Exact Support Recovery for Sparse Spikes Deconvolution

Gabriel Peyré

Joint work with Vincent Duval & Quentin Denoyelle







www.numerical-tours.com











Astrophysics (2D)

Overview

• Sparse Spikes Super-resolution

• Robust Support Recovery

• Asymptotic Positive Measure Recovery





Radon measure m on $\mathbb{T} = (\mathbb{R}/\mathbb{Z})^d$.

Discrete measure:

$$m_{a,x} = \sum_{i=1}^{N} a_i \delta_{x_i}, \ a \in \mathbb{R}^N, x \in \mathbb{T}^N$$



Radon measure m on $\mathbb{T} = (\mathbb{R}/\mathbb{Z})^d$.

Discrete measure:

$$m_{a,x} = \sum_{i=1}^{N} a_i \delta_{x_i}, \ a \in \mathbb{R}^N, x \in \mathbb{T}^N$$

Linear measurements:

$$y = \Phi(m) + w \qquad \varphi \in C^2(\mathbb{T} \times \mathbb{T})$$
$$\Phi(m) = \int_{\mathbb{T}} \varphi(x, \cdot) dm(x)$$



Radon measure m on $\mathbb{T} = (\mathbb{R}/\mathbb{Z})^d$.

Discrete measure:

$$m_{a,x} = \sum_{i=1}^{N} a_i \delta_{x_i}, \ a \in \mathbb{R}^N, x \in \mathbb{T}^N$$

Linear measurements:

$$y = \Phi(m) + w \qquad \varphi \in C^2(\mathbb{T} \times \mathbb{T})$$
$$\Phi(m) = \int_{\mathbb{T}} \varphi(x, \cdot) dm(x)$$

Example: 1-D (d = 1) convolution $\varphi(x, t) = \varphi(t - x)$







Radon measure m on $\mathbb{T} = (\mathbb{R}/\mathbb{Z})^d$.

Discrete measure:

$$m_{a,x} = \sum_{i=1}^{N} a_i \delta_{x_i}, \ a \in \mathbb{R}^N, x \in \mathbb{T}^N$$

Linear measurements:

$$y = \Phi(m) + w \qquad \varphi \in C^2(\mathbb{T} \times \mathbb{T})$$
$$\Phi(m) = \int_{\mathbb{T}} \varphi(x, \cdot) \mathrm{d}m(x)$$

Example: 1-D
$$(d = 1)$$
 convolution
 $\varphi(x, t) = \varphi(t - x)$

Minimum separation:

$$\Delta = \min_{i \neq j} |x_i - x_j|$$

$$\rightarrow \text{ Signal-dependent recovery criteria}$$









Discrete ℓ^1 regularization:

Computation grid $z = (z_k)_{k=1}^K$.







Why ℓ^1 ? " ℓ^0 ball"









sparse

convex

Grid-free regularization: total variation of measures:

 $|m|(\mathbb{T}) = \sup\left\{\int \eta \mathrm{d}m : \eta \in C(\mathbb{T}), \|\eta\|_{\infty} \leqslant 1\right\}$

Grid-free regularization: total variation of measures: $|m|(\mathbb{T}) = \sup\left\{\int \eta \mathrm{d}m \ : \ \eta \in C(\mathbb{T}), \|\eta\|_{\infty} \leqslant 1\right\}$

$$dm(x) = f(x)dx$$

$$|m|(\mathbb{T}) = \int |f| = ||f||_{L^1}$$

Grid-free regularization: total variation of measures: $|m|(\mathbb{T}) = \sup \left\{ \int \eta \mathrm{d}m \ : \ \eta \in C(\mathbb{T}), \|\eta\|_{\infty} \leqslant 1 \right\}$

$$dm(x) = f(x)dx$$

$$m_{a,x}$$

$$m_{a,x}$$

$$|m|(\mathbb{T}) = \int |f| = ||f||_{L^1}$$

$$|m_{a,x}|(\mathbb{T}) = ||a||_{\ell^1}$$













Competitors: Prony's methods (MUSIC, ESPRIT, FRI).





Overview

• Sparse Spikes Super-resolution

Robust Support Recovery

• Asymptotic Positive Measure Recovery



Low-pass filter $\operatorname{supp}(\hat{\varphi}) = [-f_c, f_c].$

When is m_0 solution of $\mathcal{P}_0(\Phi m_0)$?





 $\min \{ |m|(\mathbb{T}); \Phi m = y \} \quad (\mathcal{P}_0(y))$ mLow-pass filter supp $(\hat{\varphi}) = [-f_c, f_c].$ When is m_0 solution of $\mathcal{P}_0(\Phi m_0)$? $\Delta = 0.55/f_c$ Theorem: [Candès, Fernandez G.] $\Delta > \frac{1.26}{f_c} \Rightarrow m_0 \text{ solves } \mathcal{P}_0(\Phi m_0).$ $\Delta = 0.45/f_c$ $\Delta = 0.3/f_c$

 $\Delta = 0.1/f_c$

$$\min_{m} \{|m|(\mathbb{T}) ; \Phi m = y\} \quad (\mathcal{P}_{0}(y))$$
Low-pass filter supp $(\hat{\varphi}) = [-f_{c}, f_{c}].$
When is m_{0} solution of $\mathcal{P}_{0}(\Phi m_{0})$?
$$\overline{Theorem: [Candès, Fernandez G.]} \\ \Delta > \frac{1.26}{f_{c}} \Rightarrow m_{0} \text{ solves } \mathcal{P}_{0}(\Phi m_{0}).$$
How close to m_{0}
are solutions of $\mathcal{P}_{\lambda}(\Phi m_{0} + w)$?
$$\overline{\Delta} = 0.3/f_{c}$$

$$\Delta = 0.1/f_{c}$$

$$\min_{m} \{|m|(\mathbb{T}) ; \Phi m = y\} \quad (\mathcal{P}_{0}(y))$$
Low-pass filter supp $(\hat{\varphi}) = [-f_{c}, f_{c}].$
When is m_{0} solution of $\mathcal{P}_{0}(\Phi m_{0})$?

$$\frac{Theorem: [Candès, Fernandez G.]}{\Delta > \frac{1.26}{f_{c}} \Rightarrow m_{0} \text{ solves } \mathcal{P}_{0}(\Phi m_{0}).$$
How close to m_{0}
are solutions of $\mathcal{P}_{\lambda}(\Phi m_{0} + w)$?

Weighted L^{2} error:
$$\rightarrow [Candès, Fernandez-G. 2012]$$
Support localization:
$$\rightarrow [Fernandez-G.][de Castro 2012]$$



Open problems: Exact support recovery? General kernels?

From Primal to Dual

$$\min_{m} |m|(\mathbb{T}) + \frac{1}{2\lambda} \|\Phi m - y\|^2$$

From Primal to Dual

$$\begin{split} & \min_{m} |m|(\mathbb{T}) + \frac{1}{2\lambda} \|\Phi m - y\|^2 \end{split}$$
$$= & \min_{m} \left[\sup_{\|\eta\|_{\infty} \leqslant 1} -\langle \eta, m \rangle + \frac{1}{2\lambda} \|\Phi m - y\|^2 \right] \end{split}$$






Ideal low-pass filter: $\left| \begin{array}{c} \rightarrow \eta = \Phi^* p \text{ trigonometric polynomial.} \end{array} \right|$













Ideal low-pass filter: $\begin{vmatrix} \rightarrow \eta = \Phi^* p \text{ trigonometric polynomial.} \\ \rightarrow \text{Interpolates spikes location and sign.} \end{vmatrix}$





 $\begin{array}{l} \textit{Ideal low-pass filter:} \ \left| \begin{array}{l} \rightarrow \eta = \Phi^* p \ \text{trigonometric polynomial.} \\ \rightarrow \ \text{Interpolates spikes location and sign.} \\ \rightarrow \ |\eta(t)|^2 = 1 \text{: polynomial equation of } \sup (m). \end{array} \right.$



$$\sum_{m} \frac{|m|(\mathbb{T}) + \frac{1}{2\lambda} \|\Phi m - y\|^2}{\|\Phi^* p\|_{\infty} \leq 1} \left(p, y \right) - \frac{\lambda}{2} \|p\|^2$$







 $p \in \mathcal{D}_0(y)$





Definition: for any
$$m_0$$
 solution of $\mathcal{P}_0(y)$,
 $\eta_0 = \Phi^* p_0 = \underset{\eta = \Phi^* p \in \partial |m_0|(\mathbb{T})}{\operatorname{argmin}} \|p\|$



 η_0

Input measure: $m_0 = m_{a,x}$.



Input measure: $m_0 = m_{a,x}$.



Input measure: $m_0 = m_{a,x}$. $\eta_0 \stackrel{\text{def.}}{=} \underset{\eta = \Phi^* p}{\operatorname{argmin}} \|p\| \text{ s.t. } \begin{cases} \forall i, \ \eta(x_i) = \operatorname{sign}(a_i), \\ \|\eta\|_{\infty} \leq 1. \end{cases}$ η_0 $-\eta_V$ $\exists \eta_0 \iff m_0 \text{ solves } \mathcal{P}_0(\Phi m_0)$ $\eta_{V} \stackrel{\text{def.}}{=} \underset{\eta = \Phi^{*}p}{\operatorname{argmin}} \|p\| \text{ s.t. } \begin{cases} \forall i, \ \eta(x_{i}) = \operatorname{sign}(a_{i}), \\ \forall i, \ \eta'(x_{i}) = 0. \end{cases}$ $\eta_0 = \eta_V$ Proposition: $\eta_V = \Phi^* A_x^+(\operatorname{sign}(a); 0)$ where $A_x(b) = \sum_i b_i^1 \varphi(x_i, \cdot) + b_i^2 \varphi'(x_i, \cdot)$



Non-degenerate certificate:
$$\eta \in ND(m_{a,x})$$
:
 $\iff \forall t \notin \{x_1, \dots, x_N\}, |\eta(t)| < 1 \text{ and } \forall i, \eta''(x_i) \neq 0$



Non-degenerate certificate: $\eta \in ND(m_{a,x})$: $\iff \forall t \notin \{x_1, \dots, x_N\}, |\eta(t)| < 1 \text{ and } \forall i, \eta''(x_i) \neq 0$

Theorem: $\eta_V \in ND(m_0) \implies \eta_V = \eta_0$

Support Stability Theorem

$$\eta_{\lambda} = \Phi^* p_{\lambda} \xrightarrow{\lambda \to 0} \eta_0 = \Phi^* p_0$$

 $\operatorname{supp}(m_{\lambda}) \subset \{|\eta_{\lambda}| = 1\}$

Support Stability Theorem

$$\eta_{\lambda} = \Phi^* p_{\lambda} \xrightarrow{\lambda \to 0} \eta_0 = \Phi^* p_0$$

$$supp(m_{\lambda}) \subset \{ |\eta_{\lambda}| = 1 \}$$

$$f \eta_0 \in ND(m_0) \text{ then } supp(m_{\lambda}) \to supp(m_0)$$

Support Stability Theorem





Overview

• Sparse Spikes Super-resolution

• Robust Support Recovery

Asymptotic Positive Measure Recovery

Recovery of Positive Measures



 $\rightarrow m_0$ is recovered when there is no noise.





de Castro et al.

2011

Recovery of Positive Measures

Input measure: $m_0 = m_{a,x}$ where $a \in \mathbb{R}^N_+$. Theorem: let $\Phi m = (\int e^{-2i\pi kt} dm(t))_{k=-f_c}^{f_c}$ and $\eta_S(t) = 1 - \rho \prod_{i=1}^N \sin(\pi(t - x_i))^2$ for $N \leq f_c$ and ρ small enough, $\eta_S \in \mathcal{D}(m_0)$. $\rightarrow m_0$ is recovered when there is no noise. \longrightarrow behavior as $\forall i, x_i \rightarrow 0$? [Morgenshtern, Candès, 2015] discrete ℓ^1 robustness. [Demanet, Nguyen, 2015] discrete ℓ^0 robustness.





de Castro et

al

2011

Recovery of Positive Measures

Input measure: $m_0 = m_{a,x}$ where $a \in \mathbb{R}^N_+$. Theorem: let $\Phi m = (\int e^{-2i\pi kt} dm(t))_{k=-f_c}^{f_c}$ and $\eta_S(t) = 1 - \rho \prod_{i=1}^N \sin(\pi(t - x_i))^2$ for $N \leq f_c$ and ρ small enough, $\eta_S \in \mathcal{D}(m_0)$. $\rightarrow m_0$ is recovered when there is no noise. \longrightarrow behavior as $\forall i, x_i \rightarrow 0$? [Morgenshtern, Candès, 2015] discrete ℓ^1 robustness. [Demanet, Nguyen, 2015] discrete ℓ^0 robustness. \rightarrow noise robustness of support recovery ?





[de

Castro et

al

2011

Comparison of Certificates



Asymptotic of Vanishing Certificate

$$m_0 = m_{a,\Delta x}$$
 where $\Delta \to 0$

Vanishing Derivative pre-certificate:

$$\eta_V \stackrel{\text{def.}}{=} \operatorname{argmin} \|p\|$$

 $\eta = \Phi^* p$
s.t. $\forall i, \begin{cases} \eta(\Delta x_i) = 1, \\ \eta'(\Delta x_i) = 0. \end{cases}$



Asymptotic of Vanishing Certificate

$$m_0 = m_{a,\Delta x}$$
 where $\Delta \to 0$

Vanishing Derivative pre-certificate:

$$\eta_V \stackrel{\text{def.}}{=} \operatorname{argmin} \|p\|_{\eta = \Phi^* p}$$

s.t. $\forall i, \begin{cases} \eta(\Delta x_i) = 1, \\ \eta'(\Delta x_i) = 0. \end{cases}$



Asymptotic of Vanishing Certificate



Asymptotic Certificate

$$(2N - 1) \text{-Non degenerate:}$$
$$\eta_{W} \in \text{ND}_{N}$$
$$\iff \begin{cases} \forall t \neq 0, |\eta_{W}(t)| < 1\\ \eta_{W}^{(2N)}(0) \neq 0 \end{cases}$$



Asymptotic Certificate



Asymptotic Robustness

Theorem: If
$$\eta_W \in \mathrm{ND}_N$$
, letting $m_0 = m_{a,\Delta x}$, then
for $\left(\frac{w}{\lambda}, \frac{w}{\Delta^{2N-1}}, \frac{\lambda}{\Delta^{2N-1}}\right) = O(1)$
the solution of $\mathcal{P}_{\lambda}(y)$ for $y = \Phi(m_0) + w$ is
 $\sum_{i=1}^{N} a_i^{\star} \delta_{\Delta x_i^{\star}}$ where $\|(x, a) - (x^{\star}, a^{\star})\| = O\left(\frac{\|w\| + \lambda}{\Delta^{2N-1}}\right)$



Asymptotic Robustness

Theorem: If
$$\eta_W \in \mathrm{ND}_N$$
, letting $m_0 = m_{a,\Delta x}$, then
for $\left(\frac{w}{\lambda}, \frac{w}{\Delta^{2N-1}}, \frac{\lambda}{\Delta^{2N-1}}\right) = O(1)$
the solution of $\mathcal{P}_{\lambda}(y)$ for $y = \Phi(m_0) + w$ is
 $\sum_{i=1}^{N} a_i^* \delta_{\Delta x_i^*}$ where $\|(x, a) - (x^*, a^*)\| = O\left(\frac{\|w\| + \lambda}{\Delta^{2N-1}}\right)$
 $y = \Phi m_{a,\Delta x} + w$ [Noise: $w = \lambda w_0$.
Regularization: $\lambda = \lambda_0 \Delta^{\alpha}$
 $x_i^* \qquad \lambda_{\max} \qquad x_i^* \qquad \alpha < 2N-1$ λ_0 λ_0
When is η_W Non-degenerate ?

Proposition: one has $\eta_W^{(2N)}(0) < 0$. \longrightarrow "locally" non-degenerate.

When is η_W Non-degenerate ?

Proposition: one has $\eta_W^{(2N)}(0) < 0$. \longrightarrow "locally" non-degenerate.



Gaussian Deconvolution

Gaussian convolution:
$$\varphi(x,t) = e^{-\frac{|x-t|^2}{2\sigma^2}} \Phi(m) \stackrel{\text{def.}}{=} \int \varphi(x,\cdot) dm(x)$$

Proposition: $\eta_W(x) = e^{-\frac{x^2}{4\sigma^2}} \sum_{k=0}^{N-1} \frac{(x/2\sigma)^{2k}}{k!}$
In particular, η_W is non-degenerate.

 \longrightarrow Gaussian deconvolution is support-stable.



Laplace transform: $\varphi(x,t) = e^{-xt} \quad \Phi(m) \stackrel{\text{\tiny def.}}{=} \int \varphi(x,\cdot) dm(x) \quad \text{[with E. Soubies]}$







Total internal reflection fluorescence microscopy (TIRFM) \rightarrow multiple angles $\theta(t)$.







Deconvolution of measures:

 $\longrightarrow L^2$ errors are not well-suited.

Weak-* convergence. Optimal transport distance. Exact support estimation.



Deconvolution of measures:

 $\longrightarrow L^2$ errors are not well-suited.

Weak-* convergence. Optimal transport distance. Exact support estimation.



Low-noise behavior:

 \longrightarrow dictated by $\overline{\eta_0}$.

 $r: \longrightarrow \text{checkable via } \eta_V.$ $\longrightarrow \text{asymptotic via } \eta_W.$

Deconvolution of measures:

 $\longrightarrow L^2$ errors are not well-suited.

Weak-* convergence. Optimal transport distance.

Exact support estimation.



 \longrightarrow dictated by $\overline{\eta_0}$.

Low-noise behavior: \longrightarrow checkable via η_V .

 \rightarrow asymptotic via η_W .

Lasso on discrete grids: similar η_0 -analysis applies. \longrightarrow Relate discrete and continuous recoveries.

Deconvolution of measures:

 $\longrightarrow L^2$ errors are not well-suited.

Weak-* convergence. Optimal transport distance.

Exact support estimation.



 \longrightarrow dictated by $\overline{\eta_0}$.

Low-noise behavior: \longrightarrow checkable via η_V .

 \rightarrow asymptotic via η_W .

Lasso on discrete grids: similar η_0 -analysis applies.

 \longrightarrow Relate discrete and continuous recoveries.

Open problem: other regularizations (e.g. piecewise constant) ? see [Chambolle, Duval, Peyré, Poon 2016] for TV denoising.