Penalized likelihood estimators for Inverse Problems with Poisson data

Frank Werner[1,2], joint with Thorsten Hohage

[1]Statistical Inverse Problems in Biophysics Group
Max Planck Institute for Biophysical Chemistry

[2]Felix-Bernstein-Institute for Mathematical Statistics in the Biosciences
University of Göttingen

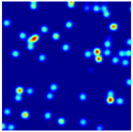
Outline

Outline

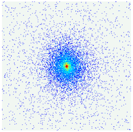
Photonic imaging I

We consider applications from photonic imaging:

- Fluorescence microscopy (in collaboration with the Lab of Stefan Hell, Nobel price for chemistry 2014)



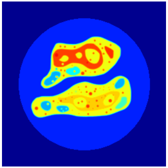
- X-ray diffraction imaging (SFB 755 'nanoscale photonic imaging' in collaboration with the Lab of Tim Salditt)



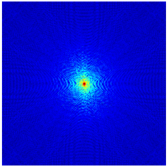
- Positron Emission Tomography
- Astronomical imaging
- ...

Photonic imaging II

- In all the aforementioned examples, the imaging process can be described (approximately) by an operator F relating the desired quantity of interest $\bar{u}$ and the ideal data $\bar{g}$:



$$\xrightarrow{F}$$



Unknown $\bar{u}$ Ideal data $\bar{g}$

- F is typically not continuously invertible due to its smoothing properties (e.g. compactness)
- Ideal data is not available, observables typically arise from measuring an energy
- At low energies the quantization of energy is the main source of noise
- Given an ideal photon detector, the observables obey a Poisson distribution

Discrete model

- Suppose the imaging procedure is modeled by a mapping $F : \mathbb{R}^n \to \mathbb{R}^m$
- Let $\bar{u} \in \mathbb{R}^n$ denote the exact solution we seek for and $\bar{g} := F(\bar{u})$, require $\bar{g} \geq 0$
- For the data $\mathbf{Y} \in \mathbb{R}^m$ the value $\mathbf{Y}_i$ is the number of photon counts in detector region $i \in \{1, ..., m\}$
- In the ideal case $\mathbf{Y} \in \mathbb{R}^m$ is a random variable such that $\mathbf{Y}_i \sim \mathrm{Poi}(\bar{g}_i)$, e.g.

$$\mathbf{P}[\mathbf{Y}_i = k] = \frac{(\bar{g}_i)^k}{k!} \exp(-\bar{g}_i).$$

Outline

Continuous model

- Now $F : \mathcal{X} \to \mathcal{Y}$ with Banach spaces $\mathcal{X}$ and $\mathcal{Y} \subset \mathbf{L}^1(\mathbb{M})$
- Consequently $\bar{u} \in \mathcal{X}$ and $\bar{g} := F(\bar{u}) \in \mathbf{L}^1(\mathbb{M})$, require $\bar{g} \geq 0$
- Say the total number of observed photons is n and their positions are $x_i \in \mathbb{M}$
- n can be influenced by the 'exposure time', mathematically described by a scaling factor $t > 0$

Data model

The observed data is a scaled **Poisson process** $G_t = \tilde{G}_t / t$, $\tilde{G}_t = \sum_{i=1}^{n} \delta_{x_i}$ with intensity $\bar{g}$, i.e. the measure $\tilde{G}_t$ satisfies the following axioms:

1. For each choice of disjoint, measurable sets $A_1, ..., A_n \subset \mathbb{M}$ the random variables $\tilde{G}_t(A_j)$ are stochastically independent.
2. $\mathbf{E}\left[\tilde{G}_t(A)\right] = \int_A t\bar{g}\,\mathrm{d}x$ for all $A \subset \mathbb{M}$ measurable.

Noise level I

- Poisson distribution incorporated, in fact it holds

$$\tilde{G}_t(A) \sim \mathrm{Poi}\left(t \int_A \bar{g}\,\mathrm{d}x\right)$$

  for all measurable $A \subset \mathbb{M}$, but ...
- ... no clear definition of the noise level so far!

Note: Similar statistical model is used by

L. Cavalier and J.-Y. Koo.
Poisson intensity estimation for tomographic data using a wavelet shrinkage approach
IEEE Transactions on Information Theory, 48(10):2794–2802, 2002.

A. Antoniadis and J. Bigot.
Poisson inverse problems.
The Annals of Statistics, 34(5):2132–2158, 2006.

Noise level II

- Recall: $\tilde{G}_t = \sum_{i=1}^{n} \delta_{x_i}$.
- For a function g let

$$\int_{\mathbb{M}} g \,\mathrm{d}\tilde{G}_t := \sum_{i=1}^{n} g\left(x_i\right)$$

- Then

$$\mathbf{E}\left[\int_{\mathbb{M}} g \,\mathrm{d}G_t\right] = \int_{\mathbb{M}} g\bar{g} \,\mathrm{d}x,$$

$$\mathbf{Var}\left[\int_{\mathbb{M}} g \,\mathrm{d}G_t\right] = \frac{1}{t^2}\mathbf{E}\left[\int_{\mathbb{M}} g^2 \,\mathrm{d}\tilde{G}_t\right] = \frac{1}{t}\int_{\mathbb{M}} g^2\bar{g} \,\mathrm{d}x.$$

- Thus the value of any bounded linear functional at the unknown quantity $\bar{g}$ can be estimated unbiasedly with a variance proportional to $\frac{1}{t}$.

$\rightsquigarrow$ For our analysis, such a property is needed uniformly!

A uniform concentration inequality for Poisson processes I

Uniform concentration inequality (Reynaud-Bouret 2003)

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

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.
Probability Theory and Related Fields, 126(1):103–153, 2003.

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

A uniform concentration inequality for Poisson processes II

Uniform concentration inequality (W., Hohage 2012)

Suppose $\mathbb{M} \subset \mathbb{R}^d$ is bounded & Lipschitz, $s > d/2$ and set

$$\mathfrak{G}(R) := \{g \in H^s(\mathbb{M}) : \|g\|_{H^s} \leq R\}.$$

Then $\exists\ c = c\left(\mathbb{M}, s, \|\bar{g}\|_{\mathbf{L}^1}\right) > 0$ such that

$$\mathbf{P}\left[\sup_{g \in \mathfrak{G}(R)} \left| \int_{\mathbb{M}} g\left(\mathrm{d}G_t - \bar{g}\,\mathrm{d}x\right)\right| \geq \frac{\rho}{\sqrt{t}}\right] \leq \exp\left(-\frac{\rho}{cR}\right)$$

for all $R \geq 1$, $t \geq 1$ and $\rho \geq cR$.

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

Outline

Regularization by projection

- Suppose $F = T$ is bounded, linear and positive definite (but not necessarily continuously invertible!), for simplicity $\mathcal{X} = \mathcal{Y} = \mathbf{L}^2(\mathbb{M})$.
- Regularization by projection: $V_n \subset \mathcal{X}$ with $\dim(V_n) < \infty$ and

$$\hat{u}_n^{\mathrm{proj}} := \operatorname*{argmin}_{u \in V_n} \| Tu - \bar{g} \|_{\mathcal{Y}}^2 \tag{1}$$

- If $\{v_1, ..., v_n\}$ is an orthonormal basis of V_n, then

$$\hat{u}_n^{\mathrm{proj}} \in V_n : \qquad \left\langle T \hat{u}_n^{\mathrm{proj}}, v_j \right\rangle = \int_{\mathbb{M}} v_j \bar{g} \, \mathrm{d}x, \qquad 1 \leq j \leq n.$$

- Define $\hat{u}_n^{\mathrm{proj}}$ also in the noisy case by replacing $\bar{g}$ by G_t.
- In principle the norm in (1) can be replaced by any other loss, but ...
- ... for the natural Poissonian choice (Kullback-Leibler divergence) this leads to problems proving existence of $\hat{u}_n^{\mathrm{proj}}$ and stability.

Some (simplified) results - Cavalier & Koo 2002

$$\hat{u}_n^{\mathrm{proj}} \in V_n : \qquad \left\langle T\hat{u}_n^{\mathrm{proj}}, v_j \right\rangle = \int_{\mathbb{M}} v_j \, \mathrm{d}G_t, \qquad 1 \leq j \leq n.$$

- $V_n =$ suitable wavelet space, $T =$ Radon transform
- The projection estimator exists and depends 'continuously' on the data
- For $t \to \infty$: If $u \in B^s_{p,q}$, then

  $$\mathbf{E}\left[\left\|\hat{u}_n^{\mathrm{proj}} - \bar{u}\right\|_{\mathcal{X}}^2\right] = \mathcal{O}\left(t^{-\frac{s}{2s+3}}\right)$$

  with an a priori choice of $n = n\,(t, s, p, q)$
- This convergence rate is optimal among all estimators (linear and nonlinear)
- For an adaptive choice of n, a $\log(t)$-factor is lost!

Note: The results of Cavalier & Koo 2002 are not based on the uniform concentration inequality, but on estimates for the Gaussian approximation.

Some (simplified) results - Antoniadis & Bigot 2006

$$\hat{u}_n^{\mathrm{proj}} \in V_n : \qquad \left\langle T\hat{u}_n^{\mathrm{proj}}, v_j \right\rangle = \int_{\mathbb{M}} v_j \,\mathrm{d}G_t, \qquad 1 \leq j \leq n.$$

- $V_n = \exp(U_n)$ with a suitable Wavelet space U_n, $T = \nu$-smoothing operator
- Corresponding estimator (if existent) is always non-negative and depends continuously on the data
- As $t \to \infty$, the estimator exists with probability 1.
- For $t \to \infty$: If $u = \exp(v)$, $v \in B^s_{p,q}$, then

  $$\mathbf{E}\left[\left\|\hat{u}_n^{\mathrm{proj}} - \bar{u}\right\|^2_{\mathcal{X}}\right] = \mathcal{O}\left(t^{-\frac{s}{2s+2\nu+d}}\right)$$

  with an a priori choice of n
- This convergence rate is optimal among all estimators
- For an adaptive choice of n, a $\log(t)$-factor is lost!

Note: Antoniadis & Bigot 2006 do use Reynaud-Bouret's concentration inequality.

Outline

Variational estimation I

- Disadvantage of the aforementioned methods: design does not rely on Poisson distribution!
- Different approach: likelihood methods!
  Minimize

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

  over all admissible u.
- Still ill-posed due to ill-posedness of the original problem $\rightsquigarrow$ penalization!

  $$\hat{u}_\alpha \in \underset{u \in \mathfrak{B}}{\operatorname{argmin}} \left[\mathcal{S}(G_t; F(u)) + \alpha \mathcal{R}(u)\right]$$

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

Variational estimation II

$$\hat{u}_\alpha \in \underset{u \in \mathfrak{B}}{\operatorname{argmin}} \left[\mathcal{S}\left(G_t; F\left(u\right)\right) + \alpha \mathcal{R}\left(u\right) \right]$$

- Main issue in the analysis: data fidelity term lacks of a triangle-type inequality!
- These methods (Tikhonov regularization) have been extensively studied in the (deterministic) Inverse Problems community: Eggermont & LaRiccia 1996, Resmerita & Anderssen 2007, Pöschl 2007, Bardsley & Laobeul 2008, Bardsley & Luttman 2009, Bardsley 2010, Flemming 2010 & 2011, Benning & Burger 2011, Lorenz & Worliczek 2013 ...
- Here: exploit deterministic results + concentration inequality to handle statistic case.

Data fidelity terms

- Negative log-likelihood for a scaled Poisson process:

$$\mathcal{S}_0\left(G_t; g\right) = \int_{\mathbb{M}} g \,\mathrm{d}x - \int_{\mathbb{M}} \ln\left(g\right) \,\mathrm{d}G_t, \qquad g \geq 0 \text{ a.e.}$$

- ideal data misfit functional for exact data $\bar{g}$ given by

$$\mathbf{E}\left[\mathcal{S}_0\left(G_t; g\right)\right] - \mathbf{E}\left[\mathcal{S}_0\left(G_t; \bar{g}\right)\right] = \int_{\mathbb{M}} \left[g - \bar{g} - \bar{g}\ln\left(\frac{g}{\bar{g}}\right)\right] \mathrm{d}x$$

  which is the Kullback-Leibler divergence $\mathbb{KL}\left(\bar{g}; g\right)$.

- we introduce a shift $\sigma > 0$ and consider

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

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

Assumptions

Assumptions on the problem

- $(\mathcal{X}, \tau_{\mathcal{X}})$ top. vector space, $\tau_{\mathcal{X}}$ weaker than norm topology, and $\mathfrak{B} \subset \mathcal{X}$ closed and convex.
- $F : \mathfrak{B} \to \mathbf{L}^1(\mathbb{M})$ with $\mathbb{M} \subset \mathbb{R}^d$ bounded & Lipschitz and
  1. $F : \mathfrak{B} \to \mathbf{L}^1(\mathbb{M})$ is $\tau_{\mathcal{X}} - \tau_\omega$-sequentially continuous.
  2. $F(u) \geq 0$ a.e. for all $u \in \mathfrak{B}$.
  3. There exists $s > d/2$ such that $F(\mathfrak{B})$ is a bounded subset of $H^s(\mathbb{M})$.

Assumptions on the method

- $\mathcal{R} : \mathfrak{B} \to (-\infty, \infty]$ is convex, proper and $\tau_{\mathcal{X}}$-sequentially lower semicontinuous.
- $\mathcal{R}$-sublevelsets $\left\{ u \in \mathcal{X} \;\middle|\; \mathcal{R}(u) \leq M \right\}$ are $\tau_{\mathcal{X}}$-sequentially pre-compact.

Consistency

$$\hat{u}_\alpha \in \underset{u \in \mathfrak{B}}{\operatorname{argmin}} \left[\mathcal{S}_\sigma\left(G_t; F(u)\right) + \alpha \mathcal{R}(u)\right]$$

Under the assumptions, for all $t > 0$ a minimizer $\hat{u}_\alpha$ exists with prob. 1.

Consistency (Hohage, W. 2015)

If α is chosen according to a rule $\bar{\alpha}$ fulfilling

$$\lim_{t\to\infty} \bar{\alpha}(t, G_t) = 0, \qquad \lim_{t\to\infty} \frac{\ln(t)}{\sqrt{t}\bar{\alpha}(t, G_t)} = 0,$$

then

$$\forall\ \varepsilon > 0\ : \qquad \lim_{t\to\infty} \mathbf{P}\left[\mathcal{D}_{\mathcal{R}}^{u^*}\left(\hat{u}_{\bar{\alpha}(t,G_t)}, \bar{u}\right) > \varepsilon\right] = 0,$$

T. Hohage and F. Werner.
Inverse Problems with Poisson Data: statistical regularization theory, applications and algorithms.
Topical review for *Inverse Problems*, in preparation, 2015

Source condition

- As the problem is ill-posed, convergence rates can only be obtained for $\bar{u}$ in a strict subset $M \subset \mathcal{X}$
- Here the set M is described by a variational inequality as source condition:

$$\beta \mathcal{D}_{\mathcal{R}}^{u^*}(u, \bar{u}) \leq \mathcal{R}(u) - \mathcal{R}(\bar{u}) + \varphi(\mathcal{T}(\bar{g}; F(u))) \tag{2}$$

  for all $u \in \mathfrak{B}$ with $\beta > 0$. φ is assumed to fulfill
  - $\varphi(0) = 0$,
  - $\varphi \nearrow$,
  - φ concave.
- Now the source set $M = M_{\mathcal{R}}^{\varphi}(\beta)$ consists of all $\bar{u} \in \mathfrak{B}$ satisfying (2).

Convergence rates

A priori convergence rates (W., Hohage 2012)

Then for $\alpha = \alpha(t)$ chosen appropriately we obtain for $\bar{u} \in M_{\mathcal{R}}^{\varphi}(\beta)$ that

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

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

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

A posteriori convergence rates (W., Hohage 2012)

For $\bar{u} \in M_{\mathcal{R}}^{\varphi}(\beta)$ we obtain

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

Outline

Iterative variational estimation I

- So far:

$$\hat{u}_\alpha \in \operatorname*{argmin}_{u \in \mathfrak{B}} \left[\mathcal{S}_\sigma\left(G_t; F(u)\right) + \alpha \mathcal{R}(u)\right]$$

- Disadvantage: If F is nonlinear, then the functional lacks of convexity!
- $\hat{u}_\alpha$ might be difficult to determine due to many local minima.
- Remedy: Combine with a Newton method! Choose $u_0 \in \mathfrak{B}$ and set

$$\hat{u}_{n+1} \in \operatorname*{argmin}_{u \in \mathfrak{B}} \left[\mathcal{S}_\sigma\left(G_t; F(\hat{u}_n) + F'[\hat{u}_n](u - \hat{u}_n)\right) + \alpha_n \mathcal{R}(u)\right]$$

- Choose the regularization parameters such that

$$\alpha_n \searrow 0, \qquad 1 \leq \frac{\alpha_n}{\alpha_{n+1}} \leq C$$

  as $n \to \infty$.
- Only free parameter: Stopping index $n \in \mathbb{N}$.

Iterative variational estimation II

$$\hat{u}_\alpha \in \operatorname*{argmin}_{u \in \mathfrak{B}} \left[\mathcal{S}_\sigma\left(G_t; F(u)\right) + \alpha \mathcal{R}(u)\right]$$

vs

$$\hat{u}_{n+1} \in \operatorname*{argmin}_{u \in \mathfrak{B}} \left[\mathcal{S}_\sigma\left(G_t; F(\hat{u}_n) + F'[\hat{u}_n](u - \hat{u}_n)\right) + \alpha_n \mathcal{R}(u)\right]$$

- Due to linearization: in each iteration a convex subproblem has to be solved.
- As we employ a Newton-method, we expect only a few iterations to be required.
- But still higher computational effort.
- Theory: Restriction on the nonlinearity required!

Nonlinearity condition

Generalized tangential cone condition

There exist constants η (later assumed to be sufficiently small) and $C \geq 1$ such that

$$\begin{aligned}\frac{1}{C}\mathcal{T}(\bar{g}; F(v)) - \eta\mathcal{T}(\bar{g}; F(u)) &\leq \mathcal{T}\left(\bar{g}; F(u) + F'(u; v-u)\right) \\ &\leq C\mathcal{T}(\bar{g}; F(v)) + \eta\mathcal{T}(\bar{g}; F(u))\end{aligned}$$

for all $u, v \in \mathfrak{B}$.

- condition is fulfilled with $\eta = 0$ if F is linear
- generalization of the tangential cone condition which is standard in inverse problems analysis
- can be weakened if the solution $\bar{u}$ is smooth enough ($\varphi(t) = \sqrt{t}$).

Convergence rates

A priori convergence rates (W., Hohage 2013)

Then for $n = n(t)$ chosen appropriately we obtain for $\bar{u} \in M_{\mathcal{R}}^{\varphi}(\beta)$ that

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

A posteriori convergence rates (W., Hohage 2013)

For n chosen by a Lepskiĭ-type rule we obtain for $\bar{u} \in M_{\mathcal{R}}^{\varphi}(\beta)$ that

$$\mathbf{E}\left[\left\|\hat{u}_{n_{\mathrm{bal}}} - \bar{u}\right\|_{\mathcal{X}}^2\right] = \mathcal{O}\left(\varphi\left(\frac{\ln(t)}{\sqrt{t}}\right)\right), \qquad t \to \infty.$$

T. Hohage and F. Werner.
Iteratively regularized Newton-type methods with general data misfit functionals and applications to Poisson data.
Numerische Mathematik 123(4), 745-779, 2013.

Consistency

$$\hat{u}_{n+1} \in \operatorname*{argmin}_{u \in \mathfrak{B}} \left[\mathcal{S}_\sigma \left(G_t; F\left(\hat{u}_n\right) + F'\left[\hat{u}_n\right]\left(u - \hat{u}_n\right)\right) + \alpha_n \mathcal{R}\left(u\right)\right]$$

- General consistency is unclear. But:
- it can be shown that for any $\bar{u}$ a variational source condition is fulfilled.
- So if the nonlinearity condition holds true we have

  $$\mathbf{E}\left[\mathcal{D}_{\mathcal{R}}^{u^*}\left(\hat{u}_n, \bar{u}\right)\right] \to 0, \qquad t \to \infty$$

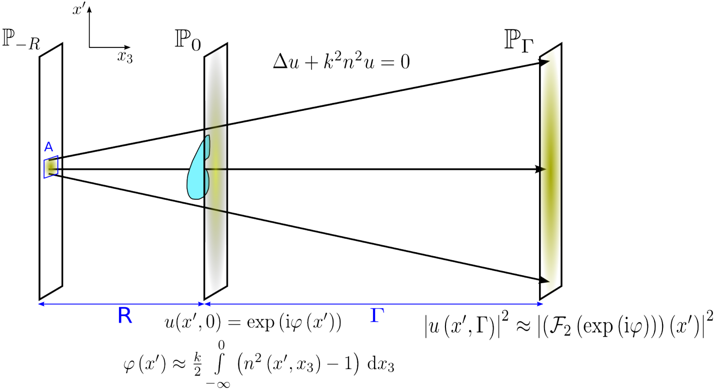
  for a specific a priori choice $n = n\left(t\right)$ and
- under additional conditions on β, $\mathcal{X}$ and $\mathcal{R}$ also

  $$\mathbf{E}\left[\left\|\hat{u}_{n_{\mathrm{bal}}} - \bar{u}\right\|_{\mathcal{X}}^2\right] \to 0, \qquad t \to \infty$$

  for the adaptive Lepskiĭ-type stopping rule.

Outline

Phase retrieval in coherent X-ray imaging

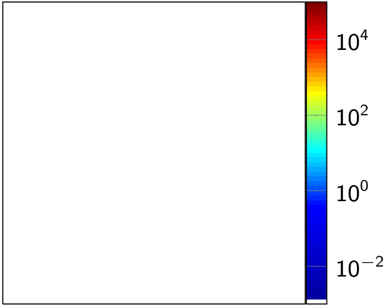


$$F : H^s(B_\rho) \longrightarrow L^\infty([-\kappa, \kappa]^2),$$

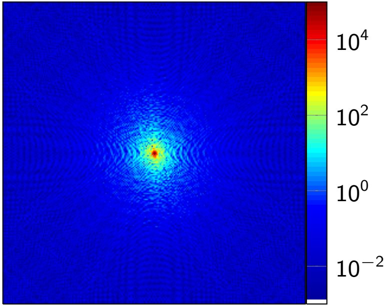
$$F(\varphi)(\xi) = \left| \int_{B_\rho} \exp(-\mathrm{i}\xi \cdot x') \exp(\mathrm{i}\varphi(x')) \,\mathrm{d}x' \right|^2 = |\mathcal{F}_2(\exp(\mathrm{i}\varphi))(\xi)|^2 .$$

Influence of t

Logarithmic plots of simulated Poisson and exact data for the phase retrieval problem:

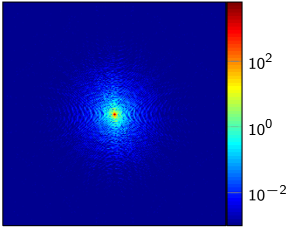


(a) simulated Poisson data; we expect 10^4 photons per time step

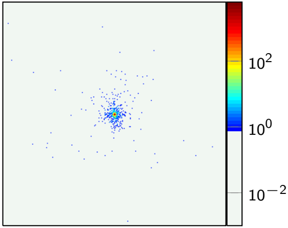


(b) exact data; total number of counts 10^6

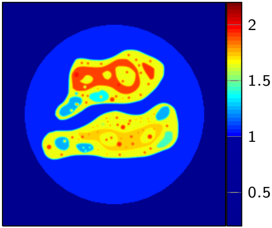
Results for $t = 10^4$



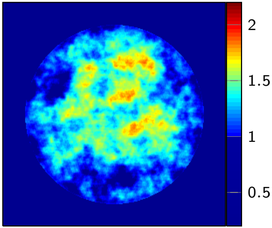
(a) exact data



(b) Poisson data



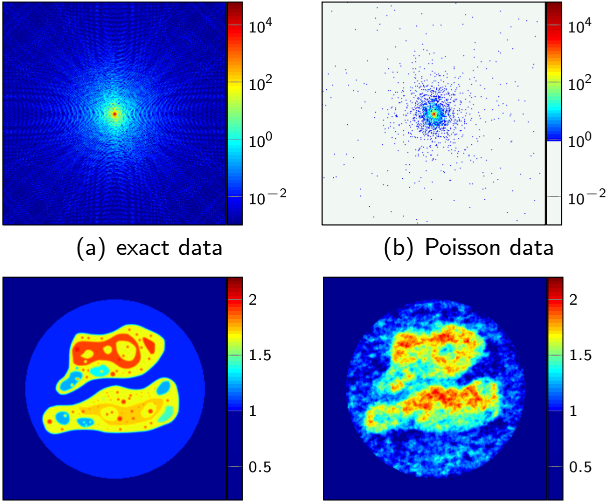
(c) exact solution



(d) reconstruction

K. Giewekemeyer et al, *Phys. Rev. A*, 83:023804, **2011**.

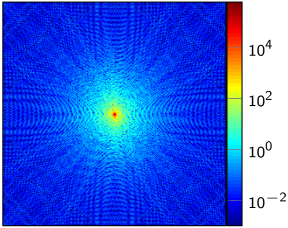
Results for $t = 10^5$



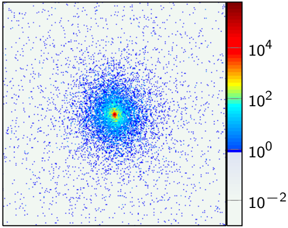
(a) exact data (b) Poisson data

(c) exact solution (d) reconstruction

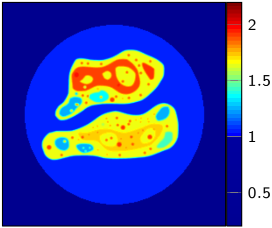
Results for $t = 10^6$



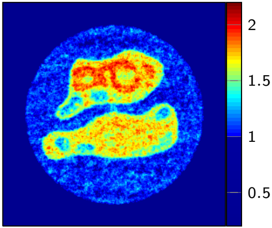
(a) exact data



(b) Poisson data



(c) exact solution



(d) reconstruction

Outline

Presented results

- sound mathematical model joining statistics and inverse problems
- review of some results for projection-type estimators
- (iterative) penalized likelihood estimators:
  - $\rightsquigarrow$ motivated by Poisson distribution
  - $\rightsquigarrow$ consistency
  - $\rightsquigarrow$ convergence rates
  - $\rightsquigarrow$ show a good performance in simulations

Thank you for your attention!