



HAL
open science

Convergence Guarantees for Unmixing PSFs over a Manifold with Non-Convex Optimization

Santos Michelena, Maxime Ferreira Da Costa, José Picheral

► **To cite this version:**

Santos Michelena, Maxime Ferreira Da Costa, José Picheral. Convergence Guarantees for Unmixing PSFs over a Manifold with Non-Convex Optimization. 2025 IEEE Statistical Signal Processing Workshop (SSP), Jul 2025, Edinburgh, United Kingdom. <10.1109/SSP64130.2025.11073360>. <hal-04964463v2>

HAL Id: hal-04964463

<https://centralesupelec.hal.science/hal-04964463v2>

Submitted on 15 Jan 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Copyright - All rights reserved

Convergence Guarantees for Unmixing PSFs over a Manifold with Non-Convex Optimization

Santos Michelena^{*†}, Maxime Ferreira Da Costa^{*}, and José Picheral^{*}

^{*}Laboratory of Signals and Systems, CentraleSupélec, Université Paris–Saclay, CNRS, Gif-sur-Yvette, France

[†]Iumtek, Orsay, France

Abstract—The problem of recovering the parameters of a mixture of spike signals convolved with different PSFs is considered. Herein, the spike support is assumed to be known, while the PSFs lie on a manifold. A non-linear least squares estimator of the mixture parameters is formulated. In the absence of noise, a lower bound on the radius of the strong basin of attraction *i.e.*, the region of convergence, is derived. Key to the analysis is the introduction of coherence and interference functions, which capture the conditioning of the PSF manifold in terms of the minimal separation of the support. Numerical experiments validate the theoretical findings. Finally, the practicality and efficacy of the non-linear least squares approach are showcased on real-world experimental spectral data from laser-induced breakdown spectroscopy.

Index Terms—non-convex optimization; PSF unmixing; manifold constraint; laser-induced breakdown spectroscopy

I. INTRODUCTION

Point spread function (PSF) unmixing is the problem of reconstructing multiple groups of spikes after their convolution with a group-dependent PSF.

Mixtures of sparse signals typically arise whenever the experimenter records a superposition of phenomena produced by different modalities. This parametric signal model is encountered in various areas of applied science and engineering, such as in super-resolution imaging [1], [2], impulse sorting from neural recordings [3], [4], multi-path channel identification in wireless communication [5], [6], or calibration-free spectroscopy [7], among others.

While the blind deconvolution problem [8], [9]—which assumes a *unique* modality—is well-studied in signal processing, the state-of-the-art in resolving sparse mixtures of unknown PSFs is limited. In [10], linear programming is first investigated when the spikes are assumed on a grid. Other approaches relying on the atomic norm minimization framework [11] remove the grid assumption and achieve stable reconstruction of both the signals and the PSFs [12]–[14]. Yet, the PSFs are constrained to lie in a predetermined known low-dimensional subspace, which may be impractical. Furthermore, the computational complexity of convex-based approaches can be prohibitive in real-world applications.

On the other hand, non-convex methods have successfully recovered sparse signals under strong structural assumptions in a scalable manner [15], [16]. When measurements are taken in the Fourier domain, non-convex optimization is shown to deconvolve well-separated sparse signals [17], [18].

In [19], the authors investigate the local geometry of blind deconvolution using non-linear least squares minimization to estimate both the PSF and the signal of interest from a single measurement and provide an estimate for the radius of the strong basin attraction of their method. However, the approach assumes the single modality is in a low-dimensional subspace. Overall, the current techniques are maladapted to many applications, such as laser-induced breakdown spectroscopy (LIBS), where the shapes of the PSFs depend on the complex non-linear physical interactions within the laser-generated plasma [20].

A. Contributions and Organization of the Paper

Motivated by application to LIBS, where accurate estimation of the plasma parameters is the crux to recover the concentration of chemical species [7], we consider the problem of unmixing spikes with *known* support—corresponding to emission wavelengths—and relax the stringent low-dimensional subspace assumption by assuming an ensemble of PSFs lying on a postulated *manifold*. We formulate a non-linear least squares estimator to recover the amplitudes of the spikes and the shape parameters of the PSFs composing the recorded signal. In the absence of noise, a lower bound on the size of the region of strong convexity around the ground truth is established in terms of *minimum separation* of the support, as well as novel *coherence and interference functions* characterizing the conditioning of PSF manifold. Numerical experiments on synthetic and LIBS data showcase our approach’s effectiveness.

The rest of the article is organized as follows. Our mixture recovery problem is introduced in Subsection I-B. Section II recalls the importance of strong basins of attraction in non-convex optimization and presents our key novel coherence and interference metrics for the PSF manifold. A lower bound on the radius of the basin of attraction of the ground truth parameters is established in terms of those metrics in Theorem 1. A sketch proof of Theorem 1 is drawn in Section III. Section IV conducts numerical simulations validating our theory, and presents an application to aluminum concentration estimation on real-world LIBS data. Finally, a conclusion is drawn in Section V.

B. Problem Formulation

We assume a parametric class of PSFs $g(\theta, \cdot)$, where $\theta \in \Theta \subset \mathbb{R}$ is a univariate shape parameter. The investigated

MF acknowledges support from ANR grant ANR-24-CE48-3094. Emails: {santos.michelena, maxime.ferreira, jose.picheral}@centralesupelec.fr.

problem is to recover a signal of the form

$$t \mapsto x(t) = \sum_{i=1}^p \sum_{\ell=1}^{L_i} \eta_{i,\ell}^* g(\theta_i^*, t - t_{i,\ell}) \quad (1)$$

from N samples taken uniformly over an the interval $I = [-\frac{T}{2}, \frac{T}{2}]$, for some $T > 0$, and at the timestamps $\{u_s = -\frac{T}{2} + \frac{T(s-1)}{N-1}\}_{s=1}^N$. In our setting, both the source locations $\{t_{1,1}, \dots, t_{1,L_1}, \dots, t_{p,1}, \dots, t_{p,L_p}\}$ and the parametric class $g(\cdot, \cdot)$ are assumed to be *known*; While the true shape parameters $\{\theta_1^*, \dots, \theta_p^*\}$ and coefficients $\{\eta_{i,\ell}^*\}$ remain *unknown* at the time of observation. Furthermore, the mixture number p and the number of spikes per modality $\{L_1, \dots, L_p\}$ are *known*. Let $\mathbf{g}_{\ell,i} = [g(\theta_i, u_s - t_{i,\ell})]_{s=1}^N$, and define the map $\theta \mapsto \mathbf{G}(\theta)$

$$\mathbf{G}(\theta) = [\mathbf{g}_{1,1} \ \dots \ \mathbf{g}_{1,L_1} \ \dots \ \mathbf{g}_{p,1} \ \dots \ \mathbf{g}_{p,L_p}]. \quad (2)$$

Furthermore, we define

$$\boldsymbol{\eta}^* = [\eta_{1,1}^* \ \dots \ \eta_{1,L_1}^* \ \dots \ \eta_{p,1}^* \ \dots \ \eta_{p,L_p}^*]^\top \in \mathbb{R}^{\bar{L}}, \quad (3)$$

where $\bar{L} = \sum_{i=1}^p L_i$. The noiseless measurements \mathbf{x}^* read

$$\mathbf{x}^* = \mathbf{G}(\boldsymbol{\theta}^*) \boldsymbol{\eta}^*. \quad (4)$$

We study the problem of estimating the parameters $\{\boldsymbol{\theta}^*, \boldsymbol{\eta}^*\}$ in (4) by solving the non-linear least-squares program

$$\arg \min_{\boldsymbol{\theta}, \boldsymbol{\eta}} \left\{ \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta}) = \frac{1}{2} \|\mathbf{G}(\boldsymbol{\theta}) \boldsymbol{\eta} - \mathbf{x}^*\|_2^2 \right\}. \quad (5)$$

For our analysis, we assume $g(\cdot, \cdot)$ is twice continuously differentiable with respect to its first argument θ , meaning that $g(\cdot, t) \in \mathcal{C}^2(\Theta)$ for all $t \in I$, and write $\partial_a g(\cdot, \cdot)$ the a -th derivative with respect to θ . Furthermore, the PSFs and their derivatives up to the second order are assumed both integrable and square-integrable over \mathbb{R} , that is

$$\forall \theta \in \Theta, \forall a \in \{0, 1, 2\}, \quad \partial_a g(\theta, \cdot) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}). \quad (6)$$

Additionally, we require that $g(\theta, \cdot)$ be even for all $\theta \in \Theta$.

II. MAIN RESULTS

A. Strong Basins of Attraction

It is clear that the ground truth parameters $(\boldsymbol{\theta}^*, \boldsymbol{\eta}^*)$ are a solution of (5) as $\mathcal{L}(\boldsymbol{\theta}^*, \boldsymbol{\eta}^*) = 0$. However, given the non-convex nature of our loss function (5), bad local minima may exist. Herein, we investigate the *local convergence* of line search methods such as gradient descent and the method of Gauss-Newton in a neighborhood of $(\boldsymbol{\theta}^*, \boldsymbol{\eta}^*)$. To that end, we seek a characterization of the strong basin of attraction of the loss (5) around the ground truth parameters.

That is, a neighborhood \mathcal{N} of $(\boldsymbol{\theta}^*, \boldsymbol{\eta}^*)$ such that there exist two constants $0 < \xi \leq \gamma < \infty$ such that the loss (5) be ξ -strongly convex and γ -smooth, *i.e.* if $(\boldsymbol{\theta}, \boldsymbol{\eta}) \in \mathcal{N}$ then

$$\forall \mathbf{u}, \xi \|\mathbf{u}\|_2^2 \leq \mathbf{u}^\top \nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta}) \mathbf{u} \leq \gamma \|\mathbf{u}\|_2^2. \quad (7)$$

Line search methods are guaranteed to converge to the global solution $(\boldsymbol{\theta}^*, \boldsymbol{\eta}^*)$ when initialized within the strong basin of attraction \mathcal{N} . However, the rate of convergence varies

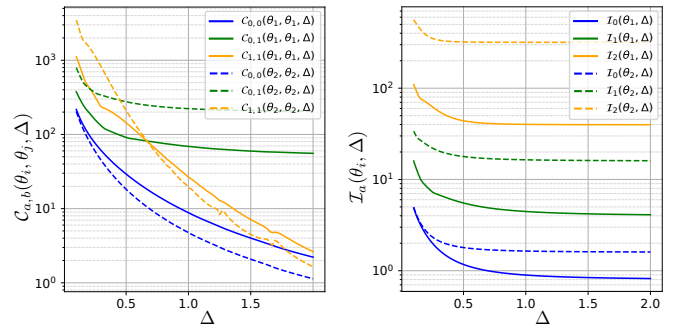


Figure 1. Graphs of the coherence functions (left) and the interference functions (right) for the class of Lorentz PSF $g(\theta, t) = \theta/(\pi(\theta^2 + t^2))$ in the range $\Delta \in [0.1, 2]$ for the parameters $\theta_1 = 0.2$ and $\theta_2 = 0.1$.

depending on the method. For example, it is well established in the literature that the method of steepest descent will converge linearly whenever the step size α is chosen as $\alpha = \gamma/\xi$ (see *e.g.* [21, Theorem 2.1]).

B. Coherence and Interference Functions

In this section, we introduce key metrics to our analysis of the radius of the basin (7). The *minimal separation* of the support Δ is known to affect the well-posedness of sparse inverse problems [22], [23] and is defined in our setting as

$$\Delta := \min_{i,j} \min_{(i,\ell) \neq (j,k)} |t_{i,\ell} - t_{j,k}|. \quad (8)$$

For $a \in \{0, 1, 2\}$, we let the vectors

$$\partial_a \mathbf{g}(\theta, t) := [\partial_a g(\theta, u_s - t)]_{s=1}^N \in \mathbb{R}^N. \quad (9)$$

We introduce the *coherence*, which characterizes the maximum correlation between the dictionary elements and their derivatives when separated by at least Δ , and is defined by

$$\mu_{a,b}(\theta_i, \theta_j, \Delta) := \sup_{|\delta| \geq \Delta} |\partial_a \mathbf{g}(\theta_i, 0)^\top \partial_b \mathbf{g}(\theta_j, \delta)|. \quad (10)$$

The *coherence function* between $\partial_a \mathbf{g}(\theta_i, \cdot)$ and $\partial_b \mathbf{g}(\theta_j, \cdot)$ is

$$\begin{aligned} & \mathcal{C}_{a,b}(\theta_i, \theta_j, \Delta) \\ & := \begin{cases} 2 \sum_{m=1}^{\infty} \mu_{a,b}(\theta_i, \theta_j, m\Delta) & \text{if } i = j \text{ and } a = b \\ 2 \sum_{m=0}^{\infty} \mu_{a,b}(\theta_i, \theta_j, m\Delta) & \text{otherwise.} \end{cases} \end{aligned} \quad (11)$$

Similarly, we define *interference functions* to capture the decay rate of $g(\theta, \cdot)$ and its derivatives when the support is separated by at least Δ , which we define as

$$\mathcal{I}_a(\theta_i, \Delta) := |\partial_a g(\theta_i, 0)| + 2 \sum_{m=1}^{\infty} \sup_{|\delta| \geq m\Delta} |\partial_a g(\theta_i, \delta)|. \quad (12)$$

For illustration purposes, Figure 1 pictures the coherence and interference functions for the class of Lorentz kernels. Of particular importance in the sequel, we remark that the coherence functions $C_{0,0}(\theta_i, \theta_i, \Delta)$ and $C_{1,1}(\theta_i, \theta_i, \Delta)$ tend to 0, while $C_{1,0}(\theta_i, \theta_j, \Delta)$ converges to a constant value as Δ increases. This property is consistent with Equation (10) and $\mu_{a,b}(\theta_i, \theta_j, \Delta) \rightarrow 0$ as $\Delta \rightarrow \infty$. Similarly, the interference functions satisfies $\mathcal{I}_a(\theta_i, \Delta) \rightarrow |\partial_a g(\theta_i, 0)|$ with $\Delta \rightarrow \infty$.

C. Main Result Statement

Denote by $\mathcal{B}_\varepsilon = \{(\boldsymbol{\theta}, \boldsymbol{\eta}) : \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|_\infty + \|\boldsymbol{\eta} - \boldsymbol{\eta}^*\|_\infty < \varepsilon\}$ and let $\eta_{i,\min}^* = \min_\ell \{\eta_{i,\ell}^*\}$ and $\eta_{i,\max}^* = \max_\ell \{\eta_{i,\ell}^*\}$. As our main result, we establish that \mathcal{B}_ε is contained in the strong basin of attraction of the ground truth for the loss (5) whenever ε is small enough. Furthermore, we express the convexity and smoothness parameters (7) as a function of the coherence and interference functions introduced in Section II-B.

Theorem 1. *Assume the maps $\theta \mapsto \mathcal{C}_{a,b}(\theta, \theta', \Delta)$ and $\theta \mapsto \mathcal{I}_a(\theta, \Delta)$ are Lipschitz with constant C_Δ , and the map $\boldsymbol{\theta} \mapsto \mathbf{G}(\boldsymbol{\theta})$ is Lipschitz in the ℓ_∞ -topology with constant K . Let*

$$c_-^* := \frac{1}{2} \min_{i \in \llbracket p \rrbracket} \left\{ \eta_{i,\min}^* \mu_{1,1}(\theta_i^*, \theta_i^*, 0), \mu_{0,0}(\theta_i^*, \theta_i^*, 0) \right\} \quad (13a)$$

$$c_+^* := \frac{3}{2} \max_{i \in \llbracket p \rrbracket} \left\{ \eta_{i,\max}^* \mu_{1,1}(\theta_i^*, \theta_i^*, 0), \mu_{0,0}(\theta_i^*, \theta_i^*, 0) \right\} \quad (13b)$$

$$r^* := \max_{i \in \llbracket p \rrbracket} \left\{ \sum_{j=1}^p \eta_{j,\max}^* \mathcal{C}_{0,1}(\theta_i^*, \theta_j^*, \Delta) + \mathcal{C}_{0,0}(\theta_i^*, \theta_j^*, \Delta), \right. \\ \left. \eta_{i,\max}^* \sum_{j=1}^p \eta_{j,\max}^* \mathcal{C}_{1,1}(\theta_i^*, \theta_j^*, \Delta) + \mathcal{C}_{1,0}(\theta_i^*, \theta_j^*, \Delta) \right\} \quad (13c)$$

$$q^* := 2N \max_{i \in \llbracket p \rrbracket} \left\{ \|\boldsymbol{\eta}_i^*\|_1 \mathcal{I}_2(\theta_i^*, \Delta) + L_i \mathcal{I}_1(\theta_i^*, \Delta) \right. \\ \left. + 2K \|\boldsymbol{\eta}^*\|_\infty + 8pC_\Delta \max\{1, \|\boldsymbol{\eta}^*\|_\infty^2\} \right\}. \quad (13d)$$

Furthermore, assume $r^* < c_-^*$, and select $\varepsilon \in [0, \bar{\varepsilon})$ with $\bar{\varepsilon} = \min \left\{ \frac{c_-^* - r^*}{q^*}, \|\boldsymbol{\eta}^*\|_\infty + \|\mathbf{G}(\boldsymbol{\theta}^*)\|_\infty \right\}$. Then there exists $N_0 \in \mathbb{N}$ such that if $N \geq N_0$ then \mathcal{B}_ε is a strong basin of attraction for the loss (5) with constants

$$\xi = c_-^* - r^* - q^* \varepsilon; \quad \gamma = c_+^* + r^* + q^* \varepsilon;$$

in Equation (7). Hence, any line search method converges given an initialization $(\boldsymbol{\theta}^0, \boldsymbol{\eta}^0) \in \mathcal{B}_{\varepsilon_0}$ with $\varepsilon_0 = \frac{c_-^* - r^*}{q^*}$.

The radius of the convexity region predicted by Theorem 1 is driven by two antagonist factors. On one hand, larger values of c_-^* and c_+^* in Equations (13a) and (13b) widen the basin and are achieved by PSFs with larger energies $\mu_{0,0}(\theta_i^*, \theta_i^*, 0)$, or by PSFs whose first derivatives with respect to θ have larger energies $\mu_{1,1}(\theta_i^*, \theta_i^*, 0)$. On the other hand, larger values of r^* from increased coherence functions or larger values of q^* from increased interference functions shrink the basin. From the definitions (11) and (12), lower coherence and interferences are obtained by increasing Δ , or by considering a parametric class of PSF $g(\theta, t)$ with faster decays when $t \rightarrow \infty$.

III. PROOF ELEMENTS OF THEOREM 1

The proofs of intermediate lemmas are omitted due to space limitation. First, the Hessian $\nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta})$ of the loss (5) writes

$$\nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta}) = \underbrace{\begin{bmatrix} \boldsymbol{\eta}^\top & \mathbf{0}_{1,\bar{L}} \\ \mathbf{0}_{\bar{L},1} & \mathbf{I}_{\bar{L}} \end{bmatrix} \begin{bmatrix} \mathbf{G}_1^\top \mathbf{G}_1 & \mathbf{G}_1^\top \mathbf{G}_0 \\ \mathbf{G}_0^\top \mathbf{G}_1 & \mathbf{G}_0^\top \mathbf{G}_0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta} & \mathbf{0}_{1,\bar{L}} \\ \mathbf{0}_{\bar{L},1} & \mathbf{I}_{\bar{L}} \end{bmatrix}}_{:=\mathbf{E}} \\ + \underbrace{\begin{bmatrix} \mathbf{r}^\top \mathbf{G}_2 \boldsymbol{\eta} & \mathbf{r}^\top \mathbf{G}_1 \\ (\mathbf{G}_1)^\top \mathbf{r} & \mathbf{0}_{\bar{L},\bar{L}} \end{bmatrix}}_{:=\mathbf{R}}, \quad (14)$$

where, for $a \in \{0, 1, 2\}$, we write $\mathbf{G}_a^i := [\partial_a \mathbf{g}_{i,1}, \dots, \partial_a \mathbf{g}_{i,L_i}] \in \mathbb{R}^{N \times L_i}$, and $\mathbf{G}_a = [\mathbf{G}_a^1, \dots, \mathbf{G}_a^p] \in \mathbb{R}^{N \times \bar{L}}$. where $\mathbf{r} = \mathbf{G}(\boldsymbol{\theta})\boldsymbol{\eta} - \mathbf{x}^*$ is the model residual. In the view of (7), we aim to derive values ξ and γ in Theorem 1 from a lower (resp. upper) bound of the spectrum from the expression of the Hessian matrix (14). To proceed, we denote by \mathbf{D} the diagonal part of \mathbf{E} . Since $\nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta})$ is symmetric, its eigenvalues are real. One may bound the spectrum of $\nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta})$ with Weyl's inequality (see e.g. [24, Theorem 8.1.5]) as

$$\lambda_{\min}(\nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta})) \geq \lambda_{\min}(\mathbf{D}) - \|\nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta}) - \mathbf{D}\|_\infty \quad (15a)$$

$$\lambda_{\max}(\nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta})) \leq \lambda_{\max}(\mathbf{D}) + \|\nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta}) - \mathbf{D}\|_\infty \quad (15b)$$

Then, values of ξ and γ can be obtained by lower (resp. upper) bounding the right-hand side of (15a) (resp. (15b)). This is achieved by controlling the three quantities $\lambda_{\max}(\mathbf{D})$, $\lambda_{\min}(\mathbf{D})$ and $\|\nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta}) - \mathbf{D}\|_\infty$ independently as functions of Δ and the error radius ε .

The eigenvalues of \mathbf{D} are controlled with Lemma 2.

Lemma 2. *Under the assumptions of Theorem 1, one has*

$$c_-^* \leq \lambda_{\min}(\mathbf{D}) \leq \lambda_{\max}(\mathbf{D}) \leq c_+^*. \quad (16)$$

Next, the quantity $\|\nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta}) - \mathbf{D}\|_\infty$ is controlled by bounding the ℓ_∞ -norm of each of the sub-block of $\nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta}) - \mathbf{D}$. As the matrices \mathbf{G}_a^i or $\mathbf{G}_a^{i^\top} \mathbf{G}_b^j$ for $a, b \in \{0, 1\}$, and for all $i, j \in \{1, \dots, p\}$ are ubiquitous in the block decomposition, it is essential to control their ℓ_∞ -norm. This is achieved in Lemma 3 by harnessing that for any $m \in \mathbb{N} \setminus \{0\}$, there exist at most two support points $\{t_{i,\ell}\}_{i,j}$ such that $(m-1)\Delta \leq |u_s - t_{i,\ell}| < m\Delta$, thus producing simple inequalities involving the coherence and interference functions.

Lemma 3. *Under the assumptions of Theorem 1, one has*

$$\|\mathbf{G}_a^i\|_\infty \leq \mathcal{I}_a(\theta_i^*, \Delta) + C_\Delta \varepsilon, \quad (17a)$$

$$\|\mathbf{G}_a^{i^\top} \mathbf{G}_b^j\|_\infty \leq \mathcal{C}_{a,b}(\theta_i^*, \theta_j^*, \Delta) + 2C_\Delta \varepsilon. \quad (17b)$$

Furthermore, from the Lipschitz assumption on the manifold map $\boldsymbol{\theta} \mapsto \mathbf{G}(\boldsymbol{\theta})$, one can control the residual error as $\|\mathbf{r}\|_\infty \leq (\|\mathbf{G}(\boldsymbol{\theta}^*)\|_\infty + K \|\boldsymbol{\eta}^*\|_\infty) \varepsilon$. Thus, by inspection of the block structure, it can be inferred with the triangle inequality that

$$\|\nabla^2 \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta}) - \mathbf{D}\|_\infty \leq r^* + q^* \varepsilon. \quad (18)$$

Substituting (16) and (18) into (15) yield the main result. ■

IV. NUMERICAL EXPERIMENTS

A. Validation of Theorem 1

In this section, we numerically validate Theorem 1. To this end, we reconstruct $p = 2$ modalities within the class of Lorentz PSFs $g(\theta, t) = \theta / (\pi(\theta^2 + t^2))$. The ground truth PSF parameters are $\theta_1^* = 2 \cdot 10^{-5}$ and $\theta_2^* = 10^{-3}$. The model orders for each modality are $L_1 = 10$ and $L_2 = 5$ with linear argument $\{\eta_{1,\ell}^* = \frac{1}{L_1}\}_{\ell=1}^{L_1}$, $\{\eta_{2,\ell}^* = \frac{1}{L_2}\}_{\ell=1}^{L_2}$.

Since $p = 2$, the restricted loss $\boldsymbol{\theta} \mapsto \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\eta}^*)$ can be visualized as a contour plot, which is pictured in Figure 2 for three different values of Δ . The lower bound on the radius of

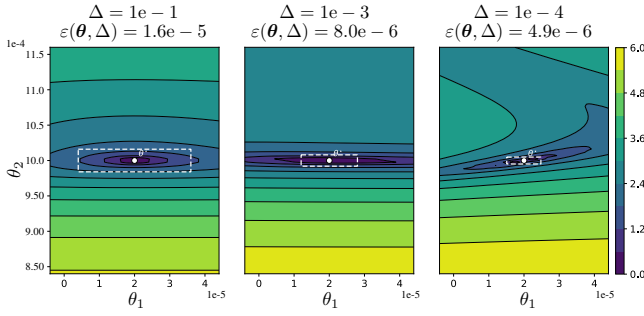


Figure 2. The level sets of the restricted map $\theta \mapsto \mathcal{L}(\theta, \eta)$, and the strong basin radii (in white/dashed) predicted by Theorem 1, for decreasing separation parameter Δ .

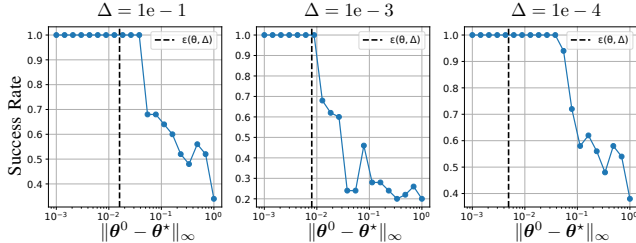


Figure 3. Success rate of SCIPY's least squares solver to recover the ground truth θ^* as a function of of the ℓ_∞ -distance between the initial guess θ^0 and the ground truth θ^* . The results are averaged over 50 trials. The lower bound on the radius of the basin predicted by Theorem 1 is shown in a dashed vertical line.

the basin of attraction guaranteed by Theorem 1 is indicated. It can be appreciated that the radius worsens as Δ decreases, resulting in a progressively smaller convergence region.

In a second experiment, we challenge the radius ε_0 predicted by Theorem 1 through Monte Carlo simulations. Specifically, we randomize an initialization θ^0 and estimate a nonlinear solver's success rate of recovering θ^* as a function of $\|\theta^* - \theta^0\|_\infty$. The results are reported in Figure 3. The Monte Carlo simulations converge with empirical probability 1 below the threshold predicted by Theorem 1, which corroborates our result. We highlight that for $\Delta = 10^{-3}$, our theoretical bound on the basin size is sharp, whereas, for smaller Δ , the bound becomes more pessimistic, owing to the extremely non-convex nature of the problem in that case.

B. Application to Laser-induced Breakdown Spectroscopy

As a final experiment, we wish to assess the performance of our method on spectroscopical data produced by a LIBS instrument. A spectrum comprises a set of spectral lines with fixed emission wavelengths. Those emission wavelengths are physical constants; hence, they are known and given for each chemical species. However, during acquisition, spectral lines become broadened by various physical interactions in the plasma and vary with the experimental setup, the chemical composition of the sample, and the species. The overarching goal of the analysis of spectroscopical data is recovering the concentration of each element/ion pair present in the sample from the observation of the emission spectrum.

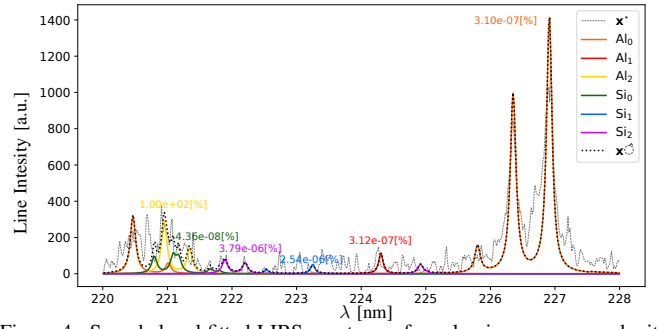


Figure 4. Sampled and fitted LIBS spectrum of an aluminum compound with corresponding ions indicated alongside their estimated concentrations.

We model the emission spectrum measurement as in Subsection I-B where p is the number of chemical species to identify in the measured chemical compound, and θ_i is the shape parameter of the i -th species, for $i \in \{1, \dots, p\}$. The species concentration vector ν^* can be modeled as a linear function of the emission amplitudes $\eta^* = \mathbf{A}\nu^*$, where \mathbf{A} is a matrix which can be estimated a priori.

We deploy our numerical method on a calibrated metallic sample composed of 97.686% aluminum and 2.313% trace elements, with silicon being the most abundant (0.896%). We aim to estimate the concentration of each aluminum element/ion pair. Since the measurement is performed using a LIBS instrument, the observation model is given by

$$\mathbf{x}_{\text{obs}} = \Xi(\mathbf{G}(\theta^*)\eta^* + \mathbf{b} + \xi), \quad (19)$$

where Ξ is the apparatus' transfer function, \mathbf{b} is the spectrum's baseline emission [25] and ξ is a noise term, which we have observed follows a χ^2 distribution in practice. In view of (1), we first perform a preprocessing step to remove the affine term $\Xi\mathbf{b}$, resulting in the modified observation $\tilde{\mathbf{x}} = \mathbf{x}_{\text{obs}} - \Xi\mathbf{b}$. We then construct \mathbf{G} using emission line location information from the NIST database for atomic spectra [26] and solve

$$\arg \min_{\theta, \eta} \left\{ \mathcal{L}(\theta, \eta) = \frac{1}{2} \|\Xi\mathbf{G}(\theta)\eta - \tilde{\mathbf{x}}\|_2^2 \right\}. \quad (20)$$

Figure 4 shows the observed and reconstructed emission spectrum, with a fit error of 15.33%, while the relative estimation error in the aluminum concentration is 0.023%.

V. CONCLUSION

In this paper, we formulated a non-convex least square problem to unmix sparse signals with known support that have been convolved with multiple PSFs lying in a univariate manifold. In Theorem 1, we have provided a lower bound on the radius of the strong basin of attraction of our numerical method in the noiseless case. Immediate extensions include extending the analysis to the presence of noise and considering recovery over a higher dimensional PSF manifold. Furthermore, harnessing sparsity constraints into η is a research direction of practical interest to the analysis of spectroscopic data for the purpose of estimating the concentration of chemical species in a compound when its composition is unknown a priori.

REFERENCES

- [1] B. Huang, W. Wang, M. Bates, and X. Zhuang, "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy," *Science*, vol. 319, no. 5864, pp. 810–813, 2008.
- [2] Y. Huang, J. Li, X. Gao, L. He, and W. Lu, "Single image super-resolution via multiple mixture prior models," *IEEE Transactions on Image Processing*, vol. 27, no. 12, pp. 5904–5917, 2018.
- [3] K. C. Knudson, J. Yates, A. Huk, and J. W. Pillow, "Inferring sparse representations of continuous signals with continuous orthogonal matching pursuit," *Advances in neural information processing systems*, vol. 27, 2014.
- [4] Y. Li, Z. L. Yu, N. Bi, Y. Xu, Z. Gu, and S.-i. Amari, "Sparse representation for brain signal processing: A tutorial on methods and applications," *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 96–106, 2014.
- [5] L. Applebaum, W. U. Bajwa, M. F. Duarte, and R. Calderbank, "Asynchronous code-division random access using convex optimization," *Physical Communication*, vol. 5, no. 2, pp. 129–147, 2012.
- [6] Y. Xie, Y. Chi, L. Applebaum, and R. Calderbank, "Compressive demodulation of mutually interfering signals," in *2012 IEEE Statistical Signal Processing Workshop (SSP)*, 2012, pp. 592–595.
- [7] Z. Hu, D. Zhang, W. Wang, *et al.*, "A review of calibration-free laser-induced breakdown spectroscopy," *TrAC Trends in Analytical Chemistry*, vol. 152, p. 116 618, 2022.
- [8] Y. Li, K. Lee, and Y. Bresler, "Identifiability in blind deconvolution with subspace or sparsity constraints," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 4266–4275, 2016.
- [9] Y. Li and Y. Bresler, "Multichannel sparse blind deconvolution on the sphere," *IEEE Transactions on Information Theory*, vol. 65, no. 11, pp. 7415–7436, 2019.
- [10] J. Romberg and R. Neelamani, "Sparse channel separation using random probes," *Inverse Problems*, vol. 26, no. 11, p. 115 015, 2010.
- [11] Y. Chi and M. Ferreira Da Costa, "Harnessing sparsity over the continuum: Atomic norm minimization for superresolution," *IEEE Signal Processing Magazine*, vol. 37, no. 2, pp. 39–57, 2020.
- [12] Y. Chi, "Guaranteed blind sparse spikes deconvolution via lifting and convex optimization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 4, pp. 782–794, Jun. 2016.
- [13] S. Li, M. B. Wakin, and G. Tang, "Atomic norm denoising for complex exponentials with unknown waveform modulations," *IEEE Transactions on Information Theory*, vol. 66, no. 6, pp. 3893–3913, 2020.
- [14] S. Chrétien and H. Tyagi, "Multi-kernel unmixing and super-resolution using the modified matrix pencil method," *Journal of Fourier Analysis and Applications*, vol. 26, no. 1, p. 18, 2020.
- [15] Y. Traonmilin and J.-F. Aujol, "The basins of attraction of the global minimizers of the non-convex sparse spike estimation problem," *Inverse Problems*, vol. 36, no. 4, p. 045 003, 2020.
- [16] Y. Traonmilin, J.-F. Aujol, P.-J. Bénéard, and A. Leclaire, "On strong basins of attractions for non-convex sparse spike estimation: Upper and lower bounds," *Journal of Mathematical Imaging and Vision*, vol. 66, no. 1, pp. 57–74, 2024.
- [17] M. Ferreira Da Costa and Y. Chi, "Local geometry of nonconvex spike deconvolution from low-pass measurements," *IEEE Journal on Selected Areas in Information Theory*, vol. 4, pp. 1–15, 2023.
- [18] J. Gabet, M. Kalra, M. Ferreira Da Costa, and K. Lee, "Global convergence of ESPRIT with preconditioned first-order methods for spike deconvolution," *arXiv preprint arXiv:2502.08035*, 2025.
- [19] X. Li, S. Ling, T. Strohmer, and K. Wei, "Rapid, robust, and reliable blind deconvolution via nonconvex optimization," *Applied and Computational Harmonic Analysis*, vol. 47, no. 3, pp. 893–934, 2019.
- [20] P. Hansen, S. Schröder, S. Kubitza, K. Rammelkamp, D. Vogt, and H.-W. Hübers, "Modeling of time-resolved libs spectra obtained in martian atmospheric conditions with a stationary plasma approach," *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 178, p. 106 115, 2021.
- [21] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd. New York, NY, USA: Springer, 2006.
- [22] M. Ferreira Da Costa and Y. Chi, "On the stable resolution limit of total variation regularization for spike deconvolution," *IEEE Transactions on Information Theory*, vol. 66, no. 11, pp. 7237–7252, 2020.
- [23] M. Hockmann and S. Kunis, "Weak sparse superresolution is well-conditioned," *SIAM Journal on Imaging Sciences*, vol. 16, no. 1, SC1–SC13, 2023.
- [24] G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins Studies in the Mathematical Sciences), Fourth edition. Baltimore: The Johns Hopkins University Press, 2013, 756 pp.
- [25] V. Mazet, C. Carteret, D. Brie, J. Idier, and B. Humbert, "Background removal from spectra by designing and minimising a non-quadratic cost function," *Chemometrics and Intelligent Laboratory Systems*, vol. 76, no. 2, pp. 121–133, 2005.
- [26] A. Kramida, Y. Ralchenko, J. Reader, and NIST ASD Team, *NIST Atomic Spectra Database (ver. 5.10)*, Online, National Institute of Standards and Technology, Available: <https://physics.nist.gov/asd>, 2025.