delta)$. I will show results for each of the above mentioned applications: deconvolution with Dirichlet, Neumann, and periodic boundary conditions, in both the Gaussian and Poisson noise cases; tomography with Gaussian noise (CT) and with Poisson noise (PET); and the nonlinear inverse problem of electrical impedance tomography (EIT) with Gaussian noise. Moreover, I will show how, through hierarchical Bayesian modeling of λ and δ, a simple MCMC sampling scheme can be derived in which samples of $\mathbf{x}$, δ, and λ (in the Gaussian noise case) are obtained, making the separate selection of the regularization parameter unnecessary.

References

[1] Johnathan M. Bardsley, *MCMC-Based Image Reconstruction with Uncertainty Quantification*, SIAM Journal on Scientific Computing, Vol. 34, No. 3, 2012, pp. A1316–A1332.

Inverse Problems with Poisson data

Frank Werner
(joint work with Thorsten Hohage and Carolin Homann)

We consider inverse problems described by an operator equation

$$F(u) = g \tag{1}$$

where the observed data is given by a Poisson process G_t with intensity $tg^\dagger$. Here $g^\dagger \in \mathbf{L}^1(\mathbb{M})$ denotes the exact photon density and $t > 0$ is a scaling parameter which can often be interpreted as exposure time. Problems of this type have been studied in [1] only for very special cases of F. From a general point of view, a promising approach to reconstruct u from G_t is Tikhonov-type regularization

$$\hat{u}_\alpha = \operatorname{argmin}_{u \in \mathfrak{B}} \left[\mathcal{S}^{\text{direct}}(G_t; F(u)) + \alpha \mathcal{R}(u) \right] \tag{2}$$

where the data fidelity term $\mathcal{S}^{\text{direct}}$ is (up to a small shift $\sigma > 0$) the negative log-likelihood functional for Poisson data and $\mathcal{R}$ is a convex penalty including a priori information on u. This method has been investigated in [6] for the setup described above and convergence rates in expectation as t tends to ∞ have been proven.

Unfortunately, the problem (2) is non-convex for nonlinear F and thus the minimizer $\hat{u}_\alpha$ is difficult to determine. Thus we consider two different approaches which are both of the Newton-type form

$$\hat{u}_{n+1} = \operatorname{argmin}_{u \in \mathfrak{B}} \left[\mathcal{S}(G_t; T(u_n) + T'[u_n](u - u_n)) + \alpha_n \mathcal{R}(u) \right]. \tag{3}$$

Now the first approach results from using $T = F$ as in (1) and $\mathcal{S} = \mathcal{S}^{\text{direct}}$. This is referred to as the **direct approach** and has been studied in [4] where also convergence and convergence rates in expectation have been proven.

The second approach results from setting $\mathfrak{F}(u) := \ln(F(u) + \sigma)$ with F as in (1) and the shift σ and to apply (3) with $T = \mathfrak{F}$ and $\mathcal{S} = \mathcal{S}^{\exp}$ which denotes an approximation of the negative log-likelihood functional for $\ln(G_t + \sigma)$. This method is referred to as the **exponential approach**.

It turns out that the inner problems are convex both for the direct and the exponential approach, which is a substantial advantage over the method (2).

In this talk we presented convergence rates for both approaches, where the results for the direct approach are improvements of those from [4] and those for the exponential approach are new. For both methods one obtains under certain conditions and for some specific stopping index n_* the convergence rate

$$\mathbb{E}\left[\mathcal{D}_{\mathcal{R}}\left(\hat{u}_{n_*}; u^\dagger\right)\right] = \mathcal{O}\left(\varphi\left(\frac{1}{\sqrt{t}}\right)\right) \tag{4}$$

as $t \to \infty$ where $\mathbb{E}$ denotes the expectation, $\mathcal{D}_{\mathcal{R}}$ the Bregman distance w.r.t. $\mathcal{R}$ and φ measures the abstract smoothness of the unknown solution $u^\dagger$ in terms of a variational inequality (see e.g. [6]). Note that the obtained convergence rate (4) coincides for both the direct and the exponential approach, but is obtained under slightly different nonlinearity conditions. In fact, for the exponential approach the

nonlinearity condition is posed on $\mathfrak{F}$ and might be not fulfilled even if F in (1) is linear.

Moreover, this talk was concerned with the numerical realization of the inner problems in (3), which have to be solved in each iteration. In [4] we solved the inner problems for the direct approach via sequential quadratic programming, but it turned out that the minimizers $\hat{u}_n$ are difficult to calculate. This might be due to the fact that in this setting a side condition in the image space is posed. As this is not the case for the exponential approach, the inner problems are strictly convex and differentiable and can hence be solved almost exactly by Newton's method applied to the gradient. Therefore we expected the results for this method to be better. We also discussed the case of other inner solvers for the direct inner problems, where for example the algorithms proposed in [2, 3] apply in principle. Unfortunately straightforward implementations of those algorithms do not perform better than the sequential quadratic programming approach. design of other algorithms for the solution of the inner problems especially for large scale inverse problems are important problems for future research.

As a real world application we considered coherent X-ray imaging. The specific experimental setup is described in [4] and after some approximations it leads to the forward operator $F(\varphi) = |\mathcal{F}(\exp(\mathrm{i}\varphi))|^2$ where $\mathcal{F}$ denotes the 2D Fourier transform. As only the modulus of the Fourier transform is observed and the phase is unknown, such problems are called phase retrieval problems. This leads to non-uniqueness, which is overcome by some constraint on the support of φ (see e.g. [5]).

Both the direct and the exponential approach performed well for the phase retrieval problem with a slight numerical advantage for the exponential approach, although it turned out that the exponential approach needs more operator evaluations for the first Newton iterations. Therefore, a hybridization of both approaches seems to be promising.

References

[1] A. Antoniadis and J. Bigot. Poisson inverse problems. *Ann. Statist.*, 34(5):2132–2158, 2006.

[2] L. M. Briceño-Arias and P. Combettes. A Monotone + Skew Splitting Model for Composite Monotone Inclusions in Duality. *SIAM J. Optim.* (2011) 21(4): 1230–1250

[3] A. Chambolle and T. Pock. A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging. *J. Math. Imaging. Vis.* (2011) 40: 120–145

[4] T. Hohage and F. Werner. Iteratively regularized Newton-type methods for general data misfit functionals and applications to Poisson data. *Numer. Math.*, 2012, to appear, DOI: 10.1007/s00211-012-0499-z.

[5] M. V. Klibanov. On the recovery of a 2-D function from the modulus of its Fourier transform. *J. Math. Anal. Appl.*, 323(2):818–843, 2006.

[6] F. Werner and T. Hohage. Convergence rates in expectation for Tikhonov-type regularization of Inverse Problems with Poisson data. *Inverse Probl.*, 28(10):104004, 2012.

Parameter identification problems with non-Gaussian noise

Christian Clason

Consider the inverse problem $S(u) = y^\delta$ for a (possibly nonlinear) operator S between two Banach spaces X and Y and noisy data y^δ. One possible approach for computing an (approximate) solution to the inverse problem is minimizing the Tikhonov functional

$$\mathcal{F}(S(u), y^\delta) + \alpha\mathcal{R}(u)$$

for an appropriate discrepancy term $\mathcal{F}$ and regularization term $\mathcal{R}$. Just as the regularization term incorporates a priori information on the solution, the discrepancy term should be chosen based on a priori information on the noise. Here, the standard L^2 data fitting term is statistically motivated by the assumption of Gaussian noise. For non-Gaussian noise, however, other data fitting terms turn out to be more appropriate. For impulsive noise (appearing in digital image acquisition, e.g., as salt-and-pepper noise) L^1 fitting is more robust. Similarly, uniform noise (e.g., arising from quantization errors) has a statistical connection to L^∞ fitting. Both formulations lead to non-differentiable problems which are challenging to solve numerically.

This talk presents an approach that combines an iterative smoothing procedure with a semismooth Newton method. Specifically, the L^1 norm is replaced with a Huber norm

$$\|u\|_\beta := \int |u(x)|_\beta \, dx, \quad |t|_\beta = \begin{cases} t - \frac{\beta}{2} & t > \beta \\ -t - \frac{\beta}{2} & t < -\beta \\ \frac{1}{2\beta} t^2 & |t| \leq \beta \end{cases}$$

which has a semismooth Fréchet derivative. The L^∞ fitting problem has the equivalent formulation

$$\min_{u,c} c + \alpha\mathcal{R}(u) \quad \text{subject to} \quad \|S(u) - y^\delta\|_{L^\infty} \leq c,$$

for which the Moreau–Yosida smoothing

$$\min_{u,c} c + \alpha\mathcal{R}(u) + \frac{\gamma}{2}\left[\|\max(0, S(u) - y^\delta - c)\|_{L^2}^2 + \|\min(0, S(u) - y^\delta + c)\|_{L^2}^2\right]$$

is introduced. Again, this functional has a semismooth Fréchet derivative. In both cases, under a standard second order condition, the semismooth Newton method converges locally superlinearly for fixed smoothing parameter β or γ, and the family of minimizers of the smoothed problems converge (subsequentially) to a minimizer of the original Tikhonov functionals as $\beta \to 0$ or $\gamma \to \infty$. The semismooth Newton method is thus combined with a continuation strategy with respect to the smoothing parameter, which in practice has a globalizing effect.

The efficiency of this approach is illustrated for the inverse potential problem of recovering u from noisy measurements of $y = S(u)$ solving $-\Delta y + uy = f$.

References

[1] C. Clason, B. Jin *A semi-smooth Newton method for nonlinear parameter identification problems with impulsive noise*, SIAM Journal on Imaging Sciences **5** (2012), 505–536.

[2] C. Clason, L^∞ *fitting for inverse problems with uniform noise*, Inverse Problems **28** (2012), 104007.

Parameter choices for total variation regularization

Elena Resmerita

(joint work with Stefan Kindermann, Lawrence Mutimbu)

Total variation regularization has been a popular approach for denoising and deblurring problems. It aims at approximating true images $u^\dagger \in BV(\Omega)$ by minimizing functionals of the form (see the Rudin-Fatemi-Osher model in [7])

$$E(u) := ||Ku - g^\delta||^2_{L^2} + \alpha TV(u),$$

where $BV(\Omega)$ is the space of bounded variation functions, TV is the total variation seminorm, g^δ are the noisy data satisfying $\|g - g^\delta\| \leq \delta$ and $\alpha > 0$ is the regularization parameter. Here $u^\dagger$ is understood as a solution of

$$\min TV(u) \quad \text{subject to} \quad Ku = g,$$

where $K : BV(\Omega) \to L^2(\Omega)$ is a linear (identity or convolution) operator. Classical choices for the regularization parameter are of a priori type, i.e., $\alpha = \alpha(\delta)$ and of a posteriori type: $\alpha = \alpha(\delta, g^\delta)$. Unfortunately, knowledge on the noise level δ is not always available. This is why parameter choice rules depending only on the noisy data g^δ, which would guarantee convergence of the minimizers u^δ_α to $u^\dagger$, would be of interest. However, such rules (called heuristic rules) cannot yield convergence in the worst case, that is,

$$\lim_\delta u^\delta_\alpha \to u \quad \text{for all } g^\delta : \|g^\delta - g\|_{L^2} \leq \delta.$$

Convergence results for heuristic rules in case of quadratic regularization in Hilbert spaces were established recently in a restricted noise case (see, e.g., [5], [4]), i.e.,

$$\lim_\delta u^\delta_\alpha \to u \quad \text{for all } g^\delta : \|g^\delta - g\|_{L^2} \leq \delta \text{ and } g^\delta - g \in \mathcal{N},$$

for some suitable set $\mathcal{N}$. The investigated rules look for a parameter α^* depending on g^δ, defined by

$$\alpha^* = argmin_{\alpha \in M} \phi(\alpha, g^\delta),$$

where M is an appropriate interval or a discrete set $M \subset \mathbb{R}_+$. They have been inspired by the seminal works [1], [2]. The paper [3] extends a couple of such rules to non-quadratic regularization in Banach spaces, addressing q-convex penalties or the ℓ^1-penalty, which do not cover the total variation case.

Our work conducts a numerical study of several noiselevel-free regularization parameter choice rules for total variation regularization of linear inverse problems. First, we review convergence results for total variation regularization when a priori and a posteriori rules are employed and point out that convergence with respect

to the strict metric d is the best that one can obtain in BV (see [6]), where d is defined on $BV(\Omega) \times BV(\Omega)$ by

$$d(u, v) := \|u - v\|_{L^1} + |TV(u) - TV(v)|.$$

Second, we propose some generalizations of two well-known heuristic parameter choice rules, the quasioptimality principle and the Hanke-Raus rules. We investigate the feasibility of these rules in one and two dimensions by numerical simulation and conclude that the discrete quasioptimality principle using the strict metric or parts of it and a Hanke-Raus rule perform well for the investigated denoising and deblurring problems.

Determining an appropriate set $\mathcal{N}$ for noise restriction in order to ensure convergence of the total variation regularization combined with the proposed rules remains an open problem.

References

[1] V. B. Glasko and Yu. A. Kriskin, *On the quasi-optimality principle for ill-posed problems in Hilbert space*, USSR Comp. Math. Math. Phys., **24** (1984), 1–7.

[2] M. Hanke and T. Raus, *A general heuristic for choosing the regularization parameter in ill-posed problems*, SIAM J. Sci. Comp., **17** (1996), 956–972.

[3] B. Jin and D. A. Lorenz, *Heuristic parameter-choice rules for convex variational regularization based on error estimates*, SIAM J. Num. Anal., **48** (2010), 1208–1229.

[4] S. Kindermann and A. Neubauer, *On the convergence of the quasiopti- mality criterion for (iterated) Tikhonov regularization*, Inverse Probl. Imaging, **2** (2008), 291–299.

[5] A. Neubauer, *The convergence of a new heuristic parameter selection criterion for general regularization methods*, Inverse Problems **24** (2008), pp. ID 055005, 10p.

[6] C. Poeschl, E. Resmerita and O. Scherzer, Discretization of variational regularization in Banach spaces, *Inverse Problems*, 26 (2010) 105017

[7] L.I. Rudin, S. Osher and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica **D 60** (1992), 259–268.

Total generalized variation and applications to inverse problems in medical imaging

Kristian Bredies

(joint work with Martin Holler, Florian Knoll, Karl Kunisch, Michael Pienn, Thomas Pock, Rudolf Stollberger and Tuomo Valkonen)

We introduce and study the concept of total generalized variation (TGV) of order k for symmetric tensor fields of order l: For $u \in L^1_{\mathrm{loc}}(\Omega, \mathrm{Sym}^l(\mathbf{R}^d))$, let

$$\mathrm{TGV}^{k,l}_\alpha(u) = \sup \left\{ \int_\Omega u \cdot \operatorname{div} v \, \mathrm{d}x \;\middle|\; v \in \mathcal{C}^k_0(\Omega, \mathrm{Sym}^{k+l}(\mathbf{R}^d)), \; \|\operatorname{div}^i v\|_\infty \le \alpha_i, \; i = 0, \dots, k-1 \right\}.$$

The scalar case, proposed in [1], already generalizes the notion of total variation, while $l \ge 1$ generalizes the notion of total deformation studied in [2]. The functional is in particular able to incorporate edge as well as higher-order smoothness

information [3]. The associated Banach spaces are shown to coincide for any order k with the space of symmetric tensor fields of bounded deformation in terms of the strong topology (see [4, 3] for the case $k = 2$). The functional-analytic properties of the latter allow to prove well-posedness of Tikhonov regularization for inverse problems $Ku = f$ where K is linear and bounded, i.e., existence and stability of minimizers of

$$\min_{u \in L^p(\Omega, \mathrm{Sym}^l(\mathbf{R}^d))} \frac{\|Ku - f\|^2}{2} + \mathrm{TGV}_\alpha^{k,l}(u).$$

We also study strict TGV-topologies reflecting the convergence notion $\|u^n - u\|_1 \to 0$ and $\mathrm{TGV}_\alpha^{k,l}(u^n) \to \mathrm{TGV}_\alpha^{k,l}(u)$ which turn out not to be equivalent to the strict TV-topology.

Furthermore, computational methods for the minimization of TGV-regularized optimization problems are presented. They base on rewriting the suitably discretized objective functional as a convex-concave saddle point problem and applying the primal-dual iteration presented in [5]. On the one hand, this results in simple, efficient and convergent methods and, on the other hand, in flexibility, for instance, with respect to the choice of the discrepancy terms (for an application outside inverse problems, see [6]). The TGV-regularization approach as well as the proposed algorithms are applied to medical imaging problems such as undersampled magnetic resonance imaging (MRI) [7], reconstruction of noisy diffusion tensor imaging data (DTI) [3] and denoising of dual energy computed tomography (dual-energy CT) images [8].

Finally, numerical experiments confirm the high reconstruction quality as well as the efficiency of TGV-based methods, see, for instance, Figures 1 and 2.

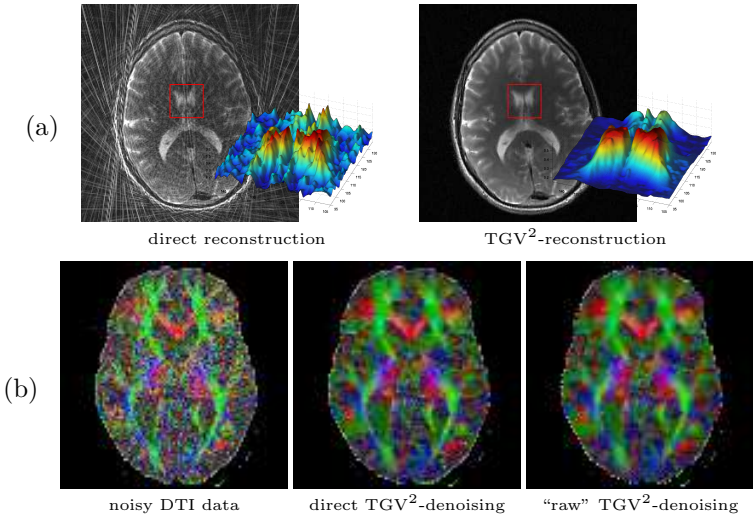


FIGURE 1. Numerical examples for TGV^2-regularization: (a) Undersampling MRI, (b) Diffusion tensor imaging (DTI).

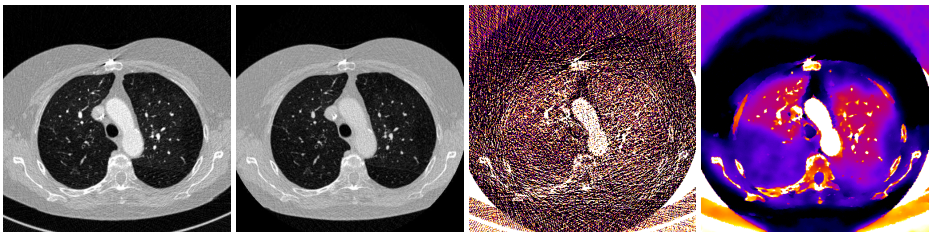


image A_0 image B_0 difference image $A_0 - B_0$ TGV^2-denoising result

FIGURE 2. TGV^2-regularization for dual-energy CT denoising.

REFERENCES

[1] K. Bredies, K. Kunisch and T. Pock, *Total generalized variation*, SIAM J. Imaging Sci., **3(3)** (2010), 492–526.

[2] K. Bredies, *Symmetric tensor fields of bounded deformation*, Ann. Mat. Pur. Appl. (2012), DOI 10.1007/s10231-011-0248-4.

[3] T. Valkonen, K. Bredies and F. Knoll, *Total generalised variation in diffusion tensor imaging*, SFB MOBIS Report 2012-003 (2012), University of Graz.

[4] K. Bredies and T. Valkonen, *Inverse problems with second-order total generalized variation constraints*, Proceedings of SampTA 2011 (2011), Singapore.

[5] A. Chambolle and T. Pock, *A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging*, J. Math. Imaging Vis. **40(1)** (2011), 120–145.

[6] K. Bredies and M. Holler, *Artifact-free JPEG decompression with total generalized variation*, Proceedings of VISAPP 2012 (2012), Rome.

[7] F. Knoll, K. Bredies, T. Pock and R. Stollberger, *Second order total generalized variation (TGV) for MRI*, Magnet. Reson. Med. **65(2)** (2011), 480–491.

[8] M. Pienn, T. R. C. Johnson, P. Kullnig, R. Stollberger, G. Kovacs, M. Tscherner, A. Olschewski, H. Olschewski and Z. Bálint, *Cardiac output determination by dynamic contrast-enhanced CT*, J. Thorac. Imag. **27** (2012), W115-W163.

One Shot Inverse Scattering

MARTIN HANKE

Given far field data of a two-dimensional time-harmonic scattered wave, reflected by an ensemble of well-separated acoustic or electromagnetic scatterers, we discuss an algorithm to approximate the far field data radiated by each of these scatterers separately. The method is based on a Galerkin procedure considering subspaces spanned by the singular vectors of 'restricted' far field operators that map local source distributions to the corresponding radiated far field patterns. Furthermore, we employ a windowed Fourier transform of the given far field to extract the required a priori knowledge directly from the data. Finally, we use the concept of scattering supports to compute (approximate) reconstructions of the scatterers from the separated far field components.

This is joint work with Roland Griesmaier (Leipzig), Thorsten Raasch (Mainz), and John Sylvester (Seattle).

References

[1] R. Griesmaier, M. Hanke, and T. Raasch, *Inverse source problems for the Helmholtz equation and the windowed Fourier transform*. SIAM J. Sci. Comput. **34** (2012), A1544-A1562.
[2] R. Griesmaier, M. Hanke, and J. Sylvester, *Far field splitting for the Helmholtz equation*. submitted (2012).
[3] S. Kusiak and J. Sylvester, *The scattering support*. Comm. Pure Appl. Math. **56** (2003), 1525–1548.
[4] J. Sylvester, *Notions of support for far fields*. Inverse Problems **22** (2006), 1273–1288.

Reporter: Carolin Homann

Participants

Dr. Ricardo J. Alonso
Department of Mathematics
Rice University
P.O. Box 1892
Houston, TX 77005-1892
UNITED STATES

Prof. Dr. Johnathan M. Bardsley
Department of Mathematical Sciences
University of Montana
Missoula, MT 59812-1032
UNITED STATES

Saskia Becker
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstr. 39
10117 Berlin
GERMANY

Prof. Dr. Marc Bonnet
ENSTA/UMA
32, Boulevard Victor
75739 Paris Cedex 15
FRANCE

Prof. Dr. Liliana Borcea
Department of Mathematics
Rice University
6100 Main Street
Houston, TX 77005-1892
UNITED STATES

Thomas Boulier
Dept. de Mathématiques et Applications
École Normale Superieure
45, rue d'Ulm
75005 Paris Cedex
FRANCE

Dr. Kristian Bredies
Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstr. 36
8010 Graz
AUSTRIA

Prof. Dr. Fioralba Cakoni
Department of Mathematical Sciences
University of Delaware
501 Ewing Hall
Newark, DE 19716-2553
UNITED STATES

Prof. Dr. Laurent Cavalier
Centre de Mathématiques et
d'Informatique
Université de Provence
39, Rue Joliot-Curie
13453 Marseille Cedex 13
FRANCE

Dr. Christian Clason
Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstr. 36
8010 Graz
AUSTRIA

Prof. Dr. Maarten de Hoop
Department of Mathematics
Purdue University
West Lafayette, IN 47907-1395
UNITED STATES

Prof. Dr. Laurent Demanet
Department of Mathematics
Massachusetts Institute of
Technology
77 Massachusetts Avenue
Cambridge, MA 02139-4307
UNITED STATES

Prof. Dr. Christine De Mol
Department of Mathematics
Université Libre de Bruxelles
CP 217 Campus Plaine
Bd. du Triomphe
1050 Bruxelles
BELGIUM

Dr. Vladimir Druskin
Schlumberger Doll Research Center
One Hampshire St.
Cambridge, MA 02139-1578
UNITED STATES

Dr. Fabian Dunker
Institut für Numerische
und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16-18
37083 Göttingen
GERMANY

Prof. Dr. Herbert Egger
Zentrum Mathematik
TU München
Boltzmannstr. 3
85748 Garching b. München
GERMANY

Prof. Dr. Josselin Garnier
Laboratoire de Probabilites et
Modeles Aleatoires
Université Paris VII
175 rue du Chevaleret
75013 Paris Cedex
FRANCE

Prof. Omar Ghattas
Institute for Computational
Engineering and Sciences (ICES)
University of Texas at Austin
1 University Station C 0200
Austin, TX 78712-1085
UNITED STATES

Laure Giovangigli
Dept. de Mathématiques et Applications
École Normale Superieure
45, rue d'Ulm
75005 Paris Cedex
FRANCE

PD. Dr. Markus Grasmair
Fakultät für Mathematik
Universität Wien
Nordbergstr. 15
1090 Wien
AUSTRIA

Dr. Fernando Guevara Vasquez
Department of Mathematics
University of Utah
155 South 1400 East
Salt Lake City, UT 84112-0090
UNITED STATES

Bernadette Hahn
FR 6.1 - Mathematik
Universität des Saarlandes
Postfach 15 11 50
66041 Saarbrücken
GERMANY

Prof. Dr. Martin Hanke-Bourgeois
FB Mathematik/Physik/Informatik
Mathematisches Institut
Johannes-Gutenberg-Universität
55099 Mainz
GERMANY

Prof. Dr. Bernd Hofmann
Fakultät für Mathematik
TU Chemnitz
Reichenhainer Str. 41
09126 Chemnitz
GERMANY

Prof. Dr. Thorsten Hohage
Institut für Numerische
und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16-18
37083 Göttingen
GERMANY

Carolin Homann
Institut für Numerische
und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16-18
37083 Göttingen
GERMANY

Prof. Dr. Barbara Kaltenbacher
Institut für Mathematik
Universität Alpen-Adria
Universitätsstr. 65-67
9020 Klagenfurt
AUSTRIA

PD. Dr. Stefan Kindermann
Institut für Industriemathematik
Universität Linz
Altenbergerstr. 69
4040 Linz
AUSTRIA

Prof. Dr. Antonio Leitao
Departamento de Matematica
Universidade Federal de Santa Catarina
Florianopolis 88040-900
BRAZIL

Prof. Dr. Dirk Alfred Lorenz
Institut für Algebra u. Analysis
TU Braunschweig
Pockelsstr. 14
38106 Braunschweig
GERMANY

Prof. Dr. Russell Luke
Institut für Numerische
und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16-18
37083 Göttingen
GERMANY

Prof. Dr. Peter Maaß
Universität Bremen
FB 3 Mathematik & Informatik
28344 Bremen
GERMANY

Dr. Alexander Mamonov
Institute for Computational
Engineering and Sciences (ICES)
University of Texas at Austin
1 University Station C
Austin, TX 78712-1085
UNITED STATES

Prof. Dr. Peter Mathe
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstr. 39
10117 Berlin
GERMANY

Prof. Dr. Shari Moskow
Department of Mathematics
Drexel University
Korman Center 269
3141 Chestnut St.
Philadelphia, PA 19104
UNITED STATES

Valeriya Naumova
Johann Radon Institute for Comput.
and Applied Mathematics (RICAM)
Austrian Academy of Sciences
Altenbergerstr. 69
4040 Linz
AUSTRIA

Prof. Dr. Sergei Pereverzyev
Johann Radon Institute
Austrian Academy of Sciences
Altenberger Straße 69
4040 Linz
AUSTRIA

Dr. Christiane Pöschl
Computational Science Center
Universität Wien
Nordbergstr. 15
1090 Wien
AUSTRIA

Prof. Dr. Roland Potthast
Deutscher Wetterdienst
Research and Development
Head Section FE12
Frankfurter Straße 135
63067 Offenbach
GERMANY

Prof. Dr. Kui Ren
Department of Mathematics
The University of Texas at Austin
1 University Station C1200
Austin, TX 78712-1082
UNITED STATES

Dr. Elena Resmerita
Institut für Mathematik
Universität Klagenfurt
Universitätsstr. 65-67
9020 Klagenfurt
AUSTRIA

Prof. Dr. Andreas Rieder
Karlsruher Institut f. Technologie (KIT)
Inst. f. Angew. & Numerische
Mathematik
Englerstr. 2
76131 Karlsruhe
GERMANY

Prof. Dr. Otmar Scherzer
Computational Science Center
Universität Wien
Nordbergstr. 15
1090 Wien
AUSTRIA

Dr. Johannes Schmidt-Hieber
École Nationale de la Statistique
e de l'Adm. Economique
ENSAE
3, avenue Pierre-Larousse
92245 Malakoff
FRANCE

Prof. Dr. John C. Schotland
Department of Mathematics
University of Michigan
East Hall, 525 E. University
Ann Arbor, MI 48109-1109
UNITED STATES

Laurent Seppecher
Dept. de Mathématiques et Applications
École Normale Superieure
45, rue d'Ulm
75005 Paris Cedex
FRANCE

Prof. Dr. Samuli Siltanen
Dept. of Mathematics & Statistics
University of Helsinki
P.O.Box 68
00014 University of Helsinki
FINLAND

Dr. Frank Werner
Institut für Numerische
und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16-18
37083 Göttingen
GERMANY

Prof. Dr. Hong-Kai Zhao
Department of Mathematics
University of California, Irvine
Irvine, CA 92697-3875
UNITED STATES