Convergence rates in expectation for Tikhonov-type regularization of Inverse Problems with Poisson data

Frank Werner

Institute for Numerical and Applied Mathematics

University of Göttingen, Germany

joint work with Thorsten Hohage

CSR 2012

Outline

Outline

Problem setup

Data: Total number N and positions $x_i \in \mathbb{M}$ of photons distributed due to a unknown photon density $g^\dagger \in \mathbf{L}^1(\mathbb{M})$.

Task: Determine the reason $u^\dagger$ of the photon density $g^\dagger$.

Note: The total number N of counted photons depends on the intensity of $g^\dagger$ as well as a parameter t interpreted as exposure time.

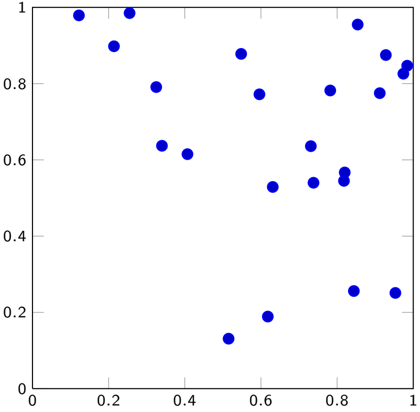
Poisson Processes

Mathematical model:

Poisson Process with intensity $tg^\dagger$, i.e.

$$\tilde{G}_t = \sum_{i=1}^{N} \delta_{x_i}$$

with the following properties:



Poisson process - Axiom I

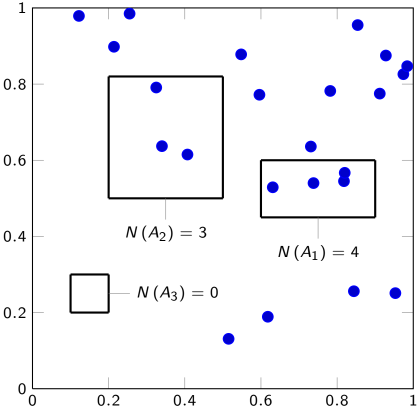
$$N(A) := \#\left\{ i \in \{1, \ldots, N\} \;\middle|\; x_i \in A \right\}$$

Independence:

For any choice of $A_1, \ldots, A_n \subset \mathbb{M}$ disjoint and measurable, the random variables

$$N(A_1), \ldots, N(A_n)$$

are independent.



Poisson process - Axiom II

$$N(A) := \#\left\{ i \in \{1, ..., N\} \;\middle|\; x_i \in A \right\}$$

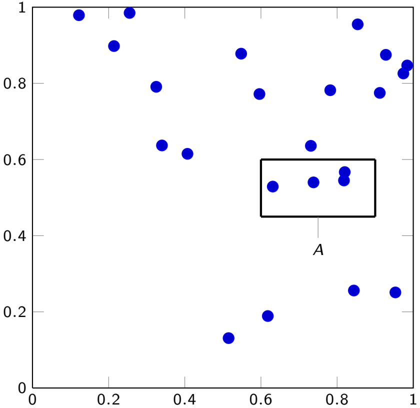
Poisson distribution:

For any measurable $A \subset \mathbb{M}$ the random variable

$$N(A)$$

is Poisson distributed with parameter

$$t \int_A g^\dagger \, \mathrm{d}x.$$



discretization / binning

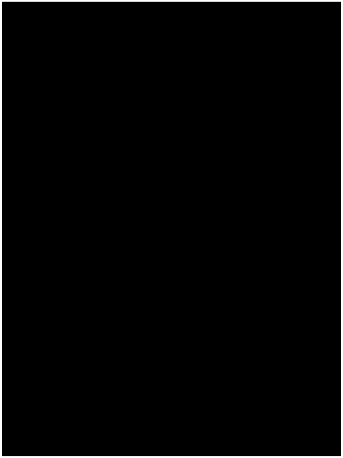
- $\mathbb{M} = \bigcup_{j=1}^{J} \mathbb{M}_j$, each $\mathbb{M}_j$ corresponds to one detector
- $g^\dagger \in \mathbf{L}^1(\mathbb{M}) \;\rightsquigarrow\; S_J : \mathbf{L}^1(\mathbb{M}) \to \mathbb{R}^J$ defined by

$$(S_J g)_j := \int_{\mathbb{M}_j} g \,\mathrm{d}x \qquad \text{and} \qquad S_J^* \underline{g} := \sum_{j=1}^{J} |\mathbb{M}_j|^{-1} \underline{g}_j, \qquad j = 1, ..., J$$

- $P_J := S_J^* S_J$ is the $\mathbf{L}^2$-orthogonal projection onto the subspace of functions, which are constant on each $\mathbb{M}_j$.
- natural extension to measures via $(S_J(\tilde{G}_t))_j = \tilde{G}_t(\mathbb{M}_j) = N(\mathbb{M}_j)$.

$\rightsquigarrow$ measured data: $\underline{g}_j^{\mathrm{obs}} = N(\mathbb{M}_j)\,, \qquad j = 1, \ldots, J$

Influence of t



We expect 20.000 photons per second

Difficulties

Model assumption: The imaging process can be described by an operator equation

$$F\left(u^\dagger\right) = g^\dagger$$

where $F : \mathfrak{B} \subset \mathcal{X} \to \mathcal{Y} \subset \mathbf{L}^1(\mathbb{M})$ is in general nonlinear and $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces.

The exact right-hand side $g^\dagger$ is unknown and in general F^{-1} is not continuous.

$\Rightarrow$ direct reconstruction impossible, regularization necessary!

Difficulties (cont.)

Several applications yield only data for small t, i.e.

- positron emission tomography (radiation exposure)
- astronomical imaging (limited observation time, motion artifacts)
- fluorescence microscopy (photobleaching)

$\Rightarrow$ use a negative log-likelihood aproach to use the infomation at hand on the Poisson distribution:

Minimize

$$u \mapsto \mathcal{S}\left(\tilde{G}_t; F(u)\right) := -\ln\left(\mathbf{P}\left[\tilde{G}_t \;\middle|\; \text{the exact photon density is } F(u)\right]\right)$$

over all admissible u.

Approach

Tackle the problem with **Tikhonov**-type regularization:

$$u_\alpha \in \underset{u \in \mathfrak{B}}{\operatorname{argmin}} \left[\mathcal{S}\left(\tilde{G}_t; F(u)\right) + \alpha \mathcal{R}(u) \right]$$

where $\mathcal{R}$ is a convex penalty term and $\alpha > 0$ a regularization parameter.

- $\mathcal{R}$ allows to incorporate a priori information on $u^\dagger$
- $\mathcal{R}$ stabilizes the reconstruction procedure
- u_α can be interpreted as a MAP estimator if $C \exp(-\alpha \mathcal{R}(u))$ models the prior

Outline

Data fidelity terms

Scaled data $G_t = \frac{1}{t} \sum_{i=1}^{N} \delta_{x_i}$, $tG_t = \tilde{G}_t$ Poisson process.

- Negative log-likelihood:

$$\mathcal{S}\left(G_t; g\right) = \int_{\mathbb{M}} g \,\mathrm{d}x - \int_{\mathbb{M}} \ln\left(g\right) \,\mathrm{d}G_t, \qquad g \geq 0 \text{ a.e.}$$

- It holds $\mathbf{E}\left[\mathcal{S}\left(G_t; g\right)\right] = \int_{\mathbb{M}} \left[g - g^\dagger \ln\left(g\right)\right] \,\mathrm{d}x$
- $\rightsquigarrow$ ideal data misfit functional for exact data $g^\dagger$ given by

$$\mathbf{E}\left[\mathcal{S}\left(G_t; g\right)\right] - \mathbf{E}\left[\mathcal{S}\left(G_t; g^\dagger\right)\right] = \int_{\mathbb{M}} \left[g - g^\dagger - g^\dagger \ln\left(\frac{g}{g^\dagger}\right)\right] \,\mathrm{d}x$$

  which is the Kullback-Leibler divergence $\mathbb{KL}\left(g^\dagger; g\right)$.
- Error at g:

$$\left|\mathcal{S}\left(G_t; g\right) - \mathbf{E}\left[\mathcal{S}\left(G_t; g^\dagger\right)\right] - \mathbb{KL}\left(g^\dagger; g\right)\right| = \left|\int_{\mathbb{M}} \ln(g) \left(\mathrm{d}G_t - g^\dagger \,\mathrm{d}x\right)\right|.$$

Controlling the error

- we want to control the error

$$\mathbf{err}(g) := \left| \mathcal{S}(G_t; g) - \mathbf{E}\left[\mathcal{S}\left(G_t; g^\dagger\right)\right] - \mathbb{KL}\left(g^\dagger; g\right)\right|$$

with $g = F(u)$

- therefore we need to control the integrals

$$\int_{\mathbb{M}} \mathfrak{g} \left(\mathrm{d}G_t - g^\dagger \,\mathrm{d}x\right)$$

where $\mathfrak{g} = \ln(F(u))$

$\rightsquigarrow$ uniform concentration inequalities!

- well-studied for white noise (e.g. Gaussian), less known for Poisson processes

Uniform concentration inequalities for Poisson processes

Uniform concentration inequality (Reynaud-Bouret 2003)

- $\{f_a\}_{a\in A}$ countable family of functions with values in $[-b, b]$
- $Z := \sup_{a\in A} \left|\int_{\mathbb{M}} f_a(x)\left(\mathrm{d}G_t - g^\dagger \mathrm{d}x\right)\right|$
- $v_0 := \sup_{a\in A} \int_{\mathbb{M}} f_a^2(x)\, g^\dagger\, \mathrm{d}x$

Then for all $\rho, \varepsilon > 0$ it holds

$$\mathbf{P}\left[Z \geq (1+\varepsilon)\,\mathbf{E}[Z] + \frac{\sqrt{12 v_0 \rho}}{\sqrt{t}} + \left(\frac{5}{4} + \frac{32}{\varepsilon}\right)\frac{b\rho}{t}\right] \leq \exp(-\rho).$$

P. Reynaud-Bouret.
Adaptive estimation of the intensity of inhomogeneous Poisson processes via concentration inequalities.
Probab. Theory Rel., 126(1):103–153, 2003.

$\rightsquigarrow$ analogue to Talagrand's concentration inequalities for empirical processes!

Uniform concentration inequalities for Poisson processes

$$\mathbf{P}\left[Z \geq (1+\varepsilon)\,\mathbf{E}\left[Z\right] + \frac{\sqrt{12 v_0 \rho}}{\sqrt{t}} + \left(\frac{5}{4} + \frac{32}{\varepsilon}\right)\frac{b\rho}{t}\right] \leq \exp\left(-\rho\right).$$

- suppose $\mathbb{M} \subset \mathbb{R}^d$ bounded and Lipschitz
- $\{f_a\}$ dense subset of $B_s\left(R\right) := \left\{\mathfrak{g} \in H^s\left(\mathbb{M}\right) \;\middle|\; \left\|\mathfrak{g}\right\|_{H^s(\mathbb{M})} \leq R\right\}$ with $s > \frac{d}{2}$ (Sobolev's embedding theorem)
- easy: $v_0 \leq R^2 C \left\|g^\dagger\right\|_{\mathbf{L}^1(\mathbb{M})}$, $C > 0$ depending only on $\mathbb{M}$ and s
- more difficult (uses periodization and Fourier expansion):

$$\mathbf{E}\left[Z\right] \leq \frac{CR}{\sqrt{t}} \|g^\dagger\|_{\mathbf{L}^1(\mathbb{M})}$$

with a constant $C > 0$ depending only on $\mathbb{M}$ and s.

Controlling the error (cont.)

Uniform concentration inequality (W., Hohage 2012)

- $\mathbb{M} \subset \mathbb{R}^d$ bounded and Lipschitz,
- $s > d/2, R > 1$.

Then there exists $C_{\mathrm{conc}} = C_{\mathrm{conc}}\left(\mathbb{M}, s, g^\dagger\right) \geq 1$ such that

$$\mathbf{P}\left[\sup_{\mathfrak{g} \in B_s(R)} \left| \int_{\mathbb{M}} \mathfrak{g} \left(\mathrm{d}G_t - g^\dagger \,\mathrm{d}x \right) \right| \leq \frac{\rho}{\sqrt{t}} \right] \geq 1 - \exp\left(-\frac{\rho}{RC_{\mathrm{conc}}} \right)$$

for all $t \geq 1$ and $\rho \geq RC_{\mathrm{conc}}$.

Controlling the error (cont.)

- Concentration inequality requires $\mathfrak{g} \in H^s(\mathbb{M}) \subset \mathbf{L}^\infty(\mathbb{M})$ due to $s > d/2$
- Error at $g = F(u)$ leads to $\mathfrak{g} = \ln(F(u))$

$\Rightarrow$ Too strong assumption!

$\rightsquigarrow$ Shift by $\sigma > 0$:

$$\mathcal{S}(G_t; g) := \int_{\mathbb{M}} g \, \mathrm{d}x - \int_{\mathbb{M}} \ln(g + \sigma)\,(\mathrm{d}G_t + \sigma \mathrm{d}x)$$

$$\mathcal{T}\left(g^\dagger; g\right) := \mathbb{KL}\left(g^\dagger + \sigma; g + \sigma\right)$$

- Then the error is given by

$$\left| \int_{\mathbb{M}} \ln(g + \sigma)\left(\mathrm{d}G_t - g^\dagger \, \mathrm{d}x\right) \right|.$$

Controlling the error (cont.)

Corollary (final concentration inequality)

- $\mathbb{M} \subset \mathbb{R}^d$ bounded and Lipschitz
- $F(u) \geq 0$ a.e. for all $u \in \mathfrak{B}$
- there exists a Sobolev index $s > \frac{d}{2}$ such that

$$R := \sup_{u \in \mathfrak{B}} \|F(u)\|_{H^s(\mathbb{M})} < \infty$$

Then there exists $C_{\mathrm{conc}} = C_{\mathrm{conc}}\left(\mathbb{M}, s, g^\dagger\right) \geq 1$ such that

$$\mathbf{P}\left[\sup_{u \in \mathfrak{B}} \mathbf{err}\left(F(u)\right) \leq \frac{\rho}{\sqrt{t}}\right] \geq 1 - \exp\left(-\frac{\rho}{R \max\left\{\sigma^{-\lfloor s \rfloor - 1}, |\ln(R)|\right\} C_{\mathrm{conc}}}\right)$$

for all $t \geq 1, \rho \geq R \max\left\{\sigma^{-\lfloor s \rfloor - 1}, |\ln(R)|\right\} C_{\mathrm{conc}}$.

Proof relies on composition theorems in the Sobolev space $H^s(\mathbb{M})$.

Outline

Deterministic noise level

We have two data fidelity terms:

- $\mathcal{S}$ w.r.t. the measured data g^{obs}
- $\mathcal{T}$ w.r.t. the photon density $g^{\dagger}$

As before: consider the difference between both as noise level!

Noise level

There exist constants **err** ≥ 0 and $C_{\mathrm{err}} \geq 1$ such that

$$\mathcal{S}\left(g^{\mathrm{obs}}; g\right) - \mathcal{S}\left(g^{\mathrm{obs}}; g^{\dagger}\right) \geq \frac{1}{C_{\mathrm{err}}} \mathcal{T}\left(g^{\dagger}; g\right) - \mathbf{err}$$

for all $g \in F(\mathfrak{B})$.

Deterministic noise level (cont.)

- *Classical deterministic noise model:*
  If $\mathcal{S}(g;\hat{g}) = \mathcal{T}(g;\hat{g}) = \|g-\hat{g}\|_{\mathcal{Y}}^{r}$, then $C_{\mathrm{err}} = 2^{r-1}$ and
  $\textbf{err} = 2\left\|g^{\dagger} - g^{\mathrm{obs}}\right\|_{\mathcal{Y}}^{r}$.
- *Poisson data:*
  $C_{\mathrm{err}} = 1$ and

$$\textbf{err} \geq -\int_{\mathbb{M}} \ln\left(g^{\dagger}+\sigma\right)\left(\mathrm{d}G_t - g^{\dagger}\,\mathrm{d}x\right) + \int_{\mathbb{M}} \ln\left(F(u)+\sigma\right)\left(\mathrm{d}G_t - g^{\dagger}\,\mathrm{d}x\right)$$

  for all $u \in \mathfrak{B}$.
  Uniform concentration inequality: $\textbf{err} \leq \frac{2\rho}{\sqrt{t}}$ with probability $\geq 1 - \exp(-c\rho)$ for some constant $c > 0$.

Source condition

- Bregman distance:

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right) := \mathcal{R}\left(u\right) - \mathcal{R}\left(u^\dagger\right) - \left\langle u^*, u - u^\dagger \right\rangle$$

where $u^* \in \partial \mathcal{R}\left(u^\dagger\right) \subset \mathcal{X}'$.

- Use a variational inequality as source condition:

$$\beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right) \leq \mathcal{R}\left(u\right) - \mathcal{R}\left(u^\dagger\right) + \varphi\left(\mathcal{T}\left(g^\dagger ; F\left(u\right)\right)\right)$$

for all $u \in \mathfrak{B}$ with $\beta > 0$. φ is assumed to fulfill

  - $\varphi\left(0\right) = 0$,
  - $\varphi \nearrow$,
  - φ concave.

Source condition (cont.)

$$\beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right) \leq \mathcal{R}(u) - \mathcal{R}\left(u^\dagger\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F(u)\right)\right)$$

- does not depend on the structure of $\mathcal{X}$ and $\mathcal{Y}$
- includes structure of $\mathcal{R}$ and $\mathcal{T}$, allows for formulation in a general setup
- nonlinear F: combination of source and nonlinearity condition
- connection to conditional stability estimates

Source condition (cont.)

important special case: $\mathcal{X}$, $\mathcal{Y}$ Hilbert spaces, $\mathcal{R}(u) = \|u - u_0\|_{\mathcal{X}}^2$.

- if $\mathcal{T}(\hat{g}; g) = \|g - \hat{g}\|_{\mathcal{Y}}^2$:
  - spectral source condition + nonlinearity condition imply variational inequality
  - provided deterministic convergence analysis is optimal in case of linear F!

  J. Flemming.
  Generalized Tikhonov regularization - Basic theory and comprehensive results on convergence rates.
  PhD thesis, Chemnitz University of Technology, 2011.

- if $\mathcal{T}$ given as above ($\sim$ negative log-likelihood) and $F(\mathfrak{B}) \subset \mathbf{L}^\infty$ bounded:
  - it holds $\left\|F(u) - g^\dagger\right\|_{\mathbf{L}^2}^2 \leq C\mathcal{T}\left(g^\dagger; F(u)\right)$ for all $u \in \mathfrak{B}$

    J. M. Borwein and A. S. Lewis.
    Convergence of best entropy estimates.
    SIAM J. Optimization, 1:191–205, 1991.
  - thus spectral source condition + nonlinearity condition imply variational inequality!

Deterministic convergence analysis

Suppose

- the noise assumption is fulfilled with **err** ≥ 0 and
- the variational inequality holds true.

Theorem (error decomposition)

Then

$$\beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) \leq \frac{\mathbf{err}}{\alpha} + (-\varphi)^*\left(-\frac{1}{C_{\mathrm{err}}\alpha}\right)$$

for all $\alpha > 0$.

Fenchel conjugate:

$$(-\varphi)^*(s) = \sup_{\tau \geq 0}\left(s\tau + \varphi(\tau)\right).$$

Proof I

$$\beta\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) \leq \mathcal{R}\left(u_\alpha\right) - \mathcal{R}\left(u^\dagger\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F\left(u_\alpha\right)\right)\right)$$

$\boxed{\text{variational inequality}}$

Proof II

$$\begin{aligned}\beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) \leq& \mathcal{R}\left(u_\alpha\right) - \mathcal{R}\left(u^\dagger\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F\left(u_\alpha\right)\right)\right) \\ \leq& \frac{1}{\alpha}\left(\mathcal{S}\left(g^{\mathrm{obs}}; g^\dagger\right) - \mathcal{S}\left(g^{\mathrm{obs}}; F\left(u_\alpha\right)\right)\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F(u_\alpha)\right)\right)\end{aligned}$$

definition of u_α: $\mathcal{S}\left(g^{\mathrm{obs}}; F\left(u_\alpha\right)\right) + \alpha\mathcal{R}\left(u_\alpha\right) \leq \mathcal{S}\left(g^{\mathrm{obs}}; g^\dagger\right) + \alpha\mathcal{R}\left(u^\dagger\right)$

Proof III

$$\begin{aligned}\beta\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) \leq & \mathcal{R}\left(u_\alpha\right) - \mathcal{R}\left(u^\dagger\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F\left(u_\alpha\right)\right)\right) \\ \leq & \frac{1}{\alpha}\left(\mathcal{S}\left(g^{\mathrm{obs}}; g^\dagger\right) - \mathcal{S}\left(g^{\mathrm{obs}}; F\left(u_\alpha\right)\right)\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F(u_\alpha)\right)\right) \\ \leq & \frac{\mathbf{err}}{\alpha} - \frac{1}{C_{\mathrm{err}}\alpha}\mathcal{T}\left(g^\dagger; F\left(u_\alpha\right)\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F\left(u_\alpha\right)\right)\right)\end{aligned}$$

deterministic noise assumption: $\mathcal{S}\left(g^{\mathrm{obs}}; g\right) - \mathcal{S}\left(g^{\mathrm{obs}}; g^\dagger\right) \geq \frac{1}{C_{\mathrm{err}}}\mathcal{T}\left(g^\dagger; g\right) - \mathbf{err}$

Proof IV

$$\begin{aligned}
\beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) \leq & \mathcal{R}\left(u_\alpha\right) - \mathcal{R}\left(u^\dagger\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F\left(u_\alpha\right)\right)\right) \\
\leq & \frac{1}{\alpha}\left(\mathcal{S}\left(g^{\mathrm{obs}}; g^\dagger\right) - \mathcal{S}\left(g^{\mathrm{obs}}; F\left(u_\alpha\right)\right)\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F(u_\alpha)\right)\right) \\
\leq & \frac{\mathbf{err}}{\alpha} - \frac{1}{C_{\mathrm{err}}\alpha}\mathcal{T}\left(g^\dagger; F\left(u_\alpha\right)\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F\left(u_\alpha\right)\right)\right) \\
\leq & \frac{\mathbf{err}}{\alpha} + \sup_{\tau \geq 0}\left[\frac{\tau}{-C_{\mathrm{err}}\alpha} - (-\varphi)(\tau)\right]
\end{aligned}$$

Proof V

$$
\begin{aligned}
\beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) \leq & \mathcal{R}\left(u_\alpha\right) - \mathcal{R}\left(u^\dagger\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F\left(u_\alpha\right)\right)\right) \\
\leq & \frac{1}{\alpha}\left(\mathcal{S}\left(g^{\mathrm{obs}}; g^\dagger\right) - \mathcal{S}\left(g^{\mathrm{obs}}; F\left(u_\alpha\right)\right)\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F(u_\alpha)\right)\right) \\
\leq & \frac{\mathbf{err}}{\alpha} - \frac{1}{C_{\mathrm{err}}\alpha}\mathcal{T}\left(g^\dagger; F\left(u_\alpha\right)\right) + \varphi\left(\mathcal{T}\left(g^\dagger; F\left(u_\alpha\right)\right)\right) \\
\leq & \frac{\mathbf{err}}{\alpha} + \sup_{\tau \geq 0}\left[\frac{\tau}{-C_{\mathrm{err}}\alpha} - (-\varphi)(\tau)\right] \\
= & \frac{\mathbf{err}}{\alpha} + (-\varphi)^*\left(-\frac{1}{C_{\mathrm{err}}\alpha}\right).
\end{aligned}
$$

$\boxed{\text{definition of Fenchel conjugate: } (-\varphi)^*(s) = \sup_{\tau\geq 0}(s\tau + \varphi(\tau))}$

Deterministic convergence analysis (cont.)

$$\beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) \leq \frac{\mathbf{err}}{\alpha} + (-\varphi)^*\left(-\frac{1}{C_{\mathrm{err}}\alpha}\right)$$

Theorem (a priori rates)

The infimum of the right-hand side it attained at $\alpha = \overline{\alpha}$ if and only if

$$\frac{-1}{C_{\mathrm{err}}\overline{\alpha}} \in \partial(-\varphi)(C_{\mathrm{err}}\,\mathbf{err}) \qquad \left[\hat{=} \quad \bar{\alpha} = \frac{1}{C_{\mathrm{err}}\varphi'(C_{\mathrm{err}}\,\mathbf{err})}\right]$$

and in that case

$$\beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{\overline{\alpha}}, u^\dagger\right) \leq C_{\mathrm{err}}\varphi(\mathbf{err}).$$

Proof

Young's inequality:

$$s\tau \le f(\tau) + f^*(s) \qquad \text{for all} \qquad s,\tau \in \mathbb{R},$$
$$s\tau = f(\tau) + f^*(s) \qquad \Leftrightarrow \qquad \tau \in \partial f(s).$$

moreover $f^{**} = f$ whenever f is convex, proper and lower-semicontinuous.

$$\begin{aligned}\inf_{\alpha>0}\left[\frac{\mathbf{err}}{\alpha} + (-\varphi)^*\left(-\frac{1}{C_{\mathrm{err}}\alpha}\right)\right] &\overset{-\frac{1}{C_{\mathrm{err}}\alpha}=s}{=} -\sup_{s<0}\left[sC_{\mathrm{err}}\,\mathbf{err} - (-\varphi)^*(s)\right] \\ &= -(-\varphi)^{**}(C_{\mathrm{err}}\,\mathbf{err}) \\ &= \varphi(C_{\mathrm{err}}\,\mathbf{err}) \\ &\le C_{\mathrm{err}}\varphi(\mathbf{err})\end{aligned}$$

supremum is attained at $\alpha = \overline{\alpha}$ if and only if

$$\overline{s} \in \partial(-\varphi)(C_{\mathrm{err}}\,\mathbf{err}) \qquad \Leftrightarrow \qquad \frac{-1}{C_{\mathrm{err}}\overline{\alpha}} \in \partial(-\varphi)(C_{\mathrm{err}}\,\mathbf{err})$$

Deterministic convergence analysis (cont.)

Suppose moreover $\mathcal{X}$ Hilbert space, $\mathcal{R}(u) = \|u - u_0\|_{\mathcal{X}}^2$, $\beta \geq \frac{1}{2}$. Set

- $r > 1$
- $\alpha_j := \mathbf{err}\, r^{2j-2}$ for $j = 2, ..., m$ such that $\alpha_{m-1} < 1 \leq \alpha_m$
- $j_{\mathrm{bal}} := \max \left\{ j \leq m \;\middle|\; \left\| u_{\alpha_i} - u_{\alpha_j} \right\|_{\mathcal{X}} \leq 4\sqrt{2} r^{1-i} \text{ for all } i < j \right\}$

Theorem (a posteriori rates)

Then for $\mathbf{err} > 0$ sufficiently small:

$$\left\| u_{\alpha_{j_{\mathrm{bal}}}} - u^\dagger \right\|_{\mathcal{X}}^2 \leq 6r \min_{j=1,\dots,m} \left[\frac{\mathbf{err}}{\alpha_j} + (-\varphi)^* \left(-\frac{1}{C_{\mathrm{err}} \alpha_j} \right) \right].$$

If $\varphi^{1+\varepsilon}$ is additionally concave ($\varepsilon > 0$), then

$$\left\| u_{\alpha_{j_{\mathrm{bal}}}} - u^\dagger \right\|_{\mathcal{X}}^2 \leq 6 r^{1+\frac{1}{\varepsilon}} C_{\mathrm{err}} \varphi(\mathbf{err})$$

as $\mathbf{err} \searrow 0$.

Outline

Convergence rates for known φ

Suppose

- variational inequality holds true
- $\mathcal{X}$ Banach space, $u^\dagger \in \mathfrak{B} \subset \mathcal{X}$ bounded, closed and convex
- $\mathbb{M} \subset \mathbb{R}^d$ bounded and Lipschitz
- $F(u) \geq 0$ a.e. for all $u \in \mathfrak{B}$
- there exists a Sobolev index $s > \frac{d}{2}$ such that $F(\mathfrak{B})$ is a bounded subset of $H^s(\mathbb{M})$

A priori convergence rates (W., Hohage 2012)

Then for $\alpha = \alpha(t)$ such that $\frac{1}{\alpha} \in -\partial(-\varphi)\left(\frac{1}{\sqrt{t}}\right)$ we obtain the convergence rate

$$\mathbf{E}\left[\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right)\right] = \mathcal{O}\left(\varphi\left(\frac{1}{\sqrt{t}}\right)\right), \qquad t \to \infty.$$

Sketch of proof

- let $E_k := \left\{\sup_{u\in\mathfrak{B}} \mathbf{err}\left(F\left(u\right)\right) \leq \frac{\rho_k}{\sqrt{t}}\right\}$, $\rho_k = c^{-1}k$
  where $c \hat{=}$ constant from concentration inequality

$\rightsquigarrow$ $\mathbf{P}\left[E_k^c\right] \leq \exp\left(-c\rho_k\right) = \exp\left(-k\right)$

- $\boxed{\text{on } E_k\text{:}}$ $C_{\mathrm{err}} = 1$ and $\mathbf{err} = 2\sup_{u\in\mathfrak{B}} \mathbf{err}\left(F\left(u\right)\right) \leq 2\rho_k/\sqrt{t}$,
  i.e. deterministic convergence analysis is applicable

$$\mathbf{E}\left[\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right)\right] \leq \sum_{k=1}^{\infty} \mathbf{P}\left[E_k \setminus E_{k-1}\right] \max_{E_k} \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right)$$

$$\leq C\varphi\left(\frac{1}{\sqrt{t}}\right) \sum_{k=1}^{\infty} \mathbf{P}\left[E_k \setminus E_{k-1}\right] k^{\frac{1}{\varepsilon}}$$

sum converges due to $\mathbf{P}\left[E_k \setminus E_{k-1}\right] \leq \mathbf{P}\left[E_{k-1}^c\right] \leq \exp\left(-(k-1)\right)$

Convergence rates for unknown φ

Suppose moreover $\mathcal{X}$ Hilbert space, $\mathcal{R}(u) = \|u - u_0\|_{\mathcal{X}}^2$, $\beta \geq \frac{1}{2}$, $\varphi^{1+\varepsilon}$ concave ($\varepsilon > 0$). Set

- $r > 1$, $\tau > 0$ sufficiently large
- $\alpha_j := \frac{\tau \ln(t)}{\sqrt{t}} r^{2j-2}$ for $j = 2, \ldots, m$ such that $\alpha_{m-1} < 1 \leq \alpha_m$
- $j_{\mathrm{bal}} := \max \left\{ j \leq m \;\middle|\; \left\|u_{\alpha_i} - u_{\alpha_j}\right\|_{\mathcal{X}} \leq 4\sqrt{2} r^{1-i} \text{ for all } i < j \right\}$

A posteriori convergence rates (W., Hohage 2012)

Then we obtain

$$\mathbf{E}\left[\left\|u_{\alpha_{j_{\mathrm{bal}}}} - u^\dagger\right\|_{\mathcal{X}}^2\right] = \mathcal{O}\left(\varphi\left(\frac{\ln(t)}{\sqrt{t}}\right)\right) \qquad \text{as} \qquad t \to \infty.$$

Adaptivity causes a loss of a logarithmic factor!

A. Tsybakov.
On the best rate of adaptive estimation in some inverse problems.
C. R. Acad. Sci. Paris, 330:835–840, 2000.

Outline

Presented results

- proper setup for inverse problems with Poisson data:
  - Poisson processes
  - uniform concentration inequality
- improvements in theory:
  - convergence and convergence rates
  - generalized source conditions
  - a priori and a posteriori parameter choice
- regularization theory with general data fidelity terms

F. Werner and T. Hohage.
Convergence rates in expectation for Tikhonov-type regularization of Inverse Problems with Poisson data.
Inverse Problems, 28 104004, 2012.

Thank you for your attention!