The dual approach to non-negative super-resolution: impact on primal reconstruction accuracy

Stéphane Chrétien
National Physical Laboratory
London, UK
stephane.chretien@npl.co.uk

Andrew Thompson
National Physical Laboratory
London, UK
andrew.thompson@npl.co.uk

Bogdan Toader
University of Oxford
Oxford, UK
toader@maths.ox.ac.uk

***Abstract*—We study the problem of super-resolution, where we recover the locations and weights of non-negative point sources from a few samples of their convolution with a Gaussian kernel. It has been recently shown that exact recovery is possible by minimising the total variation norm of the measure. An alternative practical approach is to solve its dual. In this paper, we study the stability of solutions with respect to the solutions to the dual problem. In particular, we establish a relationship between perturbations in the dual variable and the primal variables around the optimiser. This is achieved by applying a quantitative version of the implicit function theorem in a non-trivial way.**

***Index Terms*—super-resolution, perturbation analysis**

I. Problem setup

In the study of non-negative super-resolution, we want to estimate a signal x which consists of a number of point sources with unknown locations and non-negative magnitudes, from only a few measurements of the convolution of x with a known kernel ϕ. This is a problem that arises in a number of applications, for example fluorescence microscopy [8], astronomy [9] and ultrasound imaging [10]. In such applications, the measurement device has limited resolution and cannot distinguish between distinct point sources in the input signal x. This is often modelled as a deconvolution problem with a Gaussian kernel.

Specifically, let x be a non-negative measure on $I = [0, 1]$ consisting of K unknown non-negative point sources:

$$x = \sum_{k=1}^{K} a_k \delta_{t_k},$$

with $a_k > 0$, for all $k = 1, \ldots, K$, and let $\mathbf{y} \in \mathbb{R}^M$ be the vector of measurements obtained by sampling the convolution of x with a known kernel ϕ (e.g. a Gaussian) at locations s_j:

$$y_j = \int_I \phi(t - s_j) x(\mathrm{d}t),$$

for all $j = 1, \ldots, M$. Note that, because x is a discrete measure, each entry in $\mathbf{y}$ is of the form:

$$y_j = \sum_{k=1}^{K} a_k \phi(t_k - s_j),$$

for all $j = 1, \ldots, M$. Let $\Phi(t) = [\phi(t - s_1), \ldots, \phi(t - s_M)]^T$, then x can be recovered by solving the following program: [1]

$$\min_{\hat{x} \geq 0} \int_I \hat{x}(\mathrm{d}t) \quad \text{subject to} \quad \mathbf{y} = \int_I \Phi(t) \hat{x}(\mathrm{d}t). \tag{1}$$

The problem of super-resolution has been studied extensively in the literature since the seminal paper [1], which addressed the case of complex amplitudes. In [2], the authors showed that when the coefficients a_k, $k = 1, \ldots, K$ are positive, which corresponds to the present setting, exact recovery is possible without separation of sources.

The dual of problem (1) is

$$\max_{\lambda \in \mathbb{R}^M} \mathbf{y}^T \lambda \quad \text{subject to} \quad \sum_{j=1}^{M} \lambda_j \phi(t - s_j) \leq 1, \quad \forall t \in I. \tag{2}$$

The dual problem (2) is a finite-dimensional problem with infinitely many constraints, known as a semi-infinite program. Such problems can be solved using a number of algorithms including exchange methods [7] and sequential quadratic programming [6]. The advantage over algorithms that solve the primal problem (for example the ADCG algorithm [3]) is working in a finite dimensional space, which simplifies the analysis.

Consider a solution λ^* of the dual problem (2) which corresponds to a dual certificate, namely a function

$$q(s) = \sum_{j=1}^{M} \lambda_j^* \phi(s - s_j), \tag{3}$$

which satisfies the conditions:

$$q(t_i) = 1, \quad \forall i = 1, \ldots, K, \tag{4}$$

$$q(s) < 1, \quad \forall s \neq t_i, \quad \forall i = 1, \ldots, K. \tag{5}$$

Then the local maximisers of $q(s)$ correspond to the source locations $\{t_k\}_{k=1}^K$ and the amplitudes $\{a_k\}_{k=1}^K$ are found by solving a linear system.

In this paper, we analyse how small perturbations of λ^* affect the local maximisers of $q(s)$ in the case when the convolution kernel is Gaussian $\phi(t) = e^{-t^2/\sigma^2}$. The outcome is a bound on how far the estimated locations t_k and magnitudes

[1] We assume that the measurements are exact. We treat the case when the measurements are noisy in the journal version of this paper.

This publication is based on work supported by the EPSRC Centre For Doctoral Training in Industrially Focused Mathematical Modelling (EP/L015803/1) in collaboration with the National Physical Laboratory and by the Alan Turing Institute under the EPSRC grant EP/N510129/1 and the Turing Seed Funding grant SF019.

a_k are from their true values obtained for exact λ^*. This gives us an insight into the size of the error in the locations and magnitudes when we apply an optimisation algorithm to the dual of the super-resolution problem.

II. Bound on the error as λ is perturbed

In this section we present our main results, namely two theorems that give bounds on the perturbations around the source locations t_k and the magnitudes a_k respectively, as the dual variable is perturbed away from the optimiser λ^*, when the convolution kernel is a Gaussian with known width σ: $\phi(t) = e^{-t^2/\sigma^2}$.

Theorem 1: **(Dependence of t on λ)** Let $\lambda^* \in \mathbb{R}^M$ be a solution of the dual program (2) with ϕ Gaussian as given above such that the dual certificate $q(s)$ defined in (3) satisfies conditions (4) and (5), λ a perturbation of λ^* in a ball of radius δ_λ and t an arbitrary local maximiser of $q_\lambda(s) = \sum_{j=1}^M \lambda_j \phi(s - s_j)$ so that for $\lambda = \lambda^*$, the corresponding local maximiser t^* is a source location of x. Then

$$|t - t^*| \leq C_{t^*} \|\lambda - \lambda^*\|_2,$$

provided that

$$\delta_\lambda \leq \frac{|q''(t^*)|^2 \sigma^3 \sqrt{e}}{4\sqrt{2}\,(2 + cR)\,M},$$

where

$$C_{t^*} = \frac{1}{4 + cR}\left[1 + \frac{2\sqrt{2M}(2 + cR)}{|q''(t^*)|\sqrt{e}}\right], \tag{6}$$

$$R = \frac{\|\lambda^*\|_2}{\sigma}, \tag{7}$$

and $c \approx 3.9036$ is a universal constant.

One of the main conclusions which can be drawn from this result is that the primal spike location error is controlled in l_∞, but degrades as a function of the number of measurements in the order of $\sqrt{M}$. Of crucial importance is the curvature of the dual certificate at the true solution: the flatter the certificate, the worse the estimation error. Our theorem also gives important information about the accuracy in the dual variable required to guarantee our upper bound on the error of recovery. This accuracy is of the inverse order of the number of measurements, which is quite a stringent constraint. Both the M and the $\sqrt{M}$ factors are a consequence of the way we bound sums of shifted copies of the kernel, namely $\sum_{j=1}^M \phi(t - s_j) \leq M \max_{t \in \mathbb{R}} \phi(t)$. Given the fast decay of the Gaussian, it is clear that this is not a tight bound. However, any bound would reflect the density of samples close to each source location.

We will now give a result regarding the perturbation of the magnitudes a_k when λ^* is perturbed. Let Φ be the matrix whose entries are defined as

$$\Phi_{ij} = \phi(t_j - s_i), \tag{8}$$

and $\mathbf{t}^*$ and $\mathbf{a}^*$ the vectors of true source locations and weights:

$$\mathbf{t}^* = [t_1, \ldots, t_K]^T, \qquad \mathbf{a}^* = [a_1, \ldots, a_K]^T.$$

When we solve (2) exactly, we obtain the source locations by finding the local maximisers of $q(s)$. Then the vector of weights $\mathbf{a}^*$ is found by solving the system

$$\Phi \mathbf{a} = \mathbf{y}.$$

When the source locations are perturbed, we denote the resulting perturbed data matrix by:

$$\tilde{\Phi} = \Phi + E, \tag{9}$$

and we calculate the vector of perturbed weights $\tilde{\mathbf{a}}$ as the solution of the least squares problem

$$\min_{\mathbf{a}} \|\tilde{\Phi}\mathbf{a} - \mathbf{y}\|_2. \tag{10}$$

The following theorem gives us a bound on the difference $\|\mathbf{a}^* - \tilde{\mathbf{a}}\|_2$ between the vector of true weights $\mathbf{a}^*$ and the vector of weights $\tilde{\mathbf{a}}$ obtained by solving the least squares problem (10) with the perturbed matrix $\tilde{\Phi}$, as a function of the difference $\|\tilde{\mathbf{t}} - \mathbf{t}^*\|_2$ between the perturbed source locations $\tilde{\mathbf{t}}$ and the true source locations $\mathbf{t}^*$.

Theorem 2: **(Dependence of a on t)** Let $\mathbf{t}^* \in [0,1]^K$ be the vector of true source locations and $\tilde{\mathbf{t}} \in [0,1]^K$ the perturbed source locations, such that:

$$\|\mathbf{t}^* - \tilde{\mathbf{t}}\|_2 < \frac{\sigma^2 \sigma_{\max}(\Phi)}{4e^{4/\sigma^2}\sqrt{M}}\left(\sqrt{1 + \frac{\sigma^2_{\min}(\Phi)}{\sigma^2_{\max}(\Phi)}} - 1\right). \tag{11}$$

Then the error between the true weights $\mathbf{a}^*$ and the perturbed weights $\tilde{\mathbf{a}}$ obtained by solving program (10) is bounded by:

$$\|\mathbf{a}^* - \tilde{\mathbf{a}}\|_2 \leq \frac{4e^{\frac{4}{\sigma^2}}\sqrt{M}\|\mathbf{a}^*\|_2}{\sigma^2 \sigma_{\min}(\Phi)}\|\tilde{\mathbf{t}} - \mathbf{t}^*\|_2 + O(\|\tilde{\mathbf{t}} - \mathbf{t}^*\|_2^2).$$

III. Proofs

In this section we present the proofs of the two theorems. Due to space limitations, we skip some details, which will be present in the journal version of the paper.

A. Proof of Theorem 1

Let t^* be an arbitrary local maximiser of the function $q(t)$ in (3), so t^* is also a source location, and λ^* the solution to (2). The key step in this proof is applying a quantitative version of the Implicit Function Theorem [11] to the function:

$$F(t, \lambda) = \sum_{j=1}^M \lambda_j \phi'(t - s_j), \tag{12}$$

where $F(t^*, \lambda^*) = 0$ because t^* is a maximizer of $q(s)$ in (3). The theorem tells us that we can express t as a function $t(\lambda)$ of λ with:

$$\partial_\lambda t(\lambda) = -\left[\partial_t F(t(\lambda), \lambda)\right]^{-1} \partial_\lambda F(t(\lambda), \lambda), \tag{13}$$

for t in a ball of radius δ_0 around t^* and for λ in a ball of radius $\delta_1 \leq \delta_0$ around λ^*, where δ_0 is chosen such that

$$\sup_{(t,\lambda) \in V_\delta} \left\| I - \left[\partial_t F(t^*, \lambda^*)\right]^{-1} \partial_t F(t, \lambda) \right\| \leq \frac{1}{2}, \tag{14}$$

where $V_\delta = \left\{(t,\lambda) \in \mathbb{R}^{M+1} : |t - t^*| \leq \delta_0, \|\lambda - \lambda^*\| \leq \delta_0\right\}$ and δ_1 is given by

$$\delta_1 = (2M_t B_\lambda)^{-1}\delta_0, \tag{15}$$

where

$$B_\lambda = \sup_{(t,\lambda)\in V_\lambda} \|\partial_\lambda F(t,\lambda)\|_2,$$
$$M_t = \left\|\partial_t F(t^*,\lambda^*)^{-1}\right\|_2.$$

The following two lemmas give values of δ_0 and δ_1 that define balls around t^* and λ^* respectively which are included in the balls required by the Quantitative Implicit Function Theorem with radii defined in (14) and (15).

Lemma 1: **(Radius of ball around** t^***)** The condition (14) is satisfied if

$$\delta_0 = \frac{\sigma^2 |q''(t^*)|}{\sqrt{M}\left(4 + 2c \cdot \frac{\|\lambda^*\|_2}{\sigma}\right)}.$$

Lemma 2: **(Radius of ball around** λ^***)** For δ_0 from Lemma 1 and δ_1 from condition (15), we have that $\delta_\lambda < \delta_1$ if

$$\delta_\lambda = \frac{\sigma\sqrt{e}|q''(t^*)|}{2\sqrt{2M}} \cdot \delta_0.$$

Given the definition of the function F in (12), we have that

$$\partial_t F(t,\lambda) = \sum_{j=1}^{m} \lambda_j \phi''(t - s_j),$$
$$\partial_\lambda F(t,\lambda) = [\phi'(t - s_1), \quad \dots \quad, \phi'(t - s_M)]^T.$$

By applying a Taylor expansion to $t(\lambda)$ around λ^* in the region defined by δ_0 and δ_λ, we have that

$$t(\lambda) = t(\lambda^*) + \left\langle \lambda - \lambda^*, \partial_\lambda t(\lambda_\delta)\right\rangle,$$

for some λ_δ on the line segment determined by λ^* and λ, so

$$\left|t(\lambda) - t(\lambda^*)\right| \leq \|\lambda - \lambda^*\|_2 \cdot \left\|\partial_\lambda t(\lambda_\delta)\right\|_2. \tag{16}$$

Bounding the second factor in (16) above, using (13) and applying a number of manipulations, we obtain

$$\delta_t \leq \frac{C + \delta_t A}{B} \cdot \delta_\lambda. \tag{17}$$

where $\delta_t = |t(\lambda) - t(\lambda^*)|$, $\|\lambda - \lambda^*\|_2 \leq \delta_\lambda$ and

$$A = \left\|\left[\phi''(t(\lambda) - s_j)\right]_{j=1}^M\right\|_2,$$
$$B = \left|\sum_{j=1}^{M} \lambda_j^* \phi''(t(\lambda) - s_j)\right|,$$
$$C = \left\|\left[\phi'(t(\lambda_\delta) - s_j)\right]_{j=1}^M\right\|_2.$$

We now need to lower bound B and upper bound $C + \delta_t A$. For B, by applying a Taylor expansion around $t(\lambda^*) - s_j$, each term in the sum is equal to

$$\lambda_j^* \phi''(t(\lambda^*) - s_j) + \left(t(\lambda) - t(\lambda^*)\right)\lambda_j^* \phi'''(\xi_j),$$

for $j = 1, \dots, M$ and some ξ_j in the interval

$$\left[t(\lambda^*) - s_j - |t(\lambda) - t(\lambda^*)|, t(\lambda^*) - s_j + |t(\lambda) - t(\lambda^*)|\right].$$

By combining this with the reverse triangle inequality, Cauchy-Schwartz inequality and Lemma 1, we obtain

$$B \geq |q''(t^*)| \left[1 - \frac{c\|\lambda^*\|_2}{4\sigma + 2c\|\lambda^*\|_2}\right],$$

where $c \approx 3.9036$. By using the global maximum of $\phi'(t)$ and $\phi''(t)$ for $\phi(t)$ Gaussian, we have that

$$A \leq \frac{2\sqrt{M}}{\sigma^2} \quad \text{and} \quad C \leq \frac{\sqrt{2M}}{\sigma\sqrt{e}}.$$

By finally plugging the above inequalities and the result of Lemma 1 into (17), we obtain the expression for C_{t^*} in (6).

B. Proof of Theorem 2

We apply equation (4.2) from [5] with $e = 0$ (the noise in the observations) and obtain

$$\tilde{\mathbf{a}} = \mathbf{a}^* - \Phi^\dagger E \mathbf{a}^* - F^T E \mathbf{a}^*, \tag{18}$$

where $\Phi^\dagger = (\Phi^T\Phi)^{-1}\Phi^T$ is the pseudoinverse of Φ and $F = O(E)$ is the perturbation of the $\Phi^\dagger$ due to the perturbation E of Φ, namely

$$\tilde{\Phi}^\dagger = \Phi^\dagger + F^T.$$

In order to obtain an explicit expression for F, we write $\tilde{\Phi}^\dagger$:

$$\begin{aligned}\tilde{\Phi}^\dagger &= (\tilde{\Phi}^T\tilde{\Phi})^{-1}\tilde{\Phi}^T \\ &= \left[(\Phi + E)^T(\Phi + E)\right]^{-1}(\Phi + E)^T \qquad \text{by (9)} \\ &= (\Phi^T\Phi + \Delta)^{-1}(\Phi^T + E^T),\end{aligned} \tag{19}$$

where

$$\Delta = E^T\Phi + \Phi^T E + E^T E \in \mathbb{R}^{K\times K}. \tag{20}$$

We then write the first factor in (19) as

$$\begin{aligned}(\Phi^T\Phi + \Delta)^{-1} &= \left[\Phi^T\left(I + {\Phi^\dagger}^T \Delta \Phi^\dagger\right)\Phi\right]^{-1} \\ &= \Phi^\dagger \left[I + \sum_{k=1}^{\infty}(-1)^k \left({\Phi^\dagger}^T\Delta\Phi^\dagger\right)^k\right]{\Phi^\dagger}^T \\ &= (\Phi^T\Phi)^{-1} + S_\Phi,\end{aligned} \tag{21}$$

where

$$S_\Phi = \Phi^\dagger \left[\sum_{k=1}^{\infty}(-1)^k\left({\Phi^\dagger}^T\Delta\Phi^\dagger\right)^k\right]{\Phi^\dagger}^T \in \mathbb{R}^{K\times K}, \tag{22}$$

and in the second inequality in (21) we applied the Neumann series expansion to the matrix $-{\Phi^\dagger}^T\Delta\Phi^\dagger$, which converges if

$$\| - {\Phi^\dagger}^T\Delta\Phi^\dagger\|_2 < 1. \tag{23}$$

We will return to condition (23) at the end of this section. We now substitute (21) in (19), giving

$$\begin{aligned}\tilde{\Phi}^\dagger &= \left[(\Phi^T\Phi)^{-1} + S_\Phi\right](\Phi^T + E^T) \\ &= \Phi^\dagger + (\Phi^T\Phi)^{-1}E^T + S_\Phi\Phi^T + S_\Phi E^T,\end{aligned}$$

so we have that

$$F^T = (\Phi^T\Phi)^{-1}E^T + S_\Phi \Phi^T + S_\Phi E^T, \quad (24)$$

which is indeed $O(E)$, since $S_\Phi = O(\Delta)$ and $\Delta = O(E)$. We next upper bound $\|S_\Phi\|_2$. From (22) we have

$$\|S_\Phi\|_2 \leq \|\Phi^\dagger\|_2^2 \sum_{k=1}^{\infty} \|\Phi^\dagger\|_2^{2k} \|\Delta\|_2^k. \quad (25)$$

Now let D be an upper bound on $\|\Delta\|_2$, obtained by applying the triangle inequality in (20), so that

$$\|\Delta\|_2 \leq D = 2\|E\|_2\|\Phi\|_2 + \|E\|_2^2. \quad (26)$$

Then, from (25) we have

$$\|S_\Phi\|_2 \leq \|\Phi^\dagger\|_2^2 \sum_{k=1}^{\infty} \|\Phi^\dagger\|_2^{2k} D^k = \frac{D\|\Phi^\dagger\|_2^4}{1 - D\|\Phi^\dagger\|_2^2}, \quad (27)$$

where the series converges if $D\|\Phi^\dagger\|_2^2 < 1$, in which case the denominator in the last fraction above is positive. We return to this condition at the end of the section. We also know that[2]

$$\|\Phi^\dagger\|_2 = \frac{1}{\sigma_{\min}(\Phi)}. \quad (28)$$

By applying the triangle inequality in (24) and then using (27) and the fact that $\|(\Phi^T\Phi)^{-1}\|_2 = 1/\sigma_{\min}^2(\Phi) = \|\Phi^\dagger\|_2^2$ (from (28)), we obtain

$$\|F\|_2 \leq \|E\|_2\|\Phi^\dagger\|_2^2 + \frac{D\|\Phi^\dagger\|_2^4}{1 - D\|\Phi^\dagger\|_2^2}\left(\|\Phi\|_2 + \|E\|_2\right), \quad (29)$$

where D is given in (26). It remains to establish an upper bound on $\|E\|_F$, and consequently on $\|E\|_2$. The following lemma gives us such a bound.

Lemma 3: **(Upper bound of $\|E\|_F$)** Let $E = \tilde{\Phi} - \Phi$ for Φ and $\tilde{\Phi}$ as defined in (8) and (9) respectively for $t_j, \tilde{t}_j \in [0, 1]$ for $j = 1, \dots, K$. Then:

$$\|E\|_F \leq \frac{4e^{\frac{4}{\sigma^2}}\sqrt{M}}{\sigma^2}\|\tilde{\mathbf{t}} - \mathbf{t}^*\|_2. \quad (30)$$

By using the triangle inequality and norm sub-multiplicativity in (18), and then substituting (29) and (30), we obtain

$$\begin{aligned}\|\mathbf{a}^* - \tilde{\mathbf{a}}\|_2 &\leq \|E\|_2\|\Phi^\dagger\|_2\|\mathbf{a}^*\|_2 + \|E\|_2^2\|\Phi^\dagger\|_2^2\|\mathbf{a}^*\|_2 \\ &\quad + \frac{\|E\|_2 D\|\Phi^\dagger\|_2^4}{1 - D\|\Phi^\dagger\|_2^2}(\|\Phi\|_2 + \|E\|_2)\|\mathbf{a}^*\|_2 \\ &\leq \frac{4e^{\frac{4}{\sigma^2}}\sqrt{M}\|\mathbf{a}^*\|_2}{\sigma^2\sigma_{\min}(\Phi)}\|\tilde{\mathbf{t}} - \mathbf{t}^*\|_2 + O(\|\tilde{\mathbf{t}} - \mathbf{t}^*\|_2^2),\end{aligned}$$

which is the bound given in Theorem 2. Note that because $\|E\|_2 = O(\|\tilde{\mathbf{t}} - \mathbf{t}^*\|_2)$ (see (30)), the first term is the only term that is $O(\|\tilde{\mathbf{t}} - \mathbf{t}^*\|_2)$ in the first inequality above, so the other terms are included in the $O(\|\tilde{\mathbf{t}} - \mathbf{t}^*\|_2^2)$ term at the end.

[2] Using the SVD $\Phi = U\Sigma V^T$, we have $\Phi^\dagger = (\Phi^T\Phi)^{-1}\Phi^T = (V\Sigma^2V^T)^{-1}V\Sigma U^T = V\Sigma^{-1}U^T$, so the conclusion follows.

Lastly, we return to condition (23), which must be satisfied in order for the bound above to hold. By using norm sub-multiplicativity and the bound on $\|\Delta\|_2$ from (26), we obtain

$$\|{\Phi^\dagger}^T\Delta\Phi^\dagger\|_2 \leq \|\Phi^\dagger\|_2^2\|E\|_2^2 + 2\|\Phi\|_2\|\Phi^\dagger\|_2^2\|E\|_2 \quad (31)$$

and by requiring that the right hand side above is less than one, we obtain a quadratic constraint on $\|E\|_2$, satisfied if

$$\|E\|_2 < \sigma_{\max}(\Phi)\left(\sqrt{1 + \frac{\sigma_{\min}^2(\Phi)}{\sigma_{\max}^2(\Phi)}} - 1\right).$$

By using the bound on $\|E\|_2$ from (30), the above holds if (11) holds. Note that by imposing this, we also ensure that the condition for the series in (27) to converge holds, since $D\|\Phi^\dagger\|_2^2$ is equal to the right hand side of (31).

IV. Conclusion

In this paper, we proved primal stability in the non-negative super-resolution problem, when addressed via convex duality. The main ingredient in our analysis is a quantitative version of the implicit function theorem, a folklore result in the theory of dynamical systems community.

Our results provide precise orders in the number of measurements for the accuracy of the solution to the convex dual problem and an ℓ_∞ error bound on the primal spike locations.

Future plans include the study of the dual approach to the noisy super-resolution problem using similar techniques to the ones developed here.

References

[1] Emmanuel J. Candès and Carlos Fernandez-Granda, "Towards a mathematical theory of super-resolution", Communications on Pure and Applied Mathematics, 67(6), 906–956 (2014)

[2] Geoffrey Schiebinger, Elina Robeva and Benjamin Recht, "Superresolution without separation", Information and Inference: A Journal of the IMA, 7(1), 1–30 (2018)

[3] Nicholas Boyd, Geoffrey Schiebinger and Benjamin Recht, "The alternating descent conditional gradient method for sparse inverse problems", SIAM Journal on Optimization, 27(2), 616–639 (2017)

[4] Yurii Nesterov, "Introductory Lectures on Convex Optimization: A Basic Course", Springer Publishing Company (2004)

[5] G. W. Stewart, "Perturbation theory and least squares with errors in the variables", Contemporary Mathematics 112: Statistical Analysis of Measurement Error Models and Applications, American Mathematical Society, 171–181 (1990)

[6] Marco López and Georg Still, "Semi-infinite programming", European Journal of Operational Research, 180(2), 491–518 (2007)

[7] Armin Eftekhari and Andrew Thompson, "Sparse inverse problems over measures: equivalence of the conditional gradient and exchange methods", To appear in SIAM Journal on Optimization, arXiv: https://arxiv.org/abs/1804.10243 (2019)

[8] Eric Betzig, George H. Patterson, Rachid Sougrat, O. Wolf Lindwasser, Scott Olenych, Juan S. Bonifacino, Michael W. Davidson, Jennifer Lippincott-Schwartz and Harald F. Hes "Imaging intracellular fluorescent proteins at nanometer resolution", Science, 313(5793), 1642–1645 (2006)

[9] Klaus G. Puschmann and Franz Kneer, "On super-resolution in astronomical imaging", Astronomy & Astrophysics, 436(1), 373–378 (2005)

[10] Ronen Tur, Yonina C. Eldar and Zvi Friedman, "Innovation rate sampling of pulse streams with application to ultrasound imaging", IEEE Transactions on Signal Processing, 59(4), 1827–1842 (2011)

[11] Carlangelo Liverani, "Implicit function theorem (a quantitative version)", retrieved January 13, 2019, from https://www.mat.uniroma2.it/ liverani/SysDyn15/app1.pdf