

PENDANTSS: PEnalized Norm-ratios Disentangling Additive Noise, Trend and Sparse Spikes

Paul Zheng, *Student Member, IEEE*, Emilie Chouzenoux, *Senior Member, IEEE*, Laurent Duval, *Senior Member, IEEE*

Abstract—Denoising, detrending, deconvolution: usual restoration tasks, traditionally decoupled. Coupled formulations entail complex ill-posed inverse problems. We propose PENDANTSS for joint trend removal and blind deconvolution of sparse peak-like signals. It blends a parsimonious prior with the hypothesis that smooth trend and noise can somewhat be separated by low-pass filtering. We combine the generalized quasi-norm ratio SOOT/SPOQ¹ sparse penalties ℓ_p/ℓ_q with the BEADS² ternary-assisted source separation algorithm. This results in a both convergent and efficient tool, with a novel Trust-Region block alternating variable metric forward-backward approach. It outperforms comparable methods, when applied to typically peaked analytical chemistry signals. Reproducible code is provided.

Index Terms—Blind deconvolution, sparse signal, trend estimation, non-convex optimization, forward-backward splitting, alternating minimization, source separation

I. INTRODUCTION AND BACKGROUND

Restoration recovers information from observations with amplitude distortion, level displacement or random disturbance. We seek estimates \hat{s} , \hat{t} and $\hat{\pi}$ from observation \mathbf{y} , under the discrete additive-convolutive degradation:

$$\mathbf{y} = \bar{s} * \bar{\pi} + \bar{t} + \mathbf{n}. \quad (1)$$

Among N sample values, a series of *spikes* (also called impulses, events, “diracs” or spectral lines) models the **first component**, the sought sparse signal $\bar{s} \in \mathbb{R}^N$. Its convolution with an unknown short-support *kernel* $\bar{\pi} \in \mathbb{R}^L$ — typically peak-shaped — yields the *peak-signal* $\bar{x} = \bar{s} * \bar{\pi} \in \mathbb{R}^N$. The **second component** $\bar{t} \in \mathbb{R}^N$ offsets the reference level, harming quantitative estimations. It can be called baseline, background, continuum, drift, or wander. We opt for *trend*, a reference above which peaks are detected, evaluated and measured. “Trends” address slowly varying amplitude shifts (due to seasonality, calibration distortion, sensor decline...), challenging its automated removal. **Third component** $\mathbf{n} \in \mathbb{R}^N$ (*noise*) gathers stochastic residuals. Given (1), the goal is to perform jointly denoising, detrending and deconvolution. Namely, given \mathbf{y} , retrieve estimations of the spiky signal, the kernel and the trend. Fig. 1 is reminiscent of standard spectral subtraction [1], and motivated here by peak-signal retrieval in

This work was supported by the European Research Council Starting Grant MAJORIS ERC-2019-STG-850925.

P. Zheng is currently with Chair of Information Theory and Data Analytics, RWTH Aachen University, Germany (paul.zheng@inda.rwth-aachen.de); work conducted while at Univ. Paris-Saclay, CentraleSupélec, CVN, Inria, Gif-sur-Yvette, France.

E. Chouzenoux is with Univ. Paris-Saclay, CentraleSupélec, CVN, Inria, Gif-sur-Yvette, France (emilie.chouzenoux@centralesupelec.fr).

L. Duval is with IFP Energies nouvelles, France (laurent.duval@ifpen.fr).

¹SOOT: Smoothed One-Over-Two / SPOQ: Smoothed p -Over- q .

²BEADS: Baseline Estimation And Denoising with Sparsity.

separative analytical chemistry (AC): chromatography, spectrometry, spectroscopy [2], where peak localization, amplitude, width or area provide useful chemical quantitative information.

Whether acquired in its natural domain [3] or after sparsification [4], noise/trend/spike models (1) cover many multidimensional issues: signal (1D), image (2D), video, volume (3D+). We focus here on 1D data common to diverse domains: Fourier spectral analysis, econometrics, stock prices, biomedical measurements (ECG, EEG, EMG), environmental observations, astronomical spectroscopy, etc.

On the one hand, joint denoising and detrending is a long-standing preprocessing question, ranging from time series analysis to imaging. Background issues are commonly solved using a host of filling, fitting and filtering methods. We refer to overviews in [5], [6], and for AC to background corrections backcor [7] and BEADS [8].

On the other hand, joint denoising and blind deconvolution matters from channel estimation in communications [9] to image deblurring [10]. We refer to [11], [12], and especially emphasize on sparsity-promoting methods like SOOT [13] and SPOQ [14], using smoothed “scale-invariant” norm ratios.

PENDANTSS original contributions are (i) a fully coupled and solvable non-convex formulation for (1) (Section II) and (ii) a novel efficient joint disentangling algorithm (forward-backward-based [15], [16]) with proved convergence (Section III), validated by its comparative performance (Section IV).

II. PROPOSED PROBLEM FORMULATION

A. BEADS peak/trend/noise separation paradigm

Estimates of $(\hat{s}, \hat{t}, \hat{\pi})$ of $(\bar{s}, \bar{t}, \bar{\pi})$ are obtained through the resolution of the penalized least squares problem

$$\underset{\substack{s, t \in \mathbb{R}^N \\ \pi \in \mathbb{R}^L}}{\text{minimize}} \frac{1}{2} \|\mathbf{y} - \pi * \mathbf{s} - \mathbf{t}\|^2 + R(\mathbf{s}, \mathbf{t}, \boldsymbol{\pi}), \quad (2)$$

with regularization term R incorporating prior knowledge. Disentangling trend and signal is tedious [17]. As in BEADS [8], we assume that the trend can be recovered from a peakless observation through a low-pass filter \mathbf{L} :

$$\hat{t} = \mathbf{L}(\mathbf{y} - \hat{\pi} * \hat{s}). \quad (3)$$

This motivates the rewriting of the data fidelity term in (2) as:

$$\begin{aligned} (\forall \mathbf{s} \in \mathbb{R}^N)(\forall \boldsymbol{\pi} \in \mathbb{R}^L) \rho(\mathbf{s}, \boldsymbol{\pi}) &= \frac{1}{2} \|\mathbf{y} - \mathbf{L}\mathbf{y} - \mathbf{H}(\boldsymbol{\pi} * \mathbf{s})\|^2 \\ &= \frac{1}{2} \|\mathbf{H}(\mathbf{y} - \boldsymbol{\pi} * \mathbf{s})\|^2, \end{aligned} \quad (4)$$

where $\mathbf{H} = \mathbf{Id}_N - \mathbf{L}$ is a high-pass filter, and \mathbf{Id}_N the identity operator of \mathbb{R}^N . We introduce a regularization term Ψ , promoting signal sparsity. We add two extra terms to constrain

estimates $\hat{\mathbf{s}}$ and $\hat{\boldsymbol{\pi}}$ to sets $C_1 \subset \mathbb{R}^N$ and $C_2 \subset \mathbb{R}^L$ assumed closed, non-empty and convex. The indicator function ι_{C_i} , $i \in \{1, 2\}$ equals zero when the value evaluated belongs to C_i , $+\infty$ otherwise. Optimization problem (2) becomes:

$$\underset{\mathbf{s} \in \mathbb{R}^N, \boldsymbol{\pi} \in \mathbb{R}^L}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{H}(\mathbf{y} - \boldsymbol{\pi} * \mathbf{s})\|^2 + \iota_{C_1}(\mathbf{s}) + \iota_{C_2}(\boldsymbol{\pi}) + \lambda \Psi(\mathbf{s}). \quad (5)$$

The estimated trend can be obtained from (3) with $\hat{\boldsymbol{\pi}}$ and $\hat{\mathbf{s}}$ obtained by (5).

B. SPOQ/SOOT norm/quasi-norm ratio penalties

Tractable penalties for sparsity characterization include homogeneous ℓ_p -norms, quasi-norms (for $0 < p < 1$), or mixed norms. We refer to [12]–[14], [18], [19] and references therein. Ratios of norms are also promising proxies, being scale-invariant [20]. We here promote sparse $\hat{\mathbf{s}}$ through the family of SPOQ norm ratio penalties, introduced in [14], as a generalization to the SOOT ratio [13]. Let $p \in]0, 2[$ and $q \in [2, +\infty[$. Smoothed approximations to the ℓ_p quasi-norm and ℓ_q norm, parameterized by constants $(\alpha, \eta) \in]0, +\infty[^2$ are defined, for every $\mathbf{s} = (s_n)_{1 \leq n \leq N} \in \mathbb{R}^N$, as:

$$\ell_{p,\alpha}(\mathbf{s}) = \left(\sum_{n=1}^N \left((s_n^2 + \alpha^2)^{p/2} - \alpha^p \right) \right)^{1/p}, \quad (6)$$

and

$$\ell_{q,\eta}(\mathbf{s}) = \left(\eta^q + \sum_{n=1}^N |s_n|^q \right)^{1/q}. \quad (7)$$

The non-convex SPOQ penalty is given, for $\beta \in]0, +\infty[$, as:

$$(\forall \mathbf{s} \in \mathbb{R}^N) \quad \Psi(\mathbf{s}) = \log \left(\frac{(\ell_{p,\alpha}^p(\mathbf{s}) + \beta^p)^{1/p}}{\ell_{q,\eta}(\mathbf{s})} \right). \quad (8)$$

Ψ is Lipschitz differentiable on \mathbb{R}^N [14, Prop. 2] and admits $\mathbf{0}_N$ as a local minimizer when [14, Prop. 1]:

$$q > 2, \quad \text{or} \quad q = 2 \quad \text{and} \quad \eta^2 \alpha^{p-2} > \beta^p. \quad (9)$$

Condition (9) is assumed throughout this paper.

III. PROPOSED OPTIMIZATION ALGORITHM

A. Problem structure

The objective function in (5) is the sum of a differentiable function (least squares + SPOQ) and terms acting separately on \mathbf{s} or $\boldsymbol{\pi}$ (i.e., indicator terms). In the differentiable part

$$(\forall \mathbf{s} \in \mathbb{R}^N)(\forall \boldsymbol{\pi} \in \mathbb{R}^L) \quad f(\mathbf{s}, \boldsymbol{\pi}) = \rho(\mathbf{s}, \boldsymbol{\pi}) + \lambda \Psi(\mathbf{s}), \quad (10)$$

with function ρ from (4) quadratic in \mathbf{s} and $\boldsymbol{\pi}$. In particular, for every $\boldsymbol{\pi} \in \mathbb{R}^L$ (resp. $\forall \mathbf{s} \in \mathbb{R}^N$), the gradient $\nabla_{\rho_1}(\cdot, \boldsymbol{\pi})$ (resp. $\nabla_{\rho_2}(\mathbf{s}, \cdot)$) of ρ with respect to its first (resp. second) variable is Lipschitz continuous with constant $\Lambda_1(\boldsymbol{\pi})$ (resp. $\Lambda_2(\mathbf{s})$). As aforementioned, $\nabla \Psi$ is Lipschitz continuous too. The second part of the objective function reads as:

$$(\forall \mathbf{s} \in \mathbb{R}^N)(\forall \boldsymbol{\pi} \in \mathbb{R}^L) \quad g(\mathbf{s}, \boldsymbol{\pi}) = \iota_{C_1}(\mathbf{s}) + \iota_{C_2}(\boldsymbol{\pi}). \quad (11)$$

In a nutshell, Problem (5) amounts to minimizing:

$$(\forall \mathbf{s} \in \mathbb{R}^N)(\forall \boldsymbol{\pi} \in \mathbb{R}^L) \quad \Omega(\mathbf{s}, \boldsymbol{\pi}) = f(\mathbf{s}, \boldsymbol{\pi}) + g(\mathbf{s}, \boldsymbol{\pi}). \quad (12)$$

B. Proposed Trust-Region PENDANTSS algorithm

The structure of (12) suggests a block alternating approach where signal \mathbf{s} and kernel $\boldsymbol{\pi}$ are updated sequentially. We hereby introduce Algorithm 1, that generalizes the BC-VMFB algorithm [16], also used in [13] for blind deconvolution.

Algorithm 1: TR-BC-VMFB for solving (5)

Settings: $K_{\max} > 0$, $\varepsilon > 0$, $\mathcal{I} > 0$, $\theta \in]0, 1[$,
 $(\gamma_{s,k})_{k \in \mathbb{N}} \in [\underline{\gamma}, 2 - \bar{\gamma}]$ and $(\gamma_{\pi,k})_{k \in \mathbb{N}} \in [\underline{\gamma}, 2 - \bar{\gamma}]$ for
some $(\underline{\gamma}, \bar{\gamma}) \in]0, +\infty[^2$, $(p, q) \in]0, 2[\times [2, +\infty[$
satisfying (9), convex sets $(C_1, C_2) \subset \mathbb{R}^N \times \mathbb{R}^L$.

Initialize: $\mathbf{s}_0 \in C_1$, $\boldsymbol{\pi}_0 \in C_2$

for $k = 0, 1, \dots$ **do**

Update of the signal

for $i = 1, \dots, \mathcal{I}$ **do**

Set TR radius $\rho_{k,i}$ using (16) with parameter θ ;

Construct MM metric $\mathbf{A}_{1,\rho_{k,i}}(\mathbf{s}_k, \boldsymbol{\pi}_k)$
using (15);

Find $\mathbf{s}_{k,i} \in C_1$ such that (17) holds.

if $\mathbf{s}_{k,i} \in \bar{\mathcal{B}}_{q,\rho_{k,i}}$ **then**

| Stop loop

end

end

$\mathbf{s}_{k+1} = \mathbf{s}_{k,i}$;

Update of the kernel

Find $\boldsymbol{\pi}_{k+1} \in C_2$ such that (19) holds.

Stopping criterion

if $\|\mathbf{s}_k - \mathbf{s}_{k+1}\| \leq \varepsilon$ or $k \geq K_{\max}$ **then**

| Stop loop

end

end

$(\hat{\mathbf{s}}, \hat{\boldsymbol{\pi}}) = (\mathbf{s}_{k+1}, \boldsymbol{\pi}_{k+1})$ and $\hat{\mathbf{t}}$ given by (3);

Result: $\hat{\mathbf{s}}, \hat{\boldsymbol{\pi}}, \hat{\mathbf{t}}$

1) *Signal update:* Let $k \in \mathbb{N}$ and $(\mathbf{s}_k, \boldsymbol{\pi}_k) \in C_1 \times C_2$. The computation of \mathbf{s}_{k+1} follows one Majoration-Minimization (MM) iteration [21]. First, we build a majorization for $\Omega(\cdot, \boldsymbol{\pi}_k)$ around \mathbf{s}_k . Second, \mathbf{s}_{k+1} is defined as a minimizer to the majorant. In practice, both steps can be approximated for speedup and robustness to numerical errors. As emphasized in [14], [22], we need the majorization to be valid only within a neighborhood of the current iterate. For $\rho \in [0, +\infty[$, the ℓ_q -ball complement set is:

$$\bar{\mathcal{B}}_{q,\rho} = \{ \mathbf{s} = (s_n)_{1 \leq n \leq N} \in \mathbb{R}^N \mid \sum_{n=1}^N |s_n|^q \geq \rho^q \}. \quad (13)$$

From [14, Prop. 2], we can show that

$$\begin{aligned} (\forall \mathbf{s} \in \bar{\mathcal{B}}_{q,\rho} \cap C_1) \quad \Omega(\mathbf{s}, \boldsymbol{\pi}_k) &\leq f(\mathbf{s}_k, \boldsymbol{\pi}_k) \\ &+ (\mathbf{s} - \mathbf{s}_k)^\top \nabla_1 f(\mathbf{s}_k, \boldsymbol{\pi}_k) + \frac{1}{2} \|\mathbf{s} - \mathbf{s}_k\|_{\mathbf{A}_{1,\rho}(\mathbf{s}_k, \boldsymbol{\pi}_k)}^2, \end{aligned} \quad (14)$$

where we define the so-called MM metric as:

$$\begin{aligned} \mathbf{A}_{1,\rho}(\mathbf{s}_k, \boldsymbol{\pi}_k) &= (\Lambda_1(\boldsymbol{\pi}_k) + \lambda \chi_{q,\rho}) \mathbf{Id}_N + \\ &\frac{\lambda}{\ell_{p,\alpha}^p(\mathbf{s}_k) + \beta^p} \text{Diag}((s_{n,k}^2 + \alpha^2)^{p/2-1})_{1 \leq n \leq N}, \end{aligned} \quad (15)$$

with the constant $\chi_{q,\rho} = (q-1)/(\eta^q + \rho^q)^{2/q}$. In (14), $\|\cdot\|_{\mathbf{A}}$ denotes the weighted Euclidean norm related to a symmetric definite positive (SDP) matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$, i.e., $\forall \mathbf{z} \in \mathbb{R}^N$, $\|\mathbf{z}\|_{\mathbf{A}} = (\mathbf{z}^\top \mathbf{A} \mathbf{z})^{1/2}$. Since inequality (14) only holds on a limited region, we introduce a Trust-Region-based (TR) loop [22], [23] to make sure that the minimizer of the majorant is indeed in the validity domain of (14). Namely, we set $\mathcal{I} > 0$, a maximum number of trials of TR approach. For $i \in \{1, \dots, \mathcal{I}\}$, we define the TR radius as:

$$\rho_{k,i} = \begin{cases} \sum_{n=1}^N |s_{n,k}|^q & \text{if } i = 1, \\ \theta \rho_{k,i-1} & \text{if } 2 \leq i \leq \mathcal{I} - 1, \\ 0 & \text{if } i = \mathcal{I}. \end{cases} \quad (16)$$

We compute the associated MM metric $\mathbf{A}_{1,\rho_{k,i}}(\mathbf{s}_k, \boldsymbol{\pi}_k)$ and define $\mathbf{s}_{k,i}$ as a minimizer of the right term in (14). The loop stops whenever $\mathbf{s}_{k,i}$ belongs to $\bar{B}_{q,\rho_{k,i}}$, which is ensured to arise in a finite number of steps according to [14]. There remains to explain how we practically compute $\mathbf{s}_{k,i}$. Depending on the choice for C_1 , the right term in (14) might not have a closed-form minimizer. Actually, as we will show, it appears sufficient for convergence purpose to search for $\mathbf{s}_{k,i} \in C_1$ satisfying the first order optimality conditions:

$$\begin{cases} (\mathbf{s}_{k,i} - \mathbf{s}_k)^\top \nabla_1 f(\mathbf{s}_k, \boldsymbol{\pi}_k) + \gamma_{s,k}^{-1} \|\mathbf{s}_{k,i} - \mathbf{s}_k\|_{\mathbf{A}_{1,\rho_{k,i}}(\mathbf{s}_k, \boldsymbol{\pi}_k)}^2 \leq 0, \\ \|\nabla_1 f(\mathbf{s}_k, \boldsymbol{\pi}_k) + \mathbf{r}_{k,i}^{(1)}\| \leq \kappa_1 \|\mathbf{s}_{k,i} - \mathbf{s}_k\|_{\mathbf{A}_{1,\rho_{k,i}}(\mathbf{s}_k, \boldsymbol{\pi}_k)} \end{cases} \quad (17)$$

for some $\mathbf{r}_{k,i}^{(1)} \in N_{C_1}(\mathbf{s}_{k,i})$ (i.e., the normal cone of C_1 at $\mathbf{s}_{k,i}$ [24]), and some $\kappa_1 > 0$. The existence of such an $\mathbf{s}_{k,i}$ can be shown from [25, Rem. 3.3]. In particular, a minimizer over C_1 of the right term in (14) satisfies (17).

2) *Kernel update*: It follows a similar approach. The main difference is that we do not use the TR loop in that case, as the function to minimize here is simpler. Let $k \in \mathbb{N}$, and $(\mathbf{s}_{k+1}, \boldsymbol{\pi}_k) \in C_1 \times C_2$. By descent lemma,

$$\begin{aligned} (\forall \boldsymbol{\pi} \in C_2) \quad \Omega(\mathbf{s}_{k+1}, \boldsymbol{\pi}) &\leq f(\mathbf{s}_{k+1}, \boldsymbol{\pi}_k) \\ + (\boldsymbol{\pi} - \boldsymbol{\pi}_k)^\top \nabla_2 f(\mathbf{s}_{k+1}, \boldsymbol{\pi}_k) &+ \frac{\Lambda_2(\mathbf{s}_{k+1})}{2} \|\boldsymbol{\pi} - \boldsymbol{\pi}_k\|^2. \end{aligned} \quad (18)$$

The new iterate $\boldsymbol{\pi}_{k+1}$ is then defined as a minimizer of the right term of (18). Hereagain, we can solve this problem in an inexact manner, that is to search for some $\boldsymbol{\pi}_{k+1} \in C_2$ satisfying

$$\begin{cases} (\boldsymbol{\pi}_{k+1} - \boldsymbol{\pi}_k)^\top \nabla_2 f(\mathbf{s}_{k+1}, \boldsymbol{\pi}_k) \\ + \gamma_{\pi,k}^{-1} \Lambda_2(\mathbf{s}_{k+1}) \|\boldsymbol{\pi}_{k+1} - \boldsymbol{\pi}_k\|^2 \leq 0, \\ \|\nabla_2 f(\mathbf{s}_{k+1}, \boldsymbol{\pi}_k) + \mathbf{r}_k^{(2)}\| \leq \kappa_2 \sqrt{\Lambda_2(\mathbf{s}_{k+1})} \|\boldsymbol{\pi}_{k+1} - \boldsymbol{\pi}_k\|, \end{cases} \quad (19)$$

for some $\mathbf{r}_k^{(2)} \in N_{C_2}(\boldsymbol{\pi}_{k+1})$ and $\kappa_2 > 0$. The existence of $\boldsymbol{\pi}_{k+1}$ can be shown from [25, Rem. 3.3]. In particular, a minimizer over C_2 of the right term in (18) satisfies (19). The kernel update can be deactivated, if the kernel is known (i.e., non blind case), Algorithm 1 then identifies with [14].

C. Convergence Result

We establish the following convergence theorem for Algorithm 1. Its proof is provided in the supplementary material.

Theorem 1. *Let $(\mathbf{s}_k)_{k \in \mathbb{N}}$ and $(\boldsymbol{\pi}_k)_{k \in \mathbb{N}}$ be sequences generated by Alg. 1. If (C_1, C_2) are semi-algebraic sets, and ∇f is Lipschitz on the domain of Ω , then the sequence $(\mathbf{s}_k, \boldsymbol{\pi}_k)_{k \in \mathbb{N}}$ converges to a critical point $(\hat{\mathbf{s}}, \hat{\boldsymbol{\pi}})$ of Problem (5).*

The above result is novel, as it extends [14, Theo.1] to the block alternating case using proof ingredients from [16], [26]. The assumption on (C_1, C_2) ensures that function Ω satisfies Kurdyka-Łojasiewicz inequality, which is essential for the proof of descent schemes in a non-convex setting [15].

IV. NUMERICAL RESULTS

A. Datasets

Two datasets A and B were considered. The original sparse signal $\bar{\mathbf{s}}$ and the observed signal \mathbf{y} are shown in Fig. 1, both of size $N = 200$. Signal \mathbf{y} is obtained from (1) where $\bar{\boldsymbol{\pi}}$ is a normalized Gaussian kernel with standard deviation 0.15 and size $L = 21$. The noise \mathbf{n} is zero-mean white Gaussian with variance σ^2 either equals 0.5 % or 1.0 % of x_{\max} defined as the maximum amplitude of $\bar{\mathbf{x}} = \bar{\boldsymbol{\pi}} * \bar{\mathbf{s}}$. Signal and kernel convolution is implemented with zero padding. Trend $\bar{\mathbf{t}}$ is taken as the low-frequency signal from [8].

B. Algorithmic settings

We set $C_1 = [0, 100]^N$, and C_2 the simplex unit set, i.e. $C_2 = \{\boldsymbol{\pi} = (\pi_\ell)_{1 \leq \ell \leq L} \in [0, +\infty[^L \text{ s.t. } \sum_{\ell=1}^L \pi_\ell = 1\}$. For such choices, the assumptions of Theorem 1 hold, and since metric (15) is diagonal, the resolution of (17) and (19) is straightforward, by [24, Prop. 24.11] and [27, Cor. 9]. Namely, for every $k \in \mathbb{N}$, and $i \in \{1, \dots, \mathcal{I}\}$,

$$\begin{cases} \mathbf{s}_{k,i} = \text{Proj}_{C_1}(\mathbf{s}_k - \gamma_{s,k} \mathbf{A}_{1,\rho_{k,i}}(\mathbf{s}_k, \boldsymbol{\pi}_k)^{-1} \nabla_1 f(\mathbf{s}_k, \boldsymbol{\pi}_k)), \\ \boldsymbol{\pi}_{k+1} = \text{Proj}_{C_2}(\boldsymbol{\pi}_k - \gamma_{\pi,k} \Lambda_2(\mathbf{s}_{k+1})^{-1} \nabla_2 f(\mathbf{s}_{k+1}, \boldsymbol{\pi}_k)). \end{cases}$$

Hereabove, Proj_{C_1} is the projection over the positive orthant, that has a simple closed form expression, while Proj_{C_2} is the projection over the simplex unit set, that can be computed using the fast procedure from [28]. For simplicity, we set constant stepsizes $\gamma_{s,k} \equiv \gamma_{\pi,k} \equiv 1.9$, thus satisfying the required range assumption. Moreover, we take $\theta = 0.5$ in the TR update, and a maximum of $\mathcal{I} = 50$ of TR trials. We use the same initialization strategy for all methods as in [13], namely $\mathbf{s}_0 \in C_1$ is a constant positive-valued signal and $\boldsymbol{\pi}_0 \in C_2$ is a centered Gaussian filter with standard deviation of 1. The stopping criterion parameters are set as $\varepsilon = \sqrt{N} \times 10^{-6}$ and $K_{\max} = 2000$.

C. Numerical results

PENDANTSS jointly performs blind deconvolution and trend removal, using SPOQ penalty. Let us recall that SOOT penalty from [13] is retrieved by setting $(p, q) = (1, 2)$ in SPOQ. Another setting will be analyzed, namely $(p, q) = (0.75, 2)$. Other choices led to similar or poorer restoration results, as also observed in [14]. In the spirit of an ablation study, we compare PENDANTSS pipeline with the state-of-the-art background estimation method backcor [7] to estimate and remove the trend, followed by the blind deconvolution

method [13], to estimate the signal \hat{s} and the kernel $\hat{\pi}$. In both cases, we either use SPOQ $(p, q) = (0.75, 2)$, or SPOQ $(p, q) = (1, 2)$ (i.e., SOOT) for promoting sparsity in \hat{s} .

We use signal-to-noise ratios to evaluate our estimations, respectively for signal (SNR_s), kernel (SNR_π) and trend (SNR_t). For instance, $\text{SNR}_s = 20 \log_{10}(\|\bar{s}\|_2 / \|\bar{s} - \hat{s}\|_2)$. Moreover, TSNR evaluates the SNR only on the support of the original sparse signal. While their support are not known in general, it reveals how peak-derived quantities (height, width, area), important for downstream quantitative chemical analysis, would be impacted by detrending and deconvolution.

Hyperparameters, e.g. regularization parameters of backcor [7] and SPOQ/SOOT parameters (λ, β, η) , are adjusted through grid search to maximize a weighted sum of SNRs for one completely known reference realization, i.e. $2\text{SNR}_s + \text{SNR}_\pi + \text{SNR}_t$, which appeared as a representative metric in our experiments. We set $\alpha = 7 \times 10^{-7}$ as recommended in [14]. In practice, (α, β, η) have little influence on performance, while the choice of λ is critical. The cutoff frequency of the low-pass filter in (3) is chosen as the best performing point over the first ten peak points of the modulus of the signal frequency spectrum. To assure the kernel is centered, a spatial shift on the estimated kernel and the sparse signal is applied as a post-processing step as spatially shifted kernels and sparse signals result in the same observed signal. A rough grid search determines the number of inner loops to maximize the SNR_s .

Table I summarizes the results of mean SNR values, and standard deviations after the “ \pm ” sign, calculated over two hundred noise realizations. Best and second best values are almost always achieved by the proposed PENDANTSS approach with $(p, q) = (0.75, 2)$ or $(1, 2)$. The difference with the baseline methods is also significant for all cases especially in terms of TSNR_s and SNR_t . One exception lies on SNR_π with dataset B with the noise level of 1.0% of x_{\max} , where the second best is achieved by the combination backcor+SPOQ. We stress out that in such problems, correct estimations of sparse signal and baseline are usually more important than kernel estimation.

Regarding parameters (p, q) , the performance of PENDANTSS is dependent on the datasets and the noise level. Considering various SPOQ parameters is indeed beneficial. According to the presented simulation results, PENDANTSS with $(p, q) = (0.75, 2)$ is better for datasets with sparser, well-separable peaks (dataset A) whereas PENDANTSS with $(p, q) = (1, 2)$ is preferable for more challenging datasets (dataset B). Graphical details on the quality of estimated peaks are provided as supplementary material. Computational cost for PENDANTSS is slightly higher than for the sequential method with backcor: in the order of 4 s. vs 1 s. for dataset A and 20 s. vs 10 s. for dataset B on a standard laptop.

V. CONCLUSION AND PERSPECTIVES

We address a complicated joint sparse signal blind deconvolution and additive trend problem. Our method handles smooth trend removal by exploiting the low-pass property and simplifies the problem into a blind deconvolution problem formulation integrating the SPOQ sparse penalty and appropriate constraints. A new block alternating algorithm with

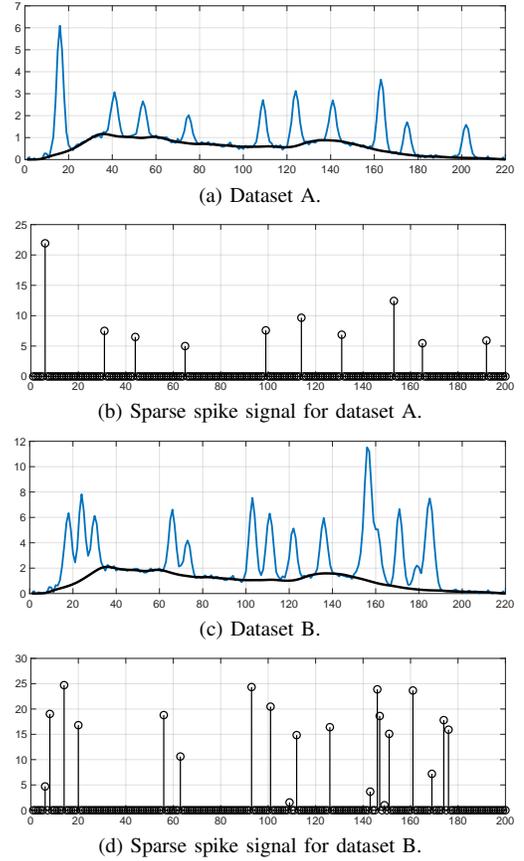


Fig. 1. Unknown sparse signal \bar{s} (b) and (d); in (a) and (c) observation y (blue) and baseline \hat{t} (black) (bottom) for datasets A and B. Signal A has 10 spikes (5.0% of sparsity) while signal B has 20 spikes (10.0% of sparsity).

TABLE I
NUMERICAL RESULTS ON DATASETS A AND B. SNR QUANTITIES IN DB.
BEST PERFORMING METHOD FOLLOWED BY **, SECOND BY *.

Noise level σ (% of x_{\max})		Dataset A		Dataset B	
		0.5%	1.0%	0.5%	1.0%
SNR_s	backcor+SOOT	29.2 \pm 0.7	28.5 \pm 1.9	14.9 \pm 4.0	11.5 \pm 4.7
	backcor+SPOQ	29.2 \pm 0.7	29.3 \pm 1.3	12.9 \pm 3.5	11.3 \pm 4.4
	PENDANTS (1, 2)	32.9 \pm 1.5*	30.9 \pm 2.2*	22.3 \pm 8.2**	17.5 \pm 8.4**
	PENDANTS (0.75, 2)	33.2 \pm 2.3**	31.0 \pm 4.2**	15.9 \pm 4.5*	12.9 \pm 4.6*
TSNR_s	backcor+SOOT	29.2 \pm 0.7	29.3 \pm 1.3	16.6 \pm 3.5	13.4 \pm 4.3
	backcor+SPOQ	29.2 \pm 0.7	29.3 \pm 1.3	15.1 \pm 3.0	13.7 \pm 3.7
	PENDANTS (1, 2)	34.1 \pm 1.4*	32.2 \pm 2.1*	24.9 \pm 8.0**	19.2 \pm 7.7**
	PENDANTS (0.75, 2)	35.4 \pm 1.7**	32.6 \pm 3.8**	17.7 \pm 4.0*	14.5 \pm 4.1*
SNR_t	backcor+SOOT	20.5 \pm 0.2	20.3 \pm 0.4	15.5 \pm 0.5	14.8 \pm 0.8
	backcor+SPOQ	20.5 \pm 0.2	20.3 \pm 0.4	15.5 \pm 0.5	14.8 \pm 0.8
	PENDANTS (1, 2)	26.9 \pm 0.5**	26.0 \pm 0.8**	22.0 \pm 0.4*	21.6 \pm 1.0**
	PENDANTS (0.75, 2)	26.9 \pm 0.6**	26.0 \pm 1.0**	24.6 \pm 0.6**	19.6 \pm 3.9*
SNR_π	backcor+SOOT	36.3 \pm 1.3	33.9 \pm 1.7	30.3 \pm 1.3	28.5 \pm 1.8
	backcor+SPOQ	36.3 \pm 1.3	34.0 \pm 1.7	33.1 \pm 1.9	31.2 \pm 2.1*
	PENDANTS (1, 2)	41.3 \pm 2.0**	34.4 \pm 2.4**	38.3 \pm 1.9**	33.6 \pm 2.2**
	PENDANTS (0.75, 2)	41.3 \pm 2.0**	34.2 \pm 2.5*	35.7 \pm 1.5*	25.4 \pm 5.5

trust region acceleration is introduced, and its convergence is established. PENDANTSS outperforms comparable methods on typical sparse analytical signals on simulation results. Further works include its validation on other sparse spike signals. The appropriate parameters for the sparsity-promoting norm ratio penalty ought to be investigated, for instance with respect to the alleged signal sparsity or peak separability. PENDANTSS Matlab code and hyper-parameter extensive analysis are available at <https://github.com/paulzhengfr/PENDANTSS>. The authors thank Vincent Mazet, Bruno Lety, the reviewers and the associate editor.

REFERENCES

- [1] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 27, no. 2, pp. 113–120, Apr. 1979.
- [2] J. Lynch, Ed., *Physico-Chemical Analysis of Industrial Catalysts. A Practical Guide to Characterisation*. Édition Technip, Sep. 2003.
- [3] A. Marmin, M. Castella, J.-C. Pesquet, and L. Duval, "Sparse signal reconstruction for nonlinear models via piecewise rational optimization," *Signal Process.*, vol. 179, no. 107835, Feb. 2021.
- [4] J. Gauthier, L. Duval, and J.-C. Pesquet, "Optimization of synthesis oversampled complex filter banks," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3827–3843, Oct. 2009.
- [5] L. Duval, L. T. Duarte, and C. Jutten, "An overview of signal processing issues in chemical sensing," in *Proc. Int. Conf. Acoust. Speech Signal Process.*, Vancouver, BC, Canada, May 26-31, 2013, pp. 8742–8746.
- [6] Z. Zhao, S. Wang, D. Wong, C. Sun, R. Yan, and X. Chen, "Robust enhanced trend filtering with unknown noise," *Signal Process.*, vol. 180, p. 107889, Mar. 2021.
- [7] V. Mazet, C. Carteret, D. Brie, J. Idier, and B. Humbert, "Background removal from spectra by designing and minimising a non-quadratic cost function," *Chemometr. Intell. Lab. Syst.*, vol. 76, no. 2, pp. 121–133, 2005.
- [8] X. Ning, I. W. Selesnick, and L. Duval, "Chromatogram baseline estimation and denoising using sparsity (BEADS)," *Chemometr. Intell. Lab. Syst.*, vol. 139, pp. 156–167, Dec. 2014.
- [9] S. Amari, S. C. Douglas, A. Cichocki, and H. H. Yang, "Multichannel blind deconvolution and equalization using the natural gradient," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Comm.*, 1997, pp. 101–104.
- [10] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, Jun. 21-25, 2011, pp. 233–240.
- [11] S. Chaudhuri, R. Velmurugan, and R. Rameshan, "Blind deconvolution methods: A review," in *Blind Image Deconvolution. Methods and Convergence*. Springer, 2014, pp. 37–60.
- [12] Q. Sun and D. Donoho, "Convex sparse blind deconvolution," *PREPRINT*, Jun. 2021.
- [13] A. Repetti, M. Q. Pham, L. Duval, E. Chouzenoux, and J.-C. Pesquet, "Euclid in a taxicab: Sparse blind deconvolution with smoothed ℓ_1/ℓ_2 regularization," *IEEE Signal Process. Lett.*, vol. 22, no. 5, pp. 539–543, May 2015.
- [14] A. Cherni, E. Chouzenoux, L. Duval, and J.-C. Pesquet, "SPOQ ℓ_p -over- ℓ_q regularization for sparse signal recovery applied to mass spectrometry," *IEEE Trans. Signal Process.*, vol. 68, pp. 6070–6084, 2020.
- [15] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *Math. Programm.*, vol. 146, no. 1-2, pp. 459–494, Aug. 2014.
- [16] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, "A block coordinate variable metric forward-backward algorithm," *J. Global Optim.*, vol. 66, no. 3, pp. 457–485, Feb. 2016.
- [17] I. W. Selesnick, H. L. Graber, D. S. Pfeil, and R. L. Barbour, "Simultaneous low-pass filtering and total variation denoising," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1109–1124, Mar. 2014.
- [18] H. Fu, M. K. Ng, M. Nikolova, and J. L. Barlow, "Efficient minimization methods of mixed ℓ_2 - ℓ_1 and ℓ_1 - ℓ_1 norms for image restoration," *SIAM J. Sci. Comput.*, vol. 27, no. 6, pp. 1881–1902, Jan. 2006. [Online]. Available: <http://dx.doi.org/10.1137/040615079>
- [19] E. Soubies, L. Blanc-Féraud, and G. Aubert, "A unified view of exact continuous penalties for ℓ_2 - ℓ_0 minimization," *SIAM J. Optim.*, vol. 27, no. 3, pp. 2034–2060, Jan. 2017.
- [20] N. Hurley and S. Rickard, "Comparing measures of sparsity," *IEEE Trans. Inform. Theory*, vol. 55, no. 10, pp. 4723–4741, Oct. 2009.
- [21] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.
- [22] E. Chouzenoux, S. Martin, and J.-C. Pesquet, "A local MM subspace method for solving constrained variational problems in image recovery," *J. Math. Imaging Vision*, 2022.
- [23] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust-Region Methods*, ser. MOS-SIAM Series on Optimization. Society for Industrial Mathematics, 2000.
- [24] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, 2nd ed., ser. CMS books in mathematics. Springer, 2011.
- [25] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, "Variable metric forward-backward algorithm for minimizing the sum of a differentiable function and a convex function," *J. Optim. Theory Appl.*, vol. 162, no. 1, pp. 107–132, Jul. 2014.
- [26] L. T. K. Hien, N. Gillis, and P. Patrinos, "Inertial block proximal methods for non-convex non-smooth optimization," in *Proc. Int. Conf. Mach. Learn.*, vol. 119, Jul. 13–18, 2020, pp. 5671–5681.
- [27] S. Becker and M. J. Fadili, "A quasi-Newton proximal splitting method," in *Proc. Ann. Conf. Neur. Inform. Proc. Syst.*, vol. 2, Dec. 3-6, 2012, pp. 2618–2626.
- [28] L. Condat, "Fast projection onto the simplex and the ℓ_1 ball," *Math. Programm.*, vol. 158, no. 1-2, pp. 575–585, 2016.

PENDANTSS: Supplementary Material

I. PROOF OF THEOREM 1 FOR ALGORITHM 1

We first provide a useful majorant metric matrix property.

Lemma 1 *There exists $(\underline{\lambda}, \bar{\lambda}) \in]0, +\infty[^2$ such that for every $k \in \mathbb{N}$, and for every $i \in \{1, \dots, \mathcal{I}\}$,*

$$\begin{cases} \underline{\lambda} \mathbf{Id}_N \preceq \mathbf{A}_{1, \rho_{k, i}}(\mathbf{s}_k, \boldsymbol{\pi}_k) \preceq \bar{\lambda} \mathbf{Id}_N, \\ \underline{\lambda} \leq \Lambda_2(\mathbf{s}_k) \leq \bar{\lambda}. \end{cases} \quad (\text{A1})$$

Proof. Direct consequence of [14, Prop. 2] and [13, Prop. 1]. \square
We then show that Algorithm 1 satisfies two essential descent properties, that are key for the convergence analysis.

Lemma 2 *There exists $(\mu_1, \mu_2) \in]0, +\infty[^2$ such that, for every $k \in \mathbb{N}$, the following descent properties hold:*

$$\Omega(\mathbf{s}_{k+1}, \boldsymbol{\pi}_k) \leq \Omega(\mathbf{s}_k, \boldsymbol{\pi}_k) - \frac{\mu_1}{2} \|\mathbf{s}_{k+1} - \mathbf{s}_k\|^2, \quad (\text{A2})$$

$$\Omega(\mathbf{s}_{k+1}, \boldsymbol{\pi}_{k+1}) \leq \Omega(\mathbf{s}_{k+1}, \boldsymbol{\pi}_k) - \frac{\mu_2}{2} \|\boldsymbol{\pi}_{k+1} - \boldsymbol{\pi}_k\|^2. \quad (\text{A3})$$

Proof. Let $k \in \mathbb{N}$. We remind that the objective function Ω is defined in (12), with g specified in (11). By construction, $\mathbf{s}_{k+1} \in \mathcal{B}_{q, \rho} \cap C_1$ for some $i \in \{1, \dots, \mathcal{I}\}$. Summing the majoration (14) and the first inequality in (17) yields:

$$\Omega(\mathbf{s}_{k+1}, \boldsymbol{\pi}_k) \leq f(\mathbf{s}_k, \boldsymbol{\pi}_k) - (\gamma_{s, k}^{-1} - \frac{1}{2}) \|\mathbf{s}_k - \mathbf{s}_{k+1}\|_{\mathbf{A}_{1, \rho}(\mathbf{s}_k, \boldsymbol{\pi}_k)}^2.$$

We notice that $f(\mathbf{s}_k, \boldsymbol{\pi}_k) = \Omega(\mathbf{s}_k, \boldsymbol{\pi}_k)$ since $\mathbf{s}_k \in C_1$ and $\boldsymbol{\pi}_k \in C_2$. Using Lemma 1 and the range assumption on $\gamma_{s, k}$ allows to show (A2) for $\mu_1 = \lambda \bar{\gamma} / (2 - \bar{\gamma})$. Again by construction, $\boldsymbol{\pi}_{k+1} \in C_2$. Summing (18) and (19) leads to:

$$\begin{aligned} \Omega(\mathbf{s}_{k+1}, \boldsymbol{\pi}_{k+1}) &\leq f(\mathbf{s}_{k+1}, \boldsymbol{\pi}_k) - \\ &\quad (\gamma_{\pi, k}^{-1} - \frac{1}{2}) \Lambda_2(\mathbf{s}_{k+1}) \|\boldsymbol{\pi}_{k+1} - \boldsymbol{\pi}_k\|^2. \end{aligned}$$

Here again, we use $f(\mathbf{s}_{k+1}, \boldsymbol{\pi}_k) = \Omega(\mathbf{s}_{k+1}, \boldsymbol{\pi}_k)$ as $\mathbf{s}_{k+1} \in C_1$ and $\boldsymbol{\pi}_k \in C_2$. The descent property (A3) is obtained by using Lemma 1, the range constraint on $\gamma_{\pi, k}$, and setting $\mu_2 = \lambda \bar{\gamma} (2 - \bar{\gamma})$. \square

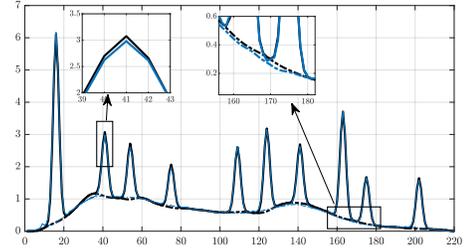
The rest of the proof of Theorem 1 is obtained by following the same lines than the one of [16, Theorem 3.1], leveraging the Lipschitz smoothness of f on the domain $C_1 \times C_2$ of Ω , and the Kurdyka-Łojasiewicz inequality satisfied by Ω .

II. ADDITIONAL RESULTS

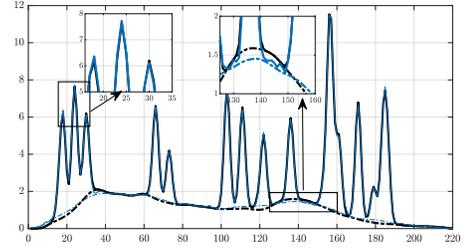
Figures 2 and 3 provide additional insights into PENDANTSS restoration. Dataset A in Figure 2-(a) presents sparse and well-isolated peaks. Accurate peak restoration is secured. Peak shapes are well recovered (left-hand zoom), and the estimated trend matches well the actual baseline. As a consequence, peak features that are computed with respect to the trend (height, area) are likely to be well-estimated with PENDANTSS. The less sparse Dataset B in Figure 2-(b) shows that the separation and the height of close peaks are accurately

matched. Some overshoot in trend estimation can be noticed. It is however not likely to drastically affect relative peak height or area computations.

Retrieved spikes are exposed in Figure 3. For Dataset A, well-separated spikes are accurately recovered using PENDANTSS. Estimated amplitudes and locations are almost indistinguishable from the original ones. This is exemplified for the less sparse Dataset B in Figure 3-(b). Isolated peaks are well-estimated. However, some spikes (for instance around index 175) for Dataset B in Figure 3-(b) remain unelucidated. Three contiguous spikes are estimated, instead of two. Such an ambiguous solution is typical to source separation problems.

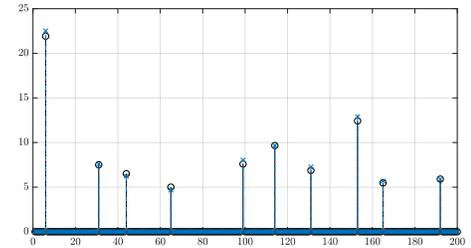


(a) Dataset A - reconstruction and trend.

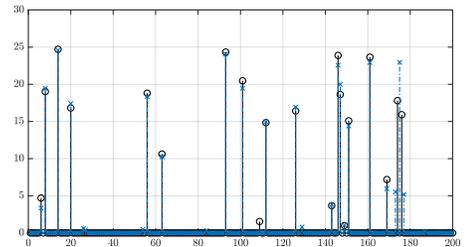


(b) Dataset B - reconstruction and trend.

Fig. 2: Ground truth (thick black line) and proposed estimation results (thin blue line) for the baseline t (dashed dot) and the signal $\mathbf{s} * \boldsymbol{\pi}$ (continuous).



(a) Dataset A - sparse spike signal.



(b) Dataset B - sparse spike signal.

Fig. 3: Ground truth (black line with circle marker) and proposed estimation results (blue line with cross marker) for sparse spike signal \mathbf{s} .