Sparse Spikes Retrieval on Thin Grids II: the Continuous Basis Pursuit

Vincent Duval[1,2], Gabriel Peyré[2,1]

[1] INRIA, MOKAPLAN

E-mail: `vincent.duval@inria.fr`

[2] CNRS and CEREMADE, Université Paris-Dauphine

E-mail: `gabriel.peyre@ceremade.dauphine.fr`

Abstract. This article analyzes the performances of the Continuous Basis Pursuit (C-BP) method for sparse super-resolution. The C-BP has been recently proposed by Ekanadham, Tranchina and Simoncelli as a refined discretization scheme for the recovery of spikes in inverse problems regularization. One of the most well known discretization scheme, the Basis Pursuit (BP, also known as Lasso) makes use of a finite dimensional ℓ^1 norm on a grid. In contrast, the C-BP rather uses a linear interpolation of the spikes position to enable the recovery of spikes between grid's points. When the thought after solution is constrained to be positive, a remarkable feature of this approach is that it retains the convexity of the initial ℓ^1 problem. For deconvolution problem, it is well known that ℓ^1-type methods (including BP and C-BP) recovers exactly the unknown sparse sum of positive Diracs when there is no noise in the measurements. Our main contribution identify a non-degeneracy condition ensuring that the support of the solution enjoy some stability when noise is added. More precisely, we show that, in the small noise regime, when the non-degeneracy condition holds, the C-B method estimates a pair of Diracs around each spike of the input measure. We also derive the asymptotic of the recovery error when the grid size tends to zero. We show some numerical illustrations of this stability to noise for both the BP and C-BP methods, and evaluate numerically these non-degeneracy conditions.

1. Introduction

This article studies a specific convex optimization approach, the Continuous Basis-Pursuit (C-BP) for sparse super-resolution. A detailed review of the litterature on super-resolution, and in particular variational regularization technics, can be found in the companion paper [17], which is dedicated to the Basis Pursuit (BP, or Lasso) method.

In the following, we focus for simplicity on the compact 1-D domain $\mathbb{T}$, but the algorithms considered (Lasso and C-BP) as well as our theoretical analysis can be extended to higher dimensional settings (see Section 1.7).

1.1. Super-resolution

We formalize the sparse super-resolution problem mathematically as the question of recovering an unknown Radon measure $m_0 \in \mathcal{M}(\mathbb{T})$ defined on the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ (i.e. an interval with periodic boundary conditions) from noisy linear observations in a separable Hilbert space $\mathcal{H}$, which we write as

$$y = \Phi(m_0) + w \in \mathcal{H} \tag{1}$$

where $w \in \mathcal{H}$ is some measurement noise. The terminology "super-resolution" refers here to the fact that $\Phi : \mathcal{M}(\mathbb{T}) \to \mathcal{H}$ is a bounded linear map defined through the integration against a *continuous* kernel function $\varphi : \mathbb{T} \to \mathcal{H}$

$$\forall x \in \mathbb{T}, \quad \Phi(m) = \int_{\mathbb{T}} \varphi(x) \mathrm{d}m(x). \tag{2}$$

The continuity of φ is crucial, since it allows to pose the problem over the space of measures, and thus allows for the possible recovery of highly localized solutions (e.g. sums of Diracs' masses). In order for our theoretical findings to hold, we furthermore require that $\varphi \in \mathscr{C}^3(\mathbb{T}, \mathcal{H})$.

A typical class of problems are so-called deconvolution problems, when the operator Φ is translation invariant. This corresponds to using $\mathcal{H} = \mathrm{L}^2(\mathbb{T})$, $\varphi(x) : x' \mapsto \tilde{\varphi}(x' - x)$ for some smooth kernel $\tilde{\varphi}$ defined on $\mathbb{T}$. This deconvolution setup is equivalent to computing the measurements over the Fourier domain. When the highest measured frequency f_c is finite, this corresponds to using $\mathcal{H} = \mathbb{C}^{2f_c+1}$ and $\varphi(x) = (e^{2ikx\pi})_{k=-f_c}^{f_c}$.

1.2. Sparse Regularization

The problem of inverting (1) is severely ill-posed. A particular example is when Φ is a low pass filter, which is a typical setting for many problems in imaging. In several applications, it makes sense to impose some sparsity assumption on the data to recover. This idea has been introduced first in the geoseismic literature, to model the layered structure of the underground using sparse sums of Dirac masses [9]. Sparse regularization has later been studied by David Donoho and co-workers, see for instance [13].

Since Φ is usually ill-posed, a regularization scheme is needed to perform its inversion. Following several recent works (see for instance [5, 11, 7]), we study here sparse regularizations induced by by the total variation of the measure (i.e. the total absolute mass)

$$|m|(\mathbb{T}) \overset{\mathrm{def.}}{=} \sup \left\{ \int_{\mathbb{T}} \psi(x) \mathrm{d}m(x) \; ; \; \psi \in \mathscr{C}(\mathbb{T}), \; \|\psi\|_\infty \leqslant 1 \right\}. \tag{3}$$

For discrete measures defined on a finite grid

$$\mathcal{G} \overset{\mathrm{def.}}{=} \{z_i \; ; \; i \in [\![0, G-1]\!]\} \subset \mathbb{T}$$

of points, which take the form

$$m_{a,\mathcal{G}} \stackrel{\text{def.}}{=} \sum_{i=0}^{G-1} a_i \delta_{z_i} \in \mathcal{M}(\mathbb{T}),$$

the total variation is equal to the usual discrete ℓ^1 norm

$$|m_{a,\mathcal{G}}|(\mathbb{T}) = \|a\|_1 = \sum_i |a_i|,$$

which is at the heart of the celebrated Basis Pursuit [] method (also known as the LASSO []). The total variation is thus the natural way to extend these intrinsically discrete and finite-dimensional method to the infinite-dimensional class of measure, to enable a grid-free regularization.

This total variation is used as a regularizer to define the

$$\min_{m \in \mathcal{M}(\mathbb{T})} \frac{1}{2} \|y - \Phi(m)\|_{\mathcal{H}}^2 + \lambda |m|(\mathbb{T}) \tag{4}$$

where $|m|(\mathbb{T})$ is the total variation of the measure m, defined as

This formulation of the recovery of sparse Radon measures has recently received lots of attention in the literature, see for instance the works of [5, 11, 7]. In the case where there is no noise, $w = 0$, it makes sense to consider $\lambda \to 0$ and to solve the following limit problem

$$\min_{m \in \mathcal{M}(\mathbb{T})} \{|m|(\mathbb{T}) \; ; \; \Phi(m) = \Phi(m_0)\} . \tag{5}$$

1.3. LASSO

[**ToDo: G: shorten this section.**]

The optimization problem (4) is convex but infinite dimensional, and while there exists solvers when Φ is measuring a finite number of Fourier frequency (see [7]), they do not scale well with the number of frequencies. Furthermore, the case of an arbitrary linear operator Φ is still difficult to handle, see [5] for an iterative scheme. The vast majority of practitioners thus approximate (4) by a finite dimensional problem computed over a finite grid $\mathcal{G} \stackrel{\text{def.}}{=} \{z_i \; ; \; i \in [\![0, G-1]\!]\} \subset \mathbb{T}$, by restricting their attention to measures of the form

$$m_{a,\mathcal{G}} \stackrel{\text{def.}}{=} \sum_{i=0}^{G-1} a_i \delta_{z_i} \in \mathcal{M}(\mathbb{T}).$$

For such a discrete measure, one has $|m|(\mathbb{T}) = \sum_{i=0}^{G-1} |a_i| = \|a\|_1$, which can be interpreted as the fact that $|\cdot|(\mathbb{T})$ is the natural extension of the ℓ^1 norm from finite dimensional vectors to the infinite dimensional space of measures. Inserting this parametrization in (4) leads to the celebrated Basis-Pursuit problem [8], which is also known as the LASSO method in statistics [30],

$$\min_{a \in \mathbb{R}^N} \frac{1}{2} \|y - \Phi_{\mathcal{G}} a\|_{\mathcal{H}}^2 + \lambda \|a\|_1 \tag{6}$$

where in the following we make use of the notations

$$\Phi_{\mathcal{G}}a \stackrel{\text{def.}}{=} \Phi(m_{a,\mathcal{G}}) = \sum_{i=0}^{G-1} a_i\varphi(z_i), \quad \text{and} \quad \Phi'_{\mathcal{G}}b \stackrel{\text{def.}}{=} \Phi'(m_{b,\mathcal{G}}) = \sum_{i=0}^{G-1} b_i\varphi'(z_i). \tag{7}$$

One can understand (6) as performing a nearest neighbor interpolation of the Dirac's locations.

This approximation is however quite crude, and we recently showed in [16] that it leads to imperfect estimation of both the number of spikes and their locations. Indeed, this problem typically recovers up to twice as many spikes as the input measures, because spikes of m_0 gets duplicated as the two nearest neighbors on the grid $\mathcal{G}$.

Note that while we focus in this paper on convex recovery method, and in particular ℓ^1-type regularization, there is a vast literature on the subject, which makes use of alternative algorithms, see for instance [25, 4] and the references therein.

1.4. Continuous Basis-Pursuit (C-BP)

To obtain a better approximation of the infinite dimensional problem, [19] proposes to perform a first order approximation of the kernel. This method assumes that the unknown measure is positive. To ease the exposition, we consider a uniform grid $\mathcal{G} \stackrel{\text{def.}}{=} \{i/G \; ; \; i \in [\![0, G-1]\!]\}$ of G points, so that the grid size is $h \stackrel{\text{def.}}{=} 1/G$. The C-BP method of [19] solves

$$\min_{(a,b)\in\mathbb{R}^G\times\mathbb{R}^G} \frac{1}{2} \left\| y - \Phi_{\mathcal{G}}a - \Phi'_{\mathcal{G}}b \right\|_{\mathcal{H}}^2 + \lambda \|a\|_1 \quad \text{subject to} \quad |b| \leqslant \frac{h}{2}\, a, \tag{8}$$

where the inequality should be understood component-wise. Note also that the obtained a is always nonnegative, hence the C-BP method is tailored for the recovery of positive measures. This is a convex optimization problem, which can be solved using traditional conic optimization methods. As detailed in Section 2.2, this problem can also be re-cast as a Lasso in dimension $2G$ with positivity constraints (see Section 2.2). Hence it can be solved using a large variety of first order proximal method, the most simple one being the Forward-Backward, see [2] and the references therein.

If $(a^\star, b^\star)$ are solutions of (8), one recovers an output discrete measure defined by

$$m^\star = \sum_{a_i^\star \neq 0} a_i^\star \delta_{x_i^\star} \quad \text{where} \quad x_i^\star \stackrel{\text{def.}}{=} ih + \frac{b_i^\star}{a_i^\star}, \tag{9}$$

where we set $\frac{b_i^\star}{a_i^\star} = 0$ whenever $a_i^\star = 0$. The rationale behind (8) is to perform a first order Taylor approximation of the operator Φ, where the variable $\tau_i \stackrel{\text{def.}}{=} b_i/a_i \in [-h/2, h/2]$ encodes the horizontal shift of the Dirac location with respect to the grid sample ih. The landmark idea introduced in [19] is that, while the optimization is non-convex with respect to the pair (a, τ), it is convex with respect to the pair (a, b).

1.5. Previous Works

Most of the early work to assess the performance of convex sparse regularization has focussed its attention on the finite dimensional case, thus considering only the LASSO problem (6). While the literature on this subject is enormous, only very few works actually deal with deterministic and highly correlated linear operators such as low-pass convolution kernels. The initial works of Donoho [13] study the Lipschitz behavior of the inverse map $y \mapsto a^\star$, where $a^\star$ is a solution of (6), as a function of the bandwidth of the bandpass filter. The first work to address the question of spikes identification (i.e. recovery of the exact location of the spikes over a discrete grid) is [15]. This work uses the analysis of ℓ^1 regularization introduced by Fuchs in [22]. This type of analysis ensures that the support of the input measure is stable under small noise perturbation of the measurements. Our finding is that this is however never the case (the support is always unstable) when the grid is thin enough, and we thus introduce the notion of "extended support", which is in some sense the smallest extension of the support which is stable. The idea of extending the support to study the recovery performance of ℓ^1 methods can be found in the work of Dossal [14] who focusses on noiseless recovery and stability in term of ℓ^2 error.

Recently, a few works have studied the theoretical properties of the recovery over measures (4). Candès and Fernandez-Granda show in [7] that this convex program does recover exactly the initial sparse measure when $w = 0$ and $\lambda \to 0$ (i.e. program (5)) under a minimum-separation condition, i.e. if the spikes are well-separated. The robustness to noisy measurements is analyzed by the same authors in [6] using an Hilbertian norm, and in [20, 1] in terms of spikes localization. The work of [28] analyzes the reconstruction error. Lastly, [16] provides a condition ensuring that (4) recovers the same number of spikes as the input measure and that the error in terms of spikes localization and elevation has the same order as the noise level. It is important to note that in the special case where m_0 is a positive measure (which is the setup considered by the C-BP method), then m_0 is always a solution to (5)**[ToDo: preciser les hypotheses?]**, as shown in [11] (see also [12] for a refined analysis of the stability to noise in this special case).

Very few works have tried to bridge the gap between these grid-free methods over the space of measures, and finite dimensional discrete approximations that are used by practitioners. These theoretical questions are however relevant from a practitioner's point of view, and we refer [24] for experimental observations of the impact of discretization and the corresponding recovery bias. The convergence (in the sense of measures) of the solutions of the discrete problem toward to ones of the grid-free problem is shown in [29], where a speed of convergence is shown using tools from semi-infinite programming [27]. The same authors show in [3] that the discretized problem achieves a similar prediction L^2 error as the grid-free method. Γ-convergence results on ℓ^1 but also ℓ^0 regularization are provided in the PhD work [23]. In [16], we have shown that solutions of the discrete LASSO problem estimate in general as much as twice the number

of spikes as the input measure. We detail in the following section how the present work gives a much more precise and general analysis of this phenomenon.

1.6. Contributions

[**ToDo: G: explain the link with the other paper.**]

Our paper is composed of two main contributions (Theorems 1 and 2) that study the robustness to noise of the support of the solution of LASSO and C-BP finite dimensional recovery problems. We stress the fact that we always suppose that the sought after sparse measure m_0 is identifiable, i.e. is the solution of the BLASSO program (5) (i.e. in the noiseless case $w = 0, \lambda = 0$). This mandatory hypothesis is now well understood, as detailed in Section 1.5, and is always true if the measure m_0 is positive, or under a minimum separation distance between the spikes.

Our main contributions study whether the support of the recovered solution is close from the one of m_0 in the presence of a small noise. Such a stability cannot hold in full generality, and requires a strengthening of the optimality condition for m_0 being identifiable, which we refers in the following as a "non-degeneracy" condition.

[**ToDo: G: expand this**]

Section 2 first studies this problem in an abstract setting, where it is reformulated as a LASSO with positivity constraints. We derive similar noise robustness properties as for the LASSO, and we characterize the extended support under a non-degeneracy condition (hypothesis (31)), see Theorem 1.

Section 3 gathers the last set of contributions. Working on a thin grid, we show the Γ-convergence of Problem (8) towards the Beurling LASSO (4) with positivity constraints and we give a fine analysis of the support of the solutions as the grid stepsize tends to zero. We also study the low noise behavior when the measure has support on the grid. Our main contribution, Theorem 2 shows that, under a non-degeneracy condition (the "Twice Non-Degenerate Source Condition", see Definition 5), the recovered spikes appear again in pairs,

$$m_\lambda = \sum_{\nu=1}^{N} \left(\alpha_{\lambda,\nu} \delta_{x_{0,\nu}+t_\nu} + \beta_{\lambda,\nu} \delta_{x_{0,\nu}+\varepsilon_\nu h/2} \right) \quad \text{where} \quad \begin{cases} \varepsilon_\nu \in \{-1, +1\}, \\ -h/2 < t_\nu < h/2. \end{cases}$$

A closed form expression for ε is given, which depends on some corresponding *natural shift* intrinsic to the measure (which differs from the one of the LASSO). The corresponding low noise regime is characterized by $\frac{\|w\|_{\mathcal{H}}}{\lambda} = O(1)$ and $\lambda = O(h^3)$.

It is worth noting that, in this setting of sparse spikes retrieval on thin grids, our contributions give important information about the structure of the recovered spikes when the noise w is small. This is especially important since, contrary to common belief, the spikes locations for LASSO and C-BP are not stable: even for an arbitrary small noise w, neither methods retrieve the correct input spikes locations.

Eventually, we illustrate in Section 5 these theoretical results with numerical experiments comparing LASSO and C-BP. We first display the evolution of the solution

path $\lambda \mapsto a_\lambda$ (a solution of (6)) and $\lambda \mapsto (a_\lambda, b_\lambda)$ (a solution of (8)). These paths are piecewise-affine, and our contributions (Theorem 2) precisely characterize the first affine segment of these paths, which perfectly matches the numerical observations.

1.7. Extensions

While we restrict here the exposition to 1-D problems, the C-BP formulation (8) can be extended to cope with measures in arbitrary dimension $d \geqslant 1$, i.e. to consider $m_0 \in \mathcal{M}(\mathbb{T}^d)$. This requires to define at each sampling grid point indexed by i a vector $b_i = (b_{i,k})_{k=1}^d \in \mathbb{R}^d$ together with the constraint $\|b_i\|_\infty \leqslant \frac{h}{2} a_i$, and also to use a matrix $\Phi'_\mathcal{G}$ defined as

$$\Phi'_\mathcal{G} b \overset{\text{def.}}{=} \sum_{i \in \mathcal{G}} \sum_{k=1}^d b_{i,k} \partial_k \varphi(x_i) \in \mathcal{H}$$

where ∂^k denote the differential operator with respect to the kth direction in $\mathbb{R}^d$. Our analysis carries over to this setting without major difficulties.

The paper [19] also proposes other interpolation schemes than a first order Taylor expansion at the grid points. In particular, they develop a "polar" interpolation which makes use of two adjacent grid points. This method seems to outperform the linear interpolation in practice, and has been employed to perform spikes sorting in neuronal recordings [18].

Extending the results we propose in the present paper to these higher dimensional settings and alternative interpolation schemes is an interesting avenue for future work.

Let us also mention that an important problem is to extend the C-BP method (9) to measures with arbitrary signs and that can even be complex-valued. Unfortunately, the corresponding constraint $|b| \leqslant |a|$ is then non-convex, which makes the mathematical analysis apparently much more involved. A non-convex and non-smooth optimization solver is proposed for this problem in [21], and shows promising practical performance for spectrum estimation.

1.8. Notations and preliminaries

The set of Radon measures (resp. positive Radon measures) is denoted by $\mathcal{M}(\mathbb{T})$ (resp. $\mathcal{M}^+(\mathbb{T})$). Endowed with the total variation norm (3), $\mathcal{M}(\mathbb{T})$ is a Banach space. Another useful topology on $\mathcal{M}(\mathbb{T})$ is the weak* topology: a sequence of measures $(m_n)_{n \in \mathbb{N}}$ weakly* converges towards $m \in \mathcal{M}(\mathbb{T})$ if and only if for all $\psi \in \mathscr{C}(\mathbb{T})$, $\lim_{n \to +\infty} \int_\mathbb{T} \psi \mathrm{d}m_n = \int_\mathbb{T} \psi \mathrm{d}m$. Any bounded subset of $\mathcal{M}(\mathbb{T})$ (for the total variation) is relatively sequentially compact for the weak* topology. Moreover the topology induced by the total variation is stronger than the weak* topology, and the total variation is sequentially lower semi-continuous for the weak* topology. Throughout the paper, given $\alpha \in \mathbb{R}^N$ and $x_0 \in \mathbb{T}^N$, the notation $m_{\alpha,x_0} \overset{\text{def.}}{=} \sum_{\nu=1}^N \alpha_\nu \delta_{x_{0,\nu}}$ hints that $\alpha_\nu \neq 0$ for all ν (contrary to the notation $m_{a,\mathcal{G}}$), and that the $x_{0,\nu}$'s are pairwise distinct.

The space $\mathcal{H}$ denotes a separable Hilbert space. As explained in the companion paper [17, Lemma 1], the operator $\Phi : \mathcal{M}(\mathbb{T}) \to \mathcal{H}$ defined in (2) is weak-* to weak continuous, and so are the operators involving the derivatives, for $k \in \mathbb{N}$,

$$\forall\, m \in \mathcal{M}(\mathbb{T}), \quad \Phi^{(k)}(m) = \int_{\mathbb{T}} \varphi^{(k)}(x) \mathrm{d}m(x).$$

Their respective adjoint operators, $\Phi^{(k),*} : \mathcal{H} \to \mathscr{C}(\mathbb{T})$, are given by $(\Phi^{(k),*}q)(t) = \langle q, \varphi^{(k)}(t) \rangle$ for all $q \in \mathcal{H}$, $t \in \mathbb{T}$. Moreover, $\Phi^{(k),*}q$ has the regularity of φ and $\frac{\mathrm{d}^k}{\mathrm{d}t^k}(\Phi^*q)(t) = (\Phi^{(k),*}q)(t)$. The $\infty, \mathcal{H}$-operator norm of $\Phi^* : \mathcal{H} \to \mathscr{C}(\mathbb{T})$ is defined as $\|\Phi^*\|_{\infty,\mathcal{H}} \stackrel{\text{def.}}{=} \sup \{\|\Phi^*w\|_\infty \;;\; w \in \mathcal{H}, \|w\|_\mathcal{H} \leqslant 1\}$ (and the $\infty, \mathcal{H}$ operator norm of a matrix is defined similarly). Given a vector $x_0 \in \mathbb{T}^N$, Φ_{x_0} refers to the linear operator $\mathbb{R}^N \to \mathcal{H}$, with

$$\forall\, \alpha \in \mathbb{R}^N, \quad \Phi_{x_0}\alpha \stackrel{\text{def.}}{=} \Phi(m_{\alpha,x_0}) = \sum_{\nu=1}^{N} \alpha_\nu \varphi(x_{0,\nu}).$$

It may also be seen as the restriction of Φ to measures supported on the set $\{x_{0,\nu} \;;\; \nu \in [\![1,\, N]\!]\}$. A similar notation is adopted for Φ'_{x_0}, resp. $\Phi^{(k)}_{x_0}$ (replacing $\varphi(x_{0,\nu})$ with $\varphi'(x_{0,\nu})$, resp. $\varphi^{(k)}(x_{0,\nu})$). The concatenation of Φ_{x_0} and Φ'_{x_0} is denoted by $\Gamma_{x_0} \stackrel{\text{def.}}{=} \left(\Phi_{x_0} \;\; \Phi'_{x_0} \right)$.

We shall rely on the notion of set convergence. Given a sequence $(C_n)_{n\in\mathbb{N}}$ of subsets of $\mathbb{T}$, we define

$$\limsup_{n\to+\infty} C_n = \left\{ x \in \mathbb{T} \;;\; \liminf_{n\to+\infty} d(x, C_n) = 0 \right\} \tag{10}$$

$$\liminf_{n\to+\infty} C_n = \left\{ x \in \mathbb{T} \;;\; \limsup_{n\to+\infty} d(x, C_n) = 0 \right\} \tag{11}$$

where d is defined by $d(x,C) = \inf_{x'\in C} |x'-x|$ and $|x-x'|$ refers to the distance between x and x' on the torus. If both sets are equal, let C be the corresponding set (then C is necessarily closed), we write

$$\lim_{n\to+\infty} C_n = C. \tag{12}$$

If the sequence $(C_n)_{n\in\mathbb{N}}$ is nondecreasing ($C_n \subset C_{n+1}$), then $\lim_{n\to\infty} C_n = \overline{\bigcup_{n\in\mathbb{N}} C_n}$, and if it is nonincreasing ($C_n \supset C_{n+1}$) then $\lim_{n\to\infty} C_n = \bigcap_{n\in\mathbb{N}} \overline{C_n}$ (where $\overline{C}$ denotes the closure of C). We refer the reader to [26] for more detail about set convergence. We shall also use this notion in Hilbert spaces, with obvious adaptations.

2. Abstract analysis of the Lasso with cone constraint

This section studies a simple variant of the Lasso with cone constraint in an abstract setting. The results stated here shall be useful in Section 3, since this variant turns out to be the Continuous Basis-Pursuit when the degradation operator is the integration of an impulse response and its derivative.

2.1. Notations

Given a parameter $h > 0$, we consider the convex cone generated by the vectors $(1, \frac{h}{2})$ and $(1, -\frac{h}{2})$,

$$\mathcal{C}_h \overset{\text{def.}}{=} \left\{ (c, d) \in \mathbb{R} \times \mathbb{R} \; ; \; c \geqslant 0 \quad \text{and} \quad -c\frac{h}{2} + |d| \leqslant 0 \right\}. \tag{13}$$

We also define the cone $\mathcal{C}_h^G$ as the set of vectors $(a, b) \in \mathbb{R}^G \times \mathbb{R}^G$ such that for all $k \in [\![0, G-1]\!]$ $(a_k, b_k) \in \mathcal{C}_h$.

Now, given a vector $(a_0, b_0) \in \mathcal{C}_h^G$ (*i.e.* $\forall k \in [\![0, G-1]\!]$, $a_{0,k} \geqslant \frac{2}{h}|b_{0,k}|$), we observe $y_0 \overset{\text{def.}}{=} Aa_0 + Bb_0$, where $A : \mathbb{R}^G \to \mathcal{H}$ and $B : \mathbb{R}^G \to \mathcal{H}$ are linear operators, or its noisy version $y = y_0 + w$ where $w \in \mathcal{H}$. To recover (a_0, b_0) from y or y_0, we consider the following reconstruction problems:

$$\min_{(a,b)\in\mathcal{C}_h^G} \frac{1}{2} \|y - Aa - Bb\|_{\mathcal{H}}^2 + \lambda \|a\|_1, \tag{$\mathcal{Q}_\lambda(y)$}$$

and for $\lambda = 0$,

$$\min_{(a,b)\in\mathcal{C}_h^G} \|a\|_1 \quad \text{such that} \quad Aa + Bb = y_0. \tag{$\mathcal{Q}_0(y_0)$}$$

Our main focus is on the support recovery properties of $(\mathcal{Q}_\lambda(y))$. Precisely, we split the "support" of $(a, b) \in \mathcal{C}_h^G$ into several parts:

$$I \overset{\text{def.}}{=} \text{supp}(a) \overset{\text{def.}}{=} \{i \in [\![0, G-1]\!] \; ; \; a_i > 0\} \tag{14}$$
$$= I^{(\mathbf{r})} \cup I^{(\mathbf{l})} \tag{15}$$

$$\text{where } I^{(\mathbf{r})} \overset{\text{def.}}{=} \left\{ i \in I \; ; \; a_i + \frac{2}{h}b_i > 0 \right\}, \quad I^{(\mathbf{l})} \overset{\text{def.}}{=} \left\{ i \in I \; ; \; a_i - \frac{2}{h}b_i > 0 \right\}. \tag{16}$$

In general $I^{(\mathbf{r})} \cap I^{(\mathbf{l})} \neq \emptyset$. If (a_λ, b_λ) is a solution of $(\mathcal{Q}_\lambda(y))$, we say that we have exact support recovery provided that $I^{(\mathbf{r})}(a_\lambda, b_\lambda) = I^{(\mathbf{r})}(a_0, b_0)$ and $I^{(\mathbf{l})}(a_\lambda, b_\lambda) = I^{(\mathbf{l})}(a_0, b_0)$.

Remark 1. The notation $I^{(\mathbf{r})}$, $I^{(\mathbf{l})}$ shall become clearer in the next section. When considering the Continuous Basis-Pursuit on a grid with stepsize $h > 0$, points i in $I^{(\mathbf{r})}$ correspond to Dirac masses which "tend to be on the right", that is they do not coincide with the left half-grid point $ih - \frac{h}{2}$. Similarly, points in $I^{(\mathbf{l})}$ correspond to Dirac masses which "tend to be on the left", as they do not coincide with the right half-grid point $ih + \frac{h}{2}$. In fact, if $i \in I^{(\mathbf{r})} \setminus I^{(\mathbf{l})}$, it correponds to a Dirac mass at the right half-grid point: $\delta_{ih+\frac{h}{2}}$, and if $i \in I^{(\mathbf{l})} \setminus I^{(\mathbf{r})}$, it correponds to a Dirac mass at the left half-grid point: $\delta_{ih-\frac{h}{2}}$. If $i \in I^{(\mathbf{r})} \cap I^{(\mathbf{l})}$, it correponds to a Dirac mass which may belong "freely" to the interval $(ih - \frac{h}{2}, ih + \frac{h}{2})$. **[ToDo: Figure]**

2.2. Parametrization as a positive LASSO

To characterize the solutions of $(\mathcal{Q}_\lambda(y))$ and $(\mathcal{Q}_0(y_0))$, it is convenient to reparametrize the problem as a LASSO with positivity constraint, writing for all

$i \in [\![0, G-1]\!]$,

$$\begin{pmatrix} a_i \\ b_i \end{pmatrix} \overset{\text{def.}}{=} \begin{pmatrix} 1 & 1 \\ \frac{h}{2} & -\frac{h}{2} \end{pmatrix} \begin{pmatrix} r_i \\ l_i \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} r_i \\ l_i \end{pmatrix} = \frac{1}{2} \begin{pmatrix} a_i + \frac{2}{h}b_i \\ a_i - \frac{2}{h}b_i \end{pmatrix}. \tag{17}$$

In the following, we define the linear map

$$H_h : \begin{pmatrix} r \\ l \end{pmatrix} \longmapsto \begin{pmatrix} a \\ b \end{pmatrix} \tag{18}$$

It is clear that $(a_i, b_i) \in \mathcal{C}_h$ if and only if $r_i \geqslant 0$ and $l_i \geqslant 0$. Moreover, given $(a, b) \in \mathcal{C}_h^G$,

$$I^c = \{i \in [\![0, G-1]\!] \; ; \; (r_i, l_i) = (0,0)\},$$
$$I^{(\mathbf{r})} = \{i \in [\![0, G-1]\!] \; ; \; r_i > 0\}, \quad \text{and} \quad I^{(\mathbf{l})} = \{i \in [\![0, G-1]\!] \; ; \; l_i > 0\}.$$

Therefore, Problems $(\mathcal{Q}_\lambda(y))$ and $(\mathcal{Q}_0(y_0))$ are respectively equivalent to the LASSO and Basis Pursuit with positivity constraint:

$$\min_{(r,l) \in (\mathbb{R}_+)^G \times (\mathbb{R}_+)^G} \lambda \left\| \begin{pmatrix} r \\ l \end{pmatrix} \right\|_1 + \frac{1}{2} \left\| y - \mathcal{A}_h \begin{pmatrix} r \\ l \end{pmatrix} \right\|_{\mathcal{H}}^2 \tag{$\tilde{\mathcal{Q}}_\lambda(y)$}$$

and

$$\min_{(r,l) \in (\mathbb{R}_+)^G \times (\mathbb{R}_+)^G} \left\| \begin{pmatrix} r \\ l \end{pmatrix} \right\|_1 \quad \text{such that} \quad \mathcal{A}_h \begin{pmatrix} r \\ l \end{pmatrix} = y_0, \tag{$\tilde{\mathcal{Q}}_0(y_0)$}$$

where $\mathcal{A}_h \overset{\text{def.}}{=} \left(A + \frac{h}{2}B \quad A - \frac{h}{2}B \right) : \mathbb{R}^{2G} \to \mathcal{H}$.

The "support recovery" of (a_0, b_0) through $(\mathcal{Q}_\lambda(y))$ is equivalent to the support recovery of (r_0, l_0) through $(\tilde{\mathcal{Q}}_\lambda(y))$. As described below, the characterization of minimizers and the support recovery properties of the LASSO with positivity constraint $(\tilde{\mathcal{Q}}_\lambda(y))$ are quite similar to those of the classical LASSO described in the companion paper [17, Section 2].

Since the regularization term is of the form

$$J(r, l) \overset{\text{def.}}{=} \sum_{i=0}^{G-1} j(r_i) + \sum_{i=0}^{G-1} j(l_i), \quad \text{with} \quad j(x) \overset{\text{def.}}{=} \sup \{qx \; ; \; q \leqslant 1\} = \begin{cases} x & \text{if } x \geqslant 0, \\ +\infty & \text{otherwise,} \end{cases}$$

its subdifferential is the product of the subdifferentials $\partial j(r_i)$ and $\partial j(l_i)$ for $1 \leqslant i \leqslant G-1$, where

$$\partial j(x) = \begin{cases} \{1\} & \text{if } x > 0, \\ (-\infty, 1] & \text{if } x = 0. \end{cases}$$

That is quite similar to the subdifferential of $|\cdot|$ at $x \in \mathbb{R}$ which is -1, $[-1, 1]$ or 1 if $x < 0$, $x = 0$ or $x > 0$ respectively. Hence, one may adapt the standard results for the LASSO to the LASSO with positivity constraint, simply by replacing the conditions $\|\eta\|_\infty \leqslant 1$ with $\max \eta \leqslant 1$ (and similarly for strict inequalities) wherever they appear. We leave the detail to the reader, and in the following, we use those results freely to derive the properties of the LASSO with cone constraint $(\mathcal{Q}_\lambda(y))$.

2.3. Optimality conditions

The optimality conditions for $(\tilde{\mathcal{Q}}_\lambda(y))$ and $(\tilde{\mathcal{Q}}_0(y_0))$, written in terms of (a_λ, b_λ), yield the following results.

Proposition 1. *Let $y \in \mathcal{H}$, $(a_\lambda, b_\lambda) \in \mathcal{C}_h^G$, and $I = I(a_\lambda, b_\lambda)$. Then (a_λ, b_λ) is a solution to $(\mathcal{Q}_\lambda(y))$ if and only if there exists $p_\lambda \in \mathcal{H}$ such that*

$$\max\left((A^* + \frac{h}{2}B^*)p_\lambda\right) \leqslant 1, \quad and \quad \max\left((A^* - \frac{h}{2}B^*)p_\lambda\right) \leqslant 1, \tag{19}$$

$$(A^*_{I^{(\mathbf{r})}} + \frac{h}{2}B^*_{I^{(\mathbf{r})}})p_\lambda = \mathbb{1}_{I^{(\mathbf{r})}}, \quad and \quad (A^*_{I^{(\mathbf{l})}} - \frac{h}{2}B^*_{I^{(\mathbf{l})}})p_\lambda = \mathbb{1}_{I^{(\mathbf{l})}}, \tag{20}$$

$$\lambda\begin{pmatrix} A^* \\ B^* \end{pmatrix} p_\lambda + \begin{pmatrix} A^* \\ B^* \end{pmatrix}(Aa_\lambda + Bb_\lambda - y) = 0. \tag{21}$$

Similarly, $(a_0, b_0) \in \mathcal{C}_h^G$ is a solution to $(\mathcal{Q}_0(y_0))$ if and only if $Aa_0 + Bb_0 = y_0$ and there exists $p \in \mathcal{H}$ such that

$$\max\left((A^* + \frac{h}{2}B^*)p\right) \leqslant 1, \quad and \quad \max\left((A^* - \frac{h}{2}B^*)p\right) \leqslant 1, \tag{22}$$

$$(A^*_{I^{(\mathbf{r})}} + \frac{h}{2}B^*_{I^{(\mathbf{r})}})p = \mathbb{1}_{I^{(\mathbf{r})}}, \quad and \quad (A^*_{I^{(\mathbf{l})}} - \frac{h}{2}B^*_{I^{(\mathbf{l})}})p \leqslant \mathbb{1}_{I^{(\mathbf{l})}}, \tag{23}$$

where $I = I(a_0, b_0)$.

If the inequalities outside the support are strict, it is possible to ensure the uniqueness of the solution.

Proposition 2. *Under the hypotheses of Proposition 1, if $\left((A + \frac{h}{2}B)_{I^{(\mathbf{r})}} \quad (A - \frac{h}{2}B)_{I^{(\mathbf{l})}}\right)$ has full rank and if p_λ (resp. p) satisfies*

$$\forall k \in I^c, \quad (A^*p_\lambda)_k + \frac{h}{2}|(B^*p_\lambda)_k| < 1, \tag{24}$$

$$\forall i \in I^{(\mathbf{l})} \setminus I^{(\mathbf{r})}, \quad ((A^* + \frac{h}{2}B^*)p_\lambda)_i < 1, \tag{25}$$

$$\forall i \in I^{(\mathbf{r})} \setminus I^{(\mathbf{l})}, \quad ((A^* - \frac{h}{2}B^*)p_\lambda)_i < 1, \tag{26}$$

then (a_λ, b_λ) (resp. (a_0, b_0)) is the unique solution to $(\mathcal{Q}_\lambda(y))$ (resp. $(\mathcal{Q}_0(y_0))$).

One may interpret the optimality conditions of Proposition 1 as the primal-dual relations between the solutions of $(\mathcal{Q}_\lambda(y))$ and $(\mathcal{Q}_0(y_0))$ with the solutions of their respective dual problems

$$\inf_{p \in D} \left\|\frac{y}{\lambda} - p\right\|_{\mathcal{H}}^2 \tag{$\mathcal{E}_\lambda(y)$}$$

$$\sup_{p \in D}\langle y, p\rangle \tag{$\mathcal{E}_0(y)$}$$

$$where \ D \overset{\text{def.}}{=} \left\{p \in \mathcal{H} \ ; \max_{k \in [\![0, G-1]\!]}(A^*p)_k + \frac{h}{2}|(B^*p)_k| \leqslant 1\right\}. \tag{27}$$

Conversely, $p \in \mathcal{H}$ is a solution to $(\mathcal{E}_\lambda(y))$ (resp. $(\mathcal{E}_0(y))$) if and only if $p \in D$ and p satisfies the conditions of Proposition 1.

2.4. Low noise behavior of C-BP

The theorem of Fuchs [22] for the LASSO describes an almost necessary and sufficient condition for the support stability of the problem at low noise. Its adaptation to the positive LASSO is straightforward. However, the criterion that it provides is not satisfied in general, and the support at low noise is strictly larger than $(I^{(\mathbf{r})}(a_0, b_0), I^{(\mathbf{l})}(a_0, b_0))$. We provide below a finer description of that support by studying the minimal norm certificate.

Definition 1 (Minimal norm certificate). *Let* $(a_0, b_0) \in \mathcal{C}_h^G$. *Its minimal norm certificate is* $\bar{\eta}_0 \overset{\text{def.}}{=} \begin{pmatrix} A^* + \frac{h}{2}B^* \\ A^* - \frac{h}{2}B^* \end{pmatrix} p_0$ *where* $p_0 \in \mathcal{H}$ *is the solution to* $(\mathcal{E}_0(y))$ *with minimal norm. The extended support is* $\text{ext}^{(\mathbf{rl})}(a_0, b_0) = \big(\text{ext}^{(\mathbf{r})}(a_0, b_0), \text{ext}^{(\mathbf{l})}(a_0, b_0)\big)$, *where*

$$\text{ext}^{(\mathbf{r})}(a_0, b_0) = \left\{ j \in [\![0, G-1]\!] \ ; \ ((A^* + \frac{h}{2}B^*)p_0)_j = 1 \right\}, \tag{28}$$

$$\text{ext}^{(\mathbf{l})}(a_0, b_0) = \left\{ j \in [\![0, G-1]\!] \ ; \ ((A^* - \frac{h}{2}B^*)p_0)_j = 1 \right\}. \tag{29}$$

From the optimality conditions, if (a_0, b_0) is a solution of $(\mathcal{Q}_0(y_0))$ then $I^{(\mathbf{r})} \subset \text{ext}^{(\mathbf{r})}(a_0, b_0)$ and $I^{(\mathbf{l})} \subset \text{ext}^{(\mathbf{l})}(a_0, b_0)$ (where $I = I(a_0, b_0)$), and p_0 can be characterized as

$$p_0 = \underset{q \in \mathcal{H}}{\text{argmin}} \left\{ \|q\| \ ; \ \begin{pmatrix} A^* + \frac{h}{2}B^* \\ A^* - \frac{h}{2}B^* \end{pmatrix} q \in \partial J(r_0, l_0) \right\}. \tag{30}$$

We are now in position to describe the behavior of $(\mathcal{Q}_\lambda(y))$ at low noise in the generic case. Some details about the proofs of the following Theorem can be found in Appendix B.

Theorem 1. *Let* $(a_0, b_0) \in \mathcal{C}_h^G \setminus \{0\}$ *be an identifiable signal,* $(J^{(\mathbf{r})}, J^{(\mathbf{l})}) \overset{\text{def.}}{=} \text{ext}^{(\mathbf{rl})}(a_0, b_0)$ *such that* $\hat{\mathcal{A}}_h \overset{\text{def.}}{=} \big((A + \frac{h}{2}B)_{J^{(\mathbf{r})}} \quad (A - \frac{h}{2}B)_{J^{(\mathbf{l})}}\big)$ *has full rank. Let*

$$\begin{pmatrix} u_{J^{(\mathbf{r})}} \\ v_{J^{(\mathbf{l})}} \end{pmatrix} \overset{\text{def.}}{=} -(\hat{\mathcal{A}}_h^* \hat{\mathcal{A}}_h)^{-1} s \quad where \quad s \overset{\text{def.}}{=} (1, \dots, 1)^* \in \mathbb{R}^{|J^{(\mathbf{r})}| + |J^{(\mathbf{l})}|},$$

and assume that the following non-degeneracy condition holds

$$\forall j \in J^{(\mathbf{r})} \setminus I^{(\mathbf{r})}, \ u_j > 0, \quad and \quad \forall j \in J^{(\mathbf{l})} \setminus I^{(\mathbf{l})}, \ v_j > 0. \tag{31}$$

Then, there exists constants $C^{(1)} > 0$, $C^{(2)} > 0$ *such that for*

$$\lambda \leqslant C^{(1)} \min \left(\left\{ b_{0,i} + \frac{2}{h}a_{0,i} \ ; \ i \in I^{(\mathbf{r})}(a_0, b_0) \right\} \cup \left\{ b_{0,i} - \frac{2}{h}a_{0,i} \ ; \ i \in I^{(\mathbf{l})}(a_0, b_0) \right\} \right) \tag{32}$$

and $\|w\|_{\mathcal{H}} \leqslant C^{(2)}\lambda$, *the solution* (a_λ, b_λ) *to* $(\mathcal{Q}_\lambda(y))$ *is unique,* $I^{(\mathbf{r})}(a_\lambda, b_\lambda) = J^{(\mathbf{r})}$, $I^{(\mathbf{l})}(a_\lambda, b_\lambda) = J^{(\mathbf{l})}$, *and it reads*

$$\begin{pmatrix} a_\lambda \\ b_\lambda \end{pmatrix} = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + H_h \hat{\mathcal{A}}_h^+ w - \lambda H_h (\hat{\mathcal{A}}_h^* \hat{\mathcal{A}}_h)^{-1} s,$$

where H_h *is defined in* (18).

3. Continuous-Basis Pursuit on thin grids

Now, we turn to the "continuous" inverse problem described in the introduction, and we assume that each α_ν $(1 \leqslant \nu \leqslant N)$ is positive. We aim at recovering m_0 using the Continuous Basis-Pursuit (C-BP) proposed in [19]. Given a grid $\mathcal{G}_n$, the goal is to reconstruct a measure $m = \sum_{i=0}^{G_n-1} a_i \delta_{ih_n+t_i}$ where $t_i \in [-\frac{h_n}{2}, \frac{h_n}{2}]$ which estimates m_0. Applying a Taylor expansion to φ and setting $b_i = t_i a_i$, the authors of [19] are led to solve

$$\min_{(a,b)\in\mathcal{C}_{h_n}^{G_n}} \frac{1}{2} \left\| y - \Phi_{\mathcal{G}_n} a - \Phi'_{\mathcal{G}_n} b \right\|_{\mathcal{H}}^2 + \lambda \|a\|_1 \qquad (\mathcal{Q}_\lambda^n(y))$$

$$\min_{(a,b)\in\mathcal{C}_{h_n}^{G_n}} \|a\|_1 \text{ such that } \Phi_{\mathcal{G}_n} a + \Phi'_{\mathcal{G}_n} b = y_0. \qquad (\mathcal{Q}_0^n(y_0))$$

which are particular instances of $(\mathcal{Q}_\lambda(y))$ and $(\mathcal{Q}_0(y_0))$, with the choice $(A, B) = (\Phi_{\mathcal{G}_n}, \Phi'_{\mathcal{G}_n})$. The dual problems are respectively:

$$\inf_{q\in D^n} \left\| \frac{y}{\lambda} - q \right\|_{\mathcal{H}}^2 \qquad (\mathcal{E}_\lambda^n(y))$$

$$\sup_{q\in D^n} \langle y_0, q \rangle \qquad (\mathcal{E}_0^n(y_0))$$

where $\quad D^n \stackrel{\text{def.}}{=} \left\{ q \in \mathcal{H} \; ; \max_{k\in[\![0, G_n-1]\!]} (\Phi^* q)(kh_n) + \frac{h_n}{2} |(\Phi^* q)'(kh_n)| \leqslant 1 \right\}, \qquad (33)$

To study the behavior of the solutions to these problems as n increases, we shall apply the results of the previous section, and in particular Lemma 3.

3.1. The Positive Beurling LASSO

As we explain below, a natural limit model of C-BP on thin grids is the *positive Beurling* LASSO,

$$\min_{m\in\mathcal{M}^+(\mathbb{T})} \frac{1}{2} \|y - \Phi m\|_{\mathcal{H}}^2 + \lambda m(\mathbb{T}), \qquad (\mathcal{Q}_\lambda^\infty(y))$$

$$\text{and} \quad \min_{m\in\mathcal{M}^+(\mathbb{T})} m(\mathbb{T}) \text{ such that } \Phi m = y_0, \qquad (\mathcal{Q}_0^\infty(y_0))$$

where $\mathcal{M}^+(\mathbb{T})$ refers to the space of positive Radon measures. The indicator function of positive measures plus the total mass may be encoded in the quantity:

$$m(\mathbb{T}) + \iota_{\mathcal{M}^+(\mathbb{T})}(m) = \sup \left\{ \int_{\mathbb{T}} \psi(t) dm(t) \; ; \; \psi \in \mathscr{C}(\mathbb{T}) \text{ and } \sup_{t\in\mathbb{T}} \psi(t) \leqslant 1 \right\}. \qquad (34)$$

The characterization of optimality, the notions of minimal norm certificates and extended support are straightfoward adaptations of those of the Beurling LASSO exposed in the companion paper [17, Section 3]. Again, it essentially amounts to replacing

condition $\|\eta\|_\infty \leqslant 1$ with $\sup_{t\in\mathbb{T}} \eta(t) \leqslant 1$ where $\eta = \Phi^* p$ for $p \in \mathcal{H}$. For instance, up to the addition of a constant, the dual problems to $(\mathcal{Q}_\lambda^\infty(y))$ and $(\mathcal{Q}_0^\infty(y_0))$ are respectively:

$$\inf_{p\in D^\infty} \left\| \frac{y}{\lambda} - p \right\|_\mathcal{H}^2 \qquad (\mathcal{E}_\lambda^\infty(y))$$

$$\sup_{p\in D^\infty} \langle y_0, p \rangle \qquad (\mathcal{E}_0^\infty(y_0))$$

$$\text{where} \quad D^\infty \stackrel{\text{def.}}{=} \left\{ p \in \mathcal{H} \ ; \ \sup_{t\in\mathbb{T}} (\Phi^* p)(t) \leqslant 1 \right\}. \qquad (35)$$

As for the Beurling LASSO, the low noise behavior of $(\mathcal{Q}_\lambda^\infty(y))$ is governed by the minimal norm solution of $(\mathcal{E}_0^\infty(y_0))$ (see [17, Section 3]). That solution being difficult to compute in general, one is led to study a "good candidate" for it, the vanishing derivatives precertificate, which can easily be computed by solving a linear system in the least square sense. In this paper, we are not directly interested in the low noise behavior of $(\mathcal{Q}_\lambda^\infty(y))$ but we shall use this precertificate as an auxiliary quantity, hence we may adopt the following definition.

Definition 2. *Let* $m_0 = \sum_{\nu=1}^N \alpha_{0,\nu} \delta_{x_{0,\nu}}$ *an identifiable measure for the problem* $(\mathcal{Q}_0^\infty(y_0))$ *such that* $\Gamma_{x_0} \stackrel{\text{def.}}{=} \left(\Phi_{x_0} \ \ \Phi'_{x_0} \right)$ *has full rank.*

The vanishing derivatives precertificate is defined as

$$\eta_{V,\infty} \stackrel{\text{def.}}{=} \Phi^* p_{V,\infty} \quad \text{where} \quad p_{V,\infty} \stackrel{\text{def.}}{=} \Gamma_{x_0}^{+,*} \begin{pmatrix} s \\ 0 \end{pmatrix}, \qquad (36)$$

where $s = (1,\ldots,1)^T \in \mathbb{R}^N$ *and* $\Gamma_{x_0}^{+,*} = \Gamma_{x_0} (\Gamma_{x_0}^* \Gamma_{x_0})^{-1}$.

3.2. The Limit Problem for Thin Grids

Let us recall that we obtain a measure from the vector $(a,b) \in \mathcal{C}_{h_n}^{G_n}$ by setting

$$m = \sum_{i=0}^{G_n-1} a_i \delta_{ih_n + b_i/a_i} \qquad (37)$$

with the convention that $b_i/a_i = 0$ if $a_i = 0$. It should be noticed that $b_i/a_i \in [-\frac{h_n}{2}, \frac{h_n}{2}]$.

In this section we prove the Γ-convergence of $(\mathcal{Q}_\lambda^n(y))$ towards $(\mathcal{Q}_\lambda^\infty(y))$. Since all the solutions to $(\mathcal{Q}_\lambda^n(y))$ belong to $X_+ \stackrel{\text{def.}}{=} \{m \in \mathcal{M}^+(\mathbb{T}) \ ; \ \lambda|m|(\mathbb{T}) \leqslant \frac{1}{2} \|y\|_\mathcal{H}^2\}$, we may restrict the problems to X_+, which is metrizable for the weak* topology. Hence, we may use the following formulation of Γ-convergence valid in metric spaces (see [10]).

Definition 3. *We say that the Problem* $(\mathcal{Q}_\lambda^n(y))$ Γ-*converges towards Problem* $(\mathcal{Q}_\lambda^\infty(y))$ *if, for all* $m \in X_+$, *the following conditions hold*

- *(Liminf inequality) for any sequence of measures* $(m_n)_{n\in\mathbb{N}} \in X_+^\mathbb{N}$ *of the form* (37) *with* $(a^{(n)}, b^{(n)}) \in \mathcal{C}_{h_n}^{G_n}$ *such that* m_n *weakly* converges towards* m,

$$\liminf_{n\to+\infty} \left(\lambda\|a^{(n)}\|_1 + \frac{1}{2} \left\| \Phi_{\mathcal{G}_n} a^{(n)} + \Phi'_{\mathcal{G}_n} b^{(n)} - y \right\|_\mathcal{H}^2 \right) \geqslant \lambda m(\mathbb{T}) + \frac{1}{2} \|\Phi m - y\|_\mathcal{H}^2.$$

- *(Limsup inequality) there exists a sequence of measures $(m_n)_{n \in \mathbb{N}} \in X_+^{\mathbb{N}}$ of the form* (37) *with $(a^{(n)}, b^{(n)}) \in \mathcal{C}_{h_n}^{G_n}$ such that m_n weakly* converges towards m and*

$$\limsup_{n \to +\infty} \left(\lambda \|a^{(n)}\|_1 + \frac{1}{2} \left\| \Phi_{\mathcal{G}_n} a^{(n)} + \Phi'_{\mathcal{G}_n} b^{(n)} - y \right\|_{\mathcal{H}}^2 \right) \leqslant \lambda m(\mathbb{T}) + \frac{1}{2} \|\Phi m - y\|_{\mathcal{H}}^2.$$

The following proposition, which is proved in Appendix C.1, states the Γ-convergence of the model and its consequences.

Proposition 3. *The Problem $(\mathcal{Q}_\lambda^n(y))$ Γ-converges towards $(\mathcal{Q}_\lambda^\infty(y))$, and*

$$\lim_{n \to +\infty} \left(\inf \left(\mathcal{Q}_\lambda^n(y) \right) \right) = \inf \left(\mathcal{Q}_\lambda^\infty(y) \right). \tag{38}$$

Each sequence $(m_{\lambda,n})_{n \in \mathbb{N}}$ such that $m_{\lambda,n}$ is a minimizer of $(\mathcal{Q}_\lambda^n(y))$ has accumulation points (for the weak*) topology, and each of these accumulation points is a minimizer of $(\mathcal{Q}_\lambda^\infty(y))$.

In particular, if the solution $m_{\lambda,\infty}$ to $(\mathcal{Q}_\lambda^\infty(y))$ is unique, the whole sequence $(m_{\lambda,n})_{n \in \mathbb{N}}$ converges towards $m_{\lambda,\infty}$.

4. Convergence of the support

Though Proposition 3 states the convergence of the solutions of $(\mathcal{Q}_\lambda^n(y))$ towards those of $(\mathcal{Q}_\lambda^\infty(y))$, it does not describe the supports of the solutions. We now study the convergence of those supports using dual certificates and the optimality conditions (Proposition 1). In this continuous context, a dual certificate is determined by a function $\eta = \Phi^* p \in \mathscr{C}(\mathbb{T})$ where $q \in \mathcal{H}$, and if (a, b) is a solution to $(\mathcal{Q}_\lambda^n(y))$,

$$I^{(\mathbf{r})} = \left\{ i \in [\![0, G_n - 1]\!] \; ; \; b_i > -\frac{h_n}{2} a_i \right\} \subset \left\{ i \in [\![0, G_n - 1]\!] \; ; \; \left(\eta + \frac{h_n}{2} \eta' \right)(ih_n) = 1 \right\},$$

$$I^{(\mathbf{l})} = \left\{ i \in [\![0, G_n - 1]\!] \; ; \; b_i < \frac{h_n}{2} a_i \right\} \subset \left\{ i \in [\![0, G_n - 1]\!] \; ; \; \left(\eta - \frac{h_n}{2} \eta' \right)(ih_n) = 1 \right\}.$$

To sum up, we shall exploit the following observations

- if $\left(\eta + \frac{h_n}{2} \eta' \right)(ih_n) = 1$ but $\left(\eta - \frac{h_n}{2} \eta' \right)(ih_n) < 1$, a spike may appear at $ih_n + \frac{h_n}{2}$,
- if $\left(\eta - \frac{h_n}{2} \eta' \right)(ih_n) = 1$ but $\left(\eta + \frac{h_n}{2} \eta' \right)(ih_n) < 1$, a spike may appear at $ih_n - \frac{h_n}{2}$,
- if $\left(\eta + \frac{h_n}{2} \eta' \right)(ih_n) = 1$ and $\left(\eta - \frac{h_n}{2} \eta' \right)(ih_n) = 1$, a spike may appear anywhere in the interval $[ih_n - \frac{h_n}{2}, ih_n + \frac{h_n}{2}]$.

In the next two paragraphs, we describe the behavior of the support. Those results rely on auxiliary Lemmas in Appendix D.

4.1. Asymptotics of the Support for Fixed $\lambda > 0$

The following proposition relies on the convergence of the dual certificates (see Lemma 6 in Appendix). It states that in the generic case, one may observe up to two pairs of spikes for each spike of the solution of the positive Beurling-lasso. As before, r is chosen such that $0 < r < \frac{1}{2} \min_{\nu \neq \nu'} |x_\nu - x_{\nu'}|$.

Proposition 4. *Let $\lambda > 0$, and assume that there exists a solution $m_{\lambda,\infty}$ to $(\mathcal{Q}_\lambda^\infty(y))$ which is a sum of a finite number of Dirac masses, $m_{\lambda,\infty} = \sum_{\nu=1}^{N} \alpha_\nu \delta_{x_\nu}$ where $\alpha_\nu > 0$. Assume that $\eta_{\lambda,\infty}$ satisfies $|\eta_{\lambda,\infty}(t)| < 1$ for all $t \in \mathbb{T} \setminus \{x_1, \ldots, x_N\}$.*

Then any sequence of solution $m_{\lambda,n} = \sum_{i=0}^{G_n-1} a_{\lambda,i} \delta_{ih_n + b_{\lambda,i}/a_{\lambda,i}}$ to $(\mathcal{Q}_\lambda^n(y))$ satisfies

$$\limsup_{n \to +\infty} (\operatorname{supp} m_{\lambda,n}) \subset \{x_1, \ldots x_N\}.$$

If, moreover, $m_{\lambda,\infty}$ is the unique solution to $(\mathcal{Q}_\lambda^\infty(y))$,

$$\lim_{n \to +\infty} (\operatorname{supp}(m_{\lambda,n})) = \{x_1, \ldots x_N\}. \tag{39}$$

If, additionally, $(\eta_{\lambda,\infty})''(x_\nu) \neq 0$ for some $\nu \in \{1, \ldots, N\}$, then for all n large enough, the restriction of $m_{\lambda,n}$ to $(x_\nu - r, x_\nu + r)$ is a sum of Dirac masses whose configuration is given in Table 1, and if $(\eta_{\lambda,\infty})^{(3)}(x_\nu) \neq 0$, then only the cases indicated with $(*)$ may appear.

Remark 2. Proposition 4 states that the support of the C-BP on thin grids actually depends on the properties of the dual certificate $\eta_{\lambda,\infty}$ of the (positive) Beurling LASSO. The condition $\eta_{\lambda,\infty}''(x_\nu) \neq 0$ seems to be overwhelming, or generic, and it is ensured for instance if λ is small and the Non-Degenerate Source Condition holds (see [16]). As for the condition $\eta_{\lambda,\infty}^{(3)}(x_\nu) \neq 0$, it also seems to be generic, as there is nothing to impose $\eta_{\lambda,\infty}^{(3)}(x_\nu) = 0$ in the positive Beurling LASSO. As a result, in practice, *one does not observe all the configurations given in Table 1*, and *only the cases indicated with $(*)$ appear*, the case of two spikes being again overwhelming.

This means that when approximating the positive Beurling LASSO with the Continuous Basis-Pursuit, one generally sees two spikes instead of one, and those spikes are at successive half-grid points: $(ih + \frac{h}{2}, (i+1)h + \frac{h}{2})$ or $(ih - \frac{h}{2}, (i+1)h - \frac{h}{2})$.

4.2. Asymptotic of the Low Noise Support

Now, we focus on the low noise behavior of the model. This analysis is more difficult in whole generality, since it involves the minimal norm solutions of linear problems, for which it is non-trivial to pass to the limit. Therefore, we are led to assume that $\{x_{0,1}, \ldots, x_{0,N}\} \subset \mathcal{G}_n$ for n large enough, and the measure now reads $m_0 = \sum_{i=0}^{G_n-1} a_{0,i} \delta_{ih_n}$.

Even so, the minimal norm solutions of $(\mathcal{E}_0^n(y_0))$ do not converge towards the minimal norm solution of $(\mathcal{E}_0^\infty(y_0))$, and we introduce a new variational problem to carry the study further.

Number of Dirac masses	Saturations of the certificates $(S_{n,\nu}^{(r)}/S_{n,\nu}^{(l)})$	Possible Dirac Locations
One	$\{ih_n\}/\emptyset$ or $\emptyset/\{ih_n\}$	$ih_n + \varepsilon_n \frac{h_n}{2}$, with $\varepsilon_n \in \{-1,1\}$ $\qquad(*)$
	$\{ih_n\}/\{ih_n\}$	$ih_n + t_i$, with $-\frac{h_n}{2} \leqslant t_i \leqslant \frac{h_n}{2}$
Two	$\{(i-1)h_n, ih_n\}/\emptyset$ or $\emptyset/\{ih_n,(i+1)h_n\}$	$\left((i-\varepsilon_n)h_n + \varepsilon_n \frac{h_n}{2}, ih_n + \varepsilon_n \frac{h_n}{2} \right)$, with $\varepsilon_n \in \{-1,1\}$ $\qquad(*)$
	$\{ih_n\}/\{jh_n\}$	$\left(ih_n + \frac{h_n}{2}, jh_n - \frac{h_n}{2} \right)$, $i < j$
	$\{ih_n\}/\{jh_n,(i+1)h_n\}$ or $\{(i-1)h_n, ih_n\}/\{ih_n\}$	$\left((i-\varepsilon_n)h_n + \varepsilon_n \frac{h_n}{2}, ih_n + t_i \right)$, $\varepsilon_n \in \{-1,1\}$, $-\frac{h_n}{2} \leqslant t_i \leqslant \frac{h_n}{2}$
Three	$\{ih_n\}/\{jh_n,(j+1)h_n\}$	$\left(ih_n + \frac{h_n}{2}, jh_n - \frac{h_n}{2}, (j+1)h_n - \frac{h_n}{2} \right)$, with $i < j$
	$\{(i-1)h_n, ih_n\}/\{jh_n\}$	$\left((i-1)h_n + \frac{h_n}{2}, ih_n + \frac{h_n}{2}, jh_n - \frac{h_n}{2} \right)$, with $i < j$
	$\{(i-1)h_n, ih_n\}/\{ih_n,(i+1)h_n\}$	$\left((i-1)h_n + \frac{h_n}{2}, ih_n + t_i, (i+1)h_n - \frac{h_n}{2} \right)$, $-\frac{h_n}{2} \leqslant t_i \leqslant \frac{h_n}{2}$
Four	$\{(i-1)h_n, ih_n\}/\{jh_n,(j+1)h_n\}$	$\left((i-1)h_n + \frac{h_n}{2}, ih_n + \frac{h_n}{2}, jh_n - \frac{h_n}{2}, (j+1)h_n - \frac{h_n}{2} \right)$, $i < j$

Table 1: Number of Dirac masses that may appear if $\eta''_{\lambda,\infty}(x_\nu) \neq 0$. For the sake of the simplicity of the table, and since we focus on the saturations of dual certificates, we regard sums like $\delta_{ih_n+h_n/2} + \delta_{(i+1)h_n-h_n/2}$ as "two" Dirac masses.

Definition 4 (Third derivative precertificate). *Given $m_0 \in \mathcal{M}(\mathbb{T})$, we define the third derivative precertificate as $\eta_T \overset{\text{def.}}{=} \Phi^* p_T$ where*

$$p_T \overset{\text{def.}}{=} \underset{p \in \mathcal{H}}{\text{argmin}} \left\{ \|p\|_{\mathcal{H}} \,|\, ; \forall i \in \{1,\dots,N\},\ (\Phi^*p)(x_{0,i}) = 1, \right.$$

$$\left. (\Phi^*p)'(x_{0,i}) = 0 \ and \ (\Phi^*p)^{(3)}(x_{0,i}) = 0 \right\}, \qquad (40)$$

whenever the above set is not empty.

It is clear that the set defined in (40) is a closed convex set. It is nonempty for instance if the conditions of Lemma 7 hold. Note that p_T corresponds to a quadratic minimization under linear constraint, and can hence be computed by solving a linear system,

$$p_T = \begin{pmatrix} \Gamma_{x_0}^* \\ \Phi_{x_0}^{(3)*} \end{pmatrix}^+ \left(\begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix} \\ 0 \right) = \Gamma_{x_0}^{+,*} \begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix} - \tilde{\Pi}\Phi_{x_0}^{(3)*}(\Phi_{x_0}^{(3)*}\tilde{\Pi}\Phi_{x_0}^{(3)*})^{-1}\Phi_{x_0}^{(3)*}\Gamma_{x_0}^{+,*}\begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix}$$

$$(41)$$

$$= p_{V,\infty} - \tilde{\Pi}\Phi_{x_0}^{(3)*}(\Phi_{x_0}^{(3)*}\tilde{\Pi}\Phi_{x_0}^{(3)*})^{-1}\Phi_{x_0}^{(3)*}p_{V,\infty}, \qquad (42)$$

where $\tilde{\Pi}$ is the orthogonal projector onto $(\operatorname{Im}\Gamma_{x_0})^{\perp}$, and $\Gamma_{x_0} = \begin{pmatrix} \Phi_{x_0} & \Phi'_{x_0} \end{pmatrix}$.

Definition 5 (Twice Non-Degenerate Source Condition). *We say that m_0 satisfies the Twice Non-Degenerate Source Condition (TNDSC) if p_T in (40) is well defined and if it satisfies, for $\eta_T = \Phi^* p_T$,*

$$\forall t \in \mathbb{T} \setminus \{x_{0,1}, \ldots, x_{0,N}\}, \ \eta_T(t) < 1,$$

$$\forall \nu \in \{1, \ldots, N\}, \quad \eta_T''(x_{0,\nu}) < 0 \quad and \quad \eta_T^{(4)}(x_{0,\nu}) > 0.$$

We are now in position to provide a sufficient condition for the spikes to appear in pair, with a prediction on the location of the neighbor. The proof of this Theorem can be found in Appendix D.3.

Theorem 2. *Assume that the operator $\begin{pmatrix} \Phi_{x_0} & \Phi'_{x_0} & \Phi_{x_0}^{(3)} \end{pmatrix}$ has full rank and that the Twice Non Degenerate Source condition (Definiton 5) holds. Moreover, assume that all the components of the natural shift*

$$\rho \stackrel{\text{def.}}{=} (\Phi_{x_0}^{(3)*}\tilde{\Pi}\Phi_{x_0}^{(3)})^{-1}\Phi_{x_0}^{(3)*}\Gamma_{x_0}^{+,*} \begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix} \tag{43}$$

are nonzero. Then, for n large enough the extended support of m_0 on the grid $\mathcal{G}_n$ has the form

$$\operatorname{ext}_{\mathrm{n}}^{(\mathbf{r})}(m_0) = \{x_{0,1}, \ldots, x_{0,N}\} \cup \{x_{0,\nu} - h_n \ ; \ \nu \in [\![1, N]\!] \quad and \quad \rho_\nu > 0\}$$

$$\operatorname{ext}_{\mathrm{n}}^{(\mathbf{l})}(m_0) = \{x_{0,1}, \ldots, x_{0,N}\} \cup \{x_{0,\nu} + h_n \ ; \ \nu \in [\![1, N]\!] \quad and \quad \rho_\nu < 0\}.$$

Combining the above Theorem with Theorem 1, we get

Corollary 1. *Under the hypotheses of Theorem 2, for n large enough, there exists constants $C_n^{(1)} > 0$, $C_n^{(2)} > 0$ such that for $\lambda \leqslant C_n^{(1)} \min_{1 \leqslant \nu \leqslant N} |\alpha_{0,\nu}|$, and for all $w \in \mathcal{H}$ such that $\|w\|_{\mathcal{H}} \leqslant C_n^{(2)}\lambda$, the solution to $(\mathcal{Q}_\lambda^n(y))$ is unique, and reads $m_\lambda = \sum_{\nu=1}^N (\alpha_{\lambda,\nu}\delta_{x_{0,\nu}+t_\nu} + \beta_{\lambda,\nu}\delta_{x_{0,\nu}+\varepsilon_\nu h_n})$, where*

$$-h_n/2 < t_\nu < h_n/2 \quad and \quad \varepsilon \stackrel{\text{def.}}{=} -\operatorname{sign}(\rho).$$

Moreover, the constants $C_n^{(1)}, C_n^{(2)}$ can be chosen as $C_n^{(1)} = O(h_n{}^3)$ and $C_n^{(2)} = O(1)$, and one has

$$\left\| \begin{pmatrix} \alpha_\lambda \\ \beta_\lambda \end{pmatrix} - \begin{pmatrix} \alpha_0 \\ 0 \end{pmatrix} \right\|_\infty = O\left(\frac{w}{h_n{}^3}, \frac{\lambda}{h_n{}^3} \right). \tag{44}$$

The proof is given in Appendix D.4.

5. Numerical illustrations

In this section, we illustrate the relevance of our analysis to gain a precise understanding of the recovery performance of ℓ^1-type methods (Lasso and C-BP) for deconvolution. The code to reproduce these numerical experiments is available online‡.

5.1. Convergence of pre-certificates

In this section and in Section 5.2, we consider the deconvolution problem, that is $\varphi(x) = \tilde{\varphi}(\cdot - x)$, in the case where $\tilde{\varphi}$ is an ideal filter, i.e. whose Fourier coefficients

$$\forall\, k \in \mathbb{Z}, \quad \hat{\tilde{\varphi}}(k) \stackrel{\text{def.}}{=} \int_{\mathbb{T}} \tilde{\varphi}(t)e^{-2\mathrm{i}\pi kt}\mathrm{d}t$$

satisfy $\hat{\tilde{\varphi}}(k) = 1$ if $k \in \{-f_c, \ldots, f_c\}$ and $\hat{\tilde{\varphi}}(k) = 0$ otherwise. This allows us to implement exactly the Φ operator appearing in the LASSO and C-BP problem since $\mathrm{Im}(\Phi)$ is a finite dimensional space of dimension $Q = 2f_c + 1$, i.e. it can be represented using a matrix of size (Q, P) when evaluated on a grid of P points. In Figures 1 and 2 we used $f_c = 10$.
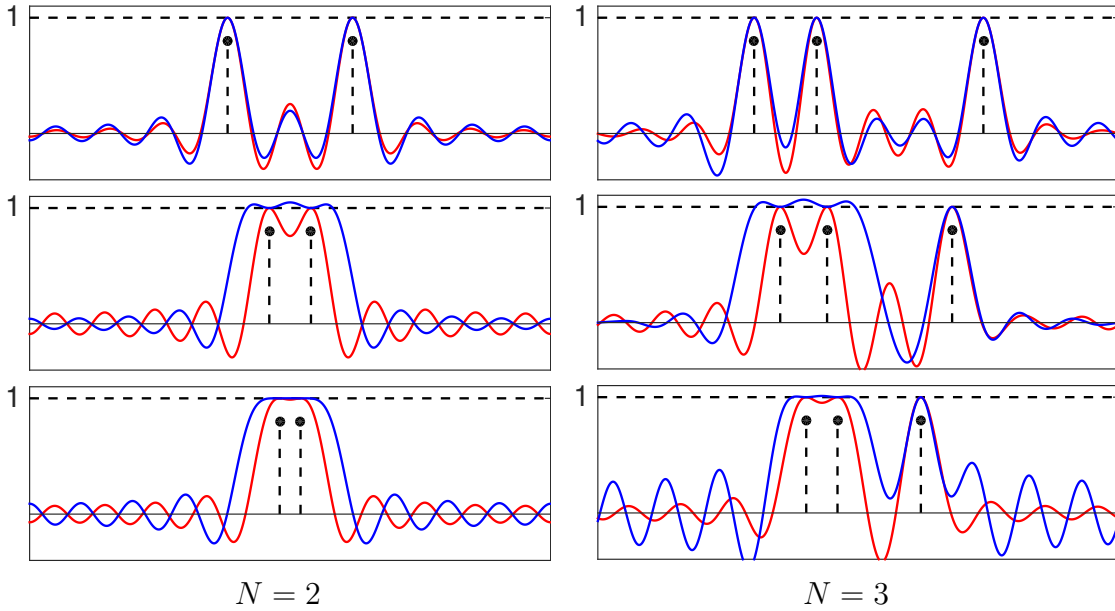

Figure 1: Display of $\eta_{V,\infty}$ (red) and η_T (blue) pre-certificate for different input positive measures m_0 (showed as black dots to symbolize the position of the Diracs).

Figure 1 illustrates for the case of two ($N = 2$) and three ($N = 3$) spikes the behavior of the vanishing pre-certificate $\eta_{V,\infty}$ (see Definition 2) useful to analyze LASSO/BLASSO problems and of the pre-certificate η_T (see Definition 4) useful to analyze C-BP problems.

We first notice that for all the (positive) input measures (i.e. whatever the spacing between the Diracs), $\eta_{V,\infty}$ is always a non-degenerate certificate (in the sense of Proposition 2), meaning that one actually has $\eta_{V,\infty} = \eta_{0,\infty}$ (where the minimal norm certificate $\eta_{0,\infty}$ is the minimal norm solution of $(\mathcal{D}_0^\infty(y_0))$ in [17]). This empirical finding is the subject of another recent work on the asymptotic of sparse recovery of positive measures when the spacing between the Diracs tends to zero [12]. Since $\eta_{0,\infty}$ is non-degenerate, one can thus apply [17, Theorem 2] to analyze the extended support of the LASSO on a thin grid (see Section 5.2 below for a numerical illustration).

For the C-BP problem, the situation is however more contrasted. We observe that when the Dirac masses are separated enough (first row) then the pre-certificate μ_T is a valid certificate, meaning the the Twice Non-Degenerate Source Condition (see Definition 5) holds. This means that Theorem 2 can be applied to analyze the extended support of C-BP on a thin grid (see Section 5.2 below for a numerical illustration). But when the Dirac masses are too close (second and third rows), one has $\|\eta_T\|_\infty > 1$, so that one cannot ensure the support stability of the C-BP solution with our result.

5.2. Extended support for deconvolution on a thin grid

We still consider the case of an ideal low pass filter. Figure 2 displays the evolution, as a function of λ (in abscissa) of the solution a_λ of the Lasso (Eq. (6)) and of the solution (a_λ, b_λ) of the C-BP ($\mathcal{Q}_\lambda^n(y)$), on a thin grid. We consider here the case of an input measure with two nearby Diracs (displayed as red/blue dots in the upper-left part of the Figure) and when there is no noise, i.e. $w = 0$. Each 1-D curve (either plain or dashed) represents the evolution of a single coefficient, e.g. $(a_\lambda)_i$, for some index i (only non-zero coefficients are displayed).
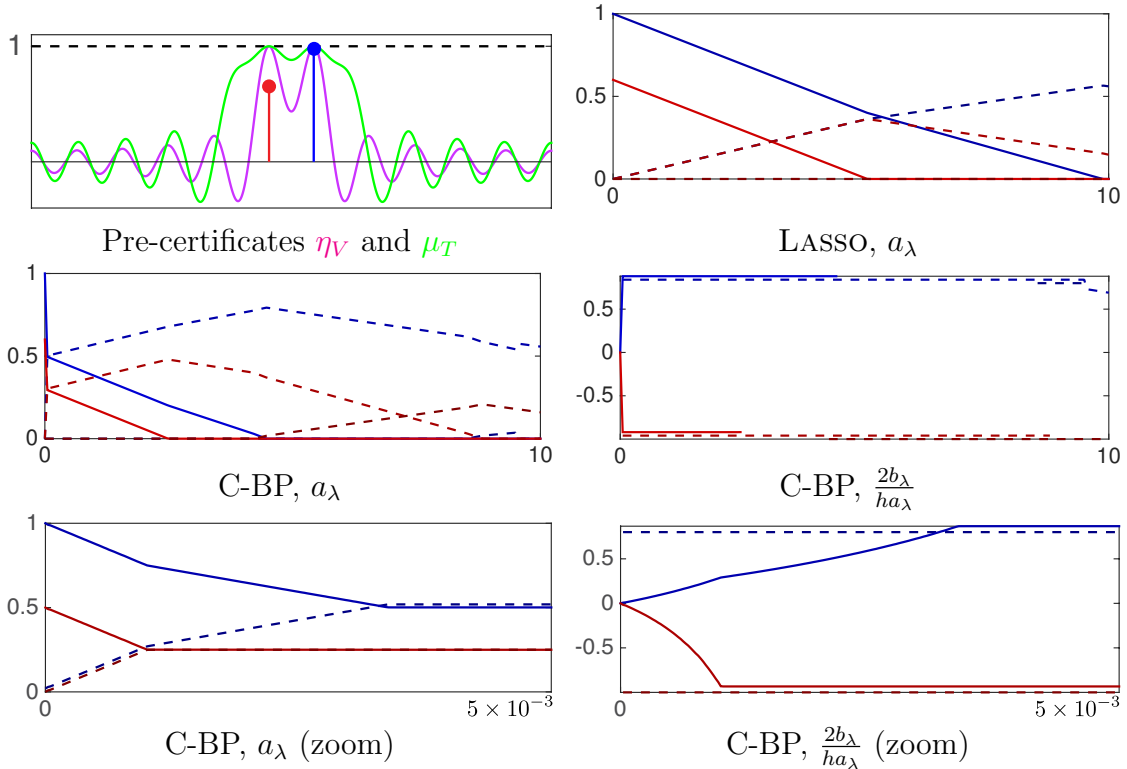

Figure 2: Display of the evolution as a function of λ of the solutions of the Lasso and C-BP problems. Note that dashed curved have been (artificially) slightly shifted to avoid that they overlap with the plain curve.

The solutions path $\lambda \mapsto a_\lambda$ (for Lasso) and $\lambda \mapsto (a_\lambda, b_\lambda)$ (for C-BP) are continuous and piecewise affine, which is to be expected since the regularizations (ℓ^1 and ℓ^1 under

conic constraints) are polyhedral. The upper-left plot in the figure displays the pre-certificate $\eta_{V,\infty}$ (in magenta, see Definition 2) and η_T (in green, see Definition 4). This shows graphically that these two precertificates are non-degenerate (in the sense of [17, Proposition 3] and Definition 5) so that the conclusions of [17, Theorem 2] and Theorem 2 hold, hence precisely describing the evolution of the solution on the extended support when λ is small. On these graphs, this corresponds to the first segment of the corresponding piecewise affine paths.

The behavior for BP agrees with our analysis. As predicted by [17, Corollary 1], there exists a range of values $0 < \lambda < \lambda_0$ on which the solution is exactly supported on the extended support J, which is composed of four spikes (the plain curve corresponds to the support I and the dashed curve corresponds to $J \backslash I$). Also, as predicted by [17, Proposition 7] in the case $w = 0$, we verify that $\lambda_0 = O(h_n)$ and that the Lipschitz constant of $\lambda \mapsto a_\lambda$ is of order $O(1/h_n)$.

In sharp contrast, the behavior for C-BP is less regular, since the range $0 < \lambda < \lambda_0$ on which the solution is supported on the extended support is shorter, as it can be clearly seen on the zoom for very small values of λ. This is in agreement with Proposition **??** which shows that λ_0 is of the order of $O(h_n{}^3)$ and that the Lipschitz constant of $\lambda \mapsto (a_\lambda, b_\lambda)$ is of order $O(1/h_n{}^3)$. On this range of small λ, as predicted by Theorem 2, the support of the solutions (which correspond to the extended support J described in Theorem 2) is composed of one pair of neighboring spikes for each original spike. For indices on the support $i \in I$, one has $|(b_\lambda)_i|/(a_\lambda)_i < h/2$ (the constraint is non-saturating, and the spike moves "freely" inside $(ih - \frac{h}{2}, ih + \frac{h}{2})$) while for indices on the extended part $i \in J \backslash I$, one has $|(b_\lambda)_i|/(a_\lambda)_i = h/2$ (the constraint is saturating, the spikes are fixed at half-grid points). Another part of the path is interesting, for λ not so small (say $\lambda > \lambda_1$), which is in fact the prominent regime in the non-zoomed figure. For this range of λ, there is still a pair of spikes for each original spike, but this time both spikes saturate, on same side. This observation should be related to Proposition 4 and Remark 2 which predict that, in the case where $\eta_{\lambda,\infty}^{(3)}(x_{\lambda,\nu}) \neq 0$, the C-BP yields either one spike or a pair of spikes with the same shift (the latter case is in fact overwhelming).

Conclusion

In this work, we have provided a precise analysis of the properties of the solution path of ℓ^1-type variational problems in the low-noise regime. This includes in particular the LASSO and the C-BP problems. A particular attention has been paid to the support set of this path, which in general cannot be expected to match the one of the sought after solution. Two striking examples support the relevance of this approach. For the deconvolution problem, we showed theoretically that in general this support is not stable, and we were able to derive in closed form the solution of the "extended support" that is twice larger, but is stable. In the compressed sensing scenario (i.e. when the operator of the inverse problem is random), we showed numerically how to leverage our theoretical findings and analyze the growth of the extended support size as the number of

measurements diminishes. This analysis opens the doors for many new developments to better understand this extended support, both for deterministic operators (e.g. Radon transform in medical imaging) and random ones.

Acknowledgements

Appendix A. Asymptotic expansion of the inverse of a Gram matrix

In this Appendix, we gather some useful lemmas on the asymptotic behavior of inverse Gram matrices. The proof of the following lemma can be found in [17].

Lemma 1. *Let $A\colon \mathbb{R}^N \to \mathcal{H}$, $B\colon \mathbb{R}^N \to \mathbb{R}^N$ be linear operators such that A has full rank and B is invertible. Then $(AB)^+ = B^{-1}A^+$.*

Lemma 2. *Let $A, B, C, C_h\colon \mathbb{R}^N \to \mathcal{H}$ be linear operators such that $C_h = C + o(1)$ for $h > 0$, and that $\begin{pmatrix} A & B & C \end{pmatrix}$ has full rank. Let $\tilde{\Pi}$ be the orthogonal projector onto $(\operatorname{Im} \begin{pmatrix} A & B \end{pmatrix})^{\perp}$, and let*

$$G_h \overset{\text{def.}}{=} \begin{pmatrix} (A + \frac{h}{2}B)^* \\ (A - \frac{h}{2}B)^* \\ (A + \frac{h}{2}B + h^3 C_h)^* \end{pmatrix} \begin{pmatrix} A + \frac{h}{2}B & A - \frac{h}{2}B & A + \frac{h}{2}B + h^3 C_h \end{pmatrix}.$$

Then for $h > 0$ small enough, G_h and $C^*\tilde{\Pi}C$ are invertible, and

$$G_h^{-1} \begin{pmatrix} \mathbb{1}_N \\ \mathbb{1}_N \\ \mathbb{1}_N \end{pmatrix} = -\frac{1}{h^3} \begin{pmatrix} -I_N \\ 0 \\ I_N \end{pmatrix} (C^*\tilde{\Pi}C)^{-1}C^* \begin{pmatrix} A & B \end{pmatrix}^{+,*} \begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix} + o\left(\frac{1}{h^3}\right)$$

$$\begin{pmatrix} A + \frac{h}{2}B & A - \frac{h}{2}B & A + \frac{h}{2}B + h^3 C_h \end{pmatrix}^+ = \frac{1}{h^3} \begin{pmatrix} -(C^*\tilde{\Pi}C)^{-1}C^*\tilde{\Pi} \\ 0 \\ (C^*\tilde{\Pi}C)^{-1}C^*\tilde{\Pi} \end{pmatrix} + o\left(\frac{1}{h^3}\right),$$

but

$$\begin{pmatrix} A^* + \frac{h}{2}B^* \\ A^* - \frac{h}{2}B^* \\ A^* + \frac{h}{2}B^* + h^3 C_h^* \end{pmatrix}^+ \begin{pmatrix} \mathbb{1}_N \\ \mathbb{1}_N \\ \mathbb{1}_N \end{pmatrix} = \begin{pmatrix} A^* \\ B^* \end{pmatrix}^+ \begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix} - \tilde{\Pi}C(C^*\tilde{\Pi}C)^{-1}C^* \begin{pmatrix} A^* \\ B^* \end{pmatrix}^+ \begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix}$$

$$+ o(1).$$

Proof. Observe that

$$
\begin{pmatrix} A + \frac{h}{2}B & A - \frac{h}{2}B & A + \frac{h}{2}B + h^3 C_h \end{pmatrix} = \begin{pmatrix} A & B & C_h \end{pmatrix} \operatorname{diag}\left(1, \frac{h}{2}, h^3\right) \begin{pmatrix} I_N & I_N & I_N \\ I_N & -I_N & I_N \\ 0 & 0 & I_N \end{pmatrix}
$$

As a result, for $h > 0$ small enough G_h is invertible and

$$
G_h^{-1} = \begin{pmatrix} \frac{1}{2}I_N & \frac{1}{2}I_N & -I_N \\ \frac{1}{2}I_N & -\frac{1}{2}I_N & 0 \\ 0 & 0 & I_N \end{pmatrix} \operatorname{diag}\left(1, \frac{2}{h}, \frac{1}{h^3}\right) \begin{pmatrix} A^*A & A^*B & A^*C_h \\ B^*A & B^*B & B^*C_h \\ C_h^*A & C_h^*B & C_h^*C_h \end{pmatrix}^{-1}
$$

$$
\times \operatorname{diag}\left(1, \frac{2}{h}, \frac{1}{h^3}\right) \begin{pmatrix} \frac{1}{2}I_N & \frac{1}{2}I_N & 0 \\ \frac{1}{2}I_N & -\frac{1}{2}I_N & 0 \\ -I_N & 0 & I_N \end{pmatrix}
$$

the middle matrix being invertible from the full rank assumption on $\begin{pmatrix} A & B & C \end{pmatrix}$.

Moreover, writing $\tilde{\Gamma} \overset{\text{def.}}{=} \begin{pmatrix} A & B \end{pmatrix}$ and $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \overset{\text{def.}}{=} \begin{pmatrix} \tilde{\Gamma}^*\tilde{\Gamma} & \tilde{\Gamma}^*C_h \\ C_h^*\tilde{\Gamma} & C_h^*C_h \end{pmatrix}$, we obtain

$$
\begin{pmatrix} A^*A & A^*B & A^*C_h \\ B^*A & B^*B & B^*C_h \\ C_h^*A & C_h^*B & C_h^*C_h \end{pmatrix}^{-1} = \begin{pmatrix} u & -a^{-1}bS^{-1} \\ -S^{-1}ca^{-1} & S^{-1} \end{pmatrix}
$$

where $u \overset{\text{def.}}{=} a^{-1} + a^{-1}bS^{-1}ca^{-1}$, $S \overset{\text{def.}}{=} d - ca^{-1}b = C_h^*\tilde{\Pi}C_h$, $a^{-1}bS^{-1} = (\tilde{\Gamma}^*\tilde{\Gamma})^{-1}\tilde{\Gamma}^*C_h(C_h^*\tilde{\Pi}C_h)^{-1}$, $S^{-1}ca^{-1} = (C_h^*\tilde{\Pi}C_h)^{-1}C_h^*\tilde{\Gamma}(\tilde{\Gamma}^*\tilde{\Gamma})^{-1}$, and $\tilde{\Pi}$ is the orthogonal projector onto $(\operatorname{Im}\tilde{\Gamma})^{\perp}$. Thus

$$
G_h^{-1}\begin{pmatrix} \mathbb{1}_N \\ \mathbb{1}_N \\ \mathbb{1}_N \end{pmatrix} = \begin{pmatrix} \frac{1}{2}I_N & \frac{1}{2}I_N & -I_N \\ \frac{1}{2}I_N & -\frac{1}{2}I_N & 0 \\ 0 & 0 & I_N \end{pmatrix} \operatorname{diag}\left(1, \frac{2}{h}, \frac{1}{h^3}\right) \begin{pmatrix} u & -a^{-1}bS^{-1} \\ -S^{-1}ca^{-1} & S^{-1} \end{pmatrix}\begin{pmatrix} \mathbb{1}_N \\ 0 \\ 0 \end{pmatrix}
$$

$$
= \frac{1}{h^3}\begin{pmatrix} 0 & 0 & -I_N \\ 0 & 0 & 0 \\ 0 & 0 & I_N \end{pmatrix}\begin{pmatrix} u\begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix} \\ -S^{-1}ca^{-1}\begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix} \end{pmatrix} + o\left(\frac{1}{h^3}\right)
$$

$$
= -\frac{1}{h^3}\begin{pmatrix} -I_N \\ 0 \\ I_N \end{pmatrix}(C^*\tilde{\Pi}C)^{-1}C^*\tilde{\Gamma}^{+,*}\begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix} + o\left(\frac{1}{h^3}\right).
$$

Eventually, one has

$$
\left(A + \tfrac{h}{2}B \quad A - \tfrac{h}{2}B \quad A + \tfrac{h}{2}B + h^3 C_h \right)^+
$$

$$
= \begin{pmatrix} \tfrac{1}{2}I_N & \tfrac{1}{2}I_N & -I_N \\ \tfrac{1}{2}I_N & -\tfrac{1}{2}I_N & 0 \\ 0 & 0 & I_N \end{pmatrix} \operatorname{diag}\left(1, \frac{2}{h}, \frac{1}{h^3}\right) \left(\tilde{\Gamma} \quad C_h \right)^+
$$

$$
= \frac{1}{h^3} \begin{pmatrix} 0 & 0 & -I_N \\ 0 & 0 & 0 \\ 0 & 0 & I_N \end{pmatrix} \begin{pmatrix} u & -a^{-1}bS^{-1} \\ -S^{-1}ca^{-1} & S^{-1} \end{pmatrix} \left(\tilde{\Gamma}^* \quad C_h^* \right) + o\left(\frac{1}{h^3} \right)
$$

$$
= \frac{1}{h^3} \begin{pmatrix} -(C^*\tilde{\Pi}C)^{-1}C^*\tilde{\Pi} \\ 0 \\ (C^*\tilde{\Pi}C)^{-1}C^*\tilde{\Pi} \end{pmatrix} + o\left(\frac{1}{h^3} \right),
$$

and

$$
\left(A + \tfrac{h}{2}B \quad A - \tfrac{h}{2}B \quad A + \tfrac{h}{2}B + h^3 C_h \right)^{+,*} \begin{pmatrix} \mathbb{1}_N \\ \mathbb{1}_N \\ \mathbb{1}_N \end{pmatrix}
$$

$$
= \begin{pmatrix} \tilde{\Gamma}^* \\ C_h^* \end{pmatrix}^+ \operatorname{diag}\left(1, \frac{2}{h}, \frac{1}{h^3}\right) \begin{pmatrix} \tfrac{1}{2}I_N & \tfrac{1}{2}I_N & 0 \\ \tfrac{1}{2}I_N & -\tfrac{1}{2}I_N & -I_N \\ 0 & 0 & I_N \end{pmatrix} \begin{pmatrix} \mathbb{1}_N \\ \mathbb{1}_N \\ \mathbb{1}_N \end{pmatrix}
$$

$$
= \left[\tilde{\Gamma}^{+,*} - \tilde{\Pi}C_h(C_h^*\tilde{\Pi}C_h)^{-1}C_h^*\tilde{\Gamma}^{+,*} \right] \begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix}
$$

$$
= \tilde{\Gamma}^{+,*} \begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix} - \tilde{\Pi}C(C^*\tilde{\Pi}C)^{-1}C^*\tilde{\Gamma}^{+,*} \begin{pmatrix} \mathbb{1}_N \\ 0 \end{pmatrix} + o(1)
$$

$\square$

Appendix B. Proofs for Section 2

The proof of Theorem 1 is given below. It relies on the following fundamental Lemma (which is a simplified version of [17, Lemma 4]), which characterizes μ_0.

Lemma 3. *Let* $J^{(\mathbf{r})}, J^{(\mathbf{l})} \subset [\![0, G-1]\!]$*, and* $(a_0, b_0) \in \mathcal{C}_h^G$*. Assume that* $(I^{(\mathbf{r})}, I^{(\mathbf{l})}) \overset{\text{def.}}{=} (I^{(\mathbf{r})}(a_0, b_0), I^{(\mathbf{l})}(a_0, b_0))$ *is such that*

$$
I^{(\mathbf{r})} \subset J^{(\mathbf{r})}, \quad I^{(\mathbf{l})} \subset J^{(\mathbf{l})} \quad and \quad \hat{\mathcal{A}}_h \overset{\text{def.}}{=} \left((A + \tfrac{h}{2}B)_{J^{(\mathbf{r})}} \quad (A - \tfrac{h}{2}B)_{J^{(\mathbf{l})}} \right)
$$

has full rank. Define $\begin{pmatrix} u_{J^{(\mathbf{r})}} \\ v_{J^{(\mathbf{l})}} \end{pmatrix} \overset{\text{def.}}{=} -(\hat{\mathcal{A}}_h^* \hat{\mathcal{A}}_h)^{-1}s$ *where* $s \overset{\text{def.}}{=} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^{|J^{(\mathbf{r})}|+|J^{(\mathbf{l})}|}$*. Then*

$(J^{(\mathbf{r})}, J^{(\mathbf{l})})$ *is the extended support of* (a_0, b_0) *if and only if the following two conditions hold:*

- *for all $j \in J^{(\mathbf{r})} \setminus I^{(\mathbf{r})}$, $u_j \geqslant 0$, and for all $j \in J^{(\mathbf{l})} \setminus I^{(\mathbf{l})}$, $v_j \geqslant 0$.*

- $\max \left[\begin{pmatrix} (A + \frac{h}{2}B)^*_{J^{(\mathbf{r})^c}} \\ (A - \frac{h}{2}B)^*_{J^{(\mathbf{l})^c}} \end{pmatrix} \hat{\mathcal{A}}_h \begin{pmatrix} u_{J^{(\mathbf{r})}} \\ v_{J^{(\mathbf{l})}} \end{pmatrix} \right] < 1.$

Moreover, in that case, the minimal norm certificate is given by

$$\bar{\eta}_0 = -\mathcal{A}_h{}^* (\hat{\mathcal{A}}_h^* \hat{\mathcal{A}}_h)^{-1} s.$$

[ToDo: Ecrire la preuve du theoreme 1. Ou alors changer l'enonce du theoreme 1 pour faire apparaitre les constantes c_1, ... c_5]

Appendix C. Proofs for Section 3

Appendix C.1. Proof of Proposition 3

We observe that our embedding of the discrete problem into the continuous one, that is $m_n = \sum_{i=0}^{G_n-1} a_i^{(n)} \delta_{ih_n + b_i/a_i}$, yields $\|a^{(n)}\|_1 = |m_n|(\mathbb{T}) = m_n(\mathbb{T})$. For the liminf inequality, let $(m_n)_{n \in \mathbb{N}}$ be of the form (37) which weakly* converges towards m. We notice that $\Phi'_{\mathcal{G}_n} b^{(n)} = \Phi'(\sum_{i=0}^{G_n-1} b_i^{(n)} \delta_{ih_n})$, where $\Phi' : m \mapsto \int_{\mathbb{T}} \varphi'(x) \mathrm{d}m(x)$ is continuous from $\mathcal{M}(\mathbb{T})$ to $\mathcal{H}$ (in the strong topologies). Moreover,

$$\left| \sum_{i=0}^{G_n-1} b_i^{(n)} \delta_{ih_n} \right| (\mathbb{T}) \leqslant \frac{h_n}{2} \left(\sum_{i=0}^{G_n-1} a_i^{(n)} \right) \leqslant \frac{h_n}{2\lambda} \left(\frac{1}{2} \|y\|_{\mathcal{H}}^2 + 1 \right) \to 0,$$

so that $\Phi'_{\mathcal{G}_n} b^{(n)}$ converges strongly towards 0 in $\mathcal{H}$. Additionally, $\Phi_{\mathcal{G}_n} a^{(n)} = \Phi(\sum_{i=0}^{G_n-1} a_i^{(n)} \delta_{ih_n})$ and for all $\psi \in \mathscr{C}(\mathbb{T})$,

$$\left| \left\langle \sum_{i=0}^{G_n-1} a_i^{(n)} \delta_{ih_n + b_i/a_i} - \sum_{i=0}^{G_n-1} a_i^{(n)} \delta_{ih_n}, \psi \right\rangle \right| = \left| \sum_{i=0}^{G_n-1} a_i^{(n)} (\psi(ih_n + b_i/a_i) - \psi(ih_n)) \right|$$

$$\leqslant \sum_{i=0}^{G_n-1} a_i^{(n)} \omega_\psi \left(\frac{h_n}{2} \right) \to 0$$

where $\omega_\psi : t \mapsto \sup_{|x'-x| \leqslant t} |\psi(x) - \psi(x')|$ is the modulus of continuity of ψ. As a result, $\sum_{i=0}^{G_n-1} a_i^{(n)} \delta_{ih_n} - m_n \overset{*}{\rightharpoonup} 0$ so that $\sum_{i=0}^{G_n-1} a_i^{(n)} \delta_{ih_n}$ weakly* converges to m. Hence, $\Phi_{\mathcal{G}_n} a^{(n)}$ weakly converges towards Φm in $\mathcal{H}$. To sum up, $\Phi_{\mathcal{G}_n} a^{(n)} + \Phi'_{\mathcal{G}_n} b^{(n)} - y$ weakly converges towards $\Phi m - y$ and we conclude by invoking the lower semi-continuity of both terms:

$$\liminf_{n \to +\infty} \left(\lambda \|a^{(n)}\|_1 + \frac{1}{2} \left\| \Phi_{\mathcal{G}_n} a^{(n)} + \Phi'_{\mathcal{G}_n} b^{(n)} - y \right\|_{\mathcal{H}}^2 \right) \geqslant \liminf_{n \to +\infty} \left(\lambda |m_n|(\mathbb{T}) + \frac{1}{2} \left\| \Phi_{\mathcal{G}_n} a^{(n)} + \Phi'_{\mathcal{G}_n} b^{(n)} - y \right\|_{\mathcal{H}}^2 \right)$$

$$\geqslant \lambda m(\mathbb{T}) + \frac{1}{2} \left\| \Phi m - y \right\|_{\mathcal{H}}^2.$$

For the limsup inequality, we build a recovery sequence m_n by choosing $a_k^{(n)} = m([kh_n, (k+1)h_n))$ and $b_k^{(n)} = 0$ for all $k \in [\![0, G_n - 1]\!]$. Then, for all n, $\|a^{(n)}\|_1 = m(\mathbb{T})$

and

$$\left\| \Phi \left(\sum_{i=0}^{G_n-1} a_i^{(n)} \delta_{ih_n} \right) - \Phi m \right\|_{\mathcal{H}} \leqslant \sum_{i=0}^{G_n-1} \left\| \int_{[ih_n,(i+1)h_n)} (\varphi(t) - \varphi(ih_n)) \mathrm{d}m(t) \right\|_{\mathcal{H}}$$

$$\leqslant m(\mathbb{T})\omega_\varphi(h_n) \to 0$$

since φ is uniformly continuous on $\mathbb{T}$. As a result, $\Phi m_n - y$ strongly converges towards $\Phi m - y$ and the limsup inequality is proved.

Appendix D. Proofs of Section 4

The following lemma is central in our analysis. It studies sequences of dual certificates $(\eta_n)_{n\in\mathbb{N}}$. For $0 < r < \frac{1}{2}\min_{\nu\neq\nu'}|x_\nu - x_{\nu'}|$, $\nu \in \{1, \ldots, N\}$, it will be useful to consider the sets:

$$S_{n,\nu}^{(\mathbf{r})}(r) \overset{\text{def.}}{=} \left\{ t \in \mathcal{G}_n \cap (x_\nu - r, x_\nu + r) \; ; \; \left(\eta_n + \frac{h_n}{2}\eta_n' \right)(t) = 1 \right\},$$

$$S_{n,\nu}^{(\mathbf{l})}(r) \overset{\text{def.}}{=} \left\{ t \in \mathcal{G}_n \cap (x_\nu - r, x_\nu + r) \; ; \; \left(\eta_n - \frac{h_n}{2}\eta_n' \right)(t) = 1 \right\}.$$

Lemma 4. *Let $(x_1, \ldots, x_N) \in \mathbb{T}^N$ pairwise distinct, and let $\{\eta_n\}_{n\in\mathbb{N}} \in (\mathscr{C}^3(\mathbb{T}))^\mathbb{N}$ be a sequence of functions which converges uniformly towards some η_∞ (and similarly for the derivatives) such that for all $\nu \in \{1, \ldots, N\}$, $\eta_\infty(x_\nu) = 1$ and for all $t \in \mathbb{T}\backslash\{x_1, \ldots, x_N\}$, $\eta_\infty(t) < 1$.*

(i) Then

$$\limsup_{n\to+\infty} \left\{ t \in \mathcal{G}_n \; ; \; \eta_n(t) + \frac{h_n}{2}|\eta_n'(t)| = 1 \right\} \subset \{x_1, \ldots, x_N\}. \tag{D.1}$$

In particular for $r > 0$ small enough, there exists $n_0 \in \mathbb{N}$ such that for $n \geqslant n_0$

$$\left\{ t \in \mathcal{G}_n \; ; \; \eta_n(t) + \frac{h_n}{2}|\eta_n'(t)| = 1 \right\} = \bigcup_{\nu=1}^N \left(S_{n,\nu}^{(\mathbf{r})}(r) \cup S_{n,\nu}^{(\mathbf{l})}(r) \right) \subset \bigcup_{\nu=1}^N (x_\nu - r, x_\nu + r).$$

Assume moreover that for all $n \in \mathbb{N}$ and all $t \in \mathcal{G}_n$, $\eta_n(t) + \frac{h_n}{2}|\eta_n'(t)| \leqslant 1$. For each $\nu \in \{1, \ldots, N\}$:

(ii) If $\eta_\infty''(x_\nu) \neq 0$, then there exists $n_0 \in \mathbb{N}$ such that for $n \geqslant n_0$, each set $S_{n,\nu}^{(\mathbf{r})}(r)$ and $S_{n,\nu}^{(\mathbf{l})}(r)$ is of the form $\emptyset$, $\{ih_n\}$, or $\{ih_n, (i+1)h_n\}$, and if both sets are nonempty:

$$\max S_{n,\nu}^{(\mathbf{r})}(r) \leqslant \min S_{n,\nu}^{(\mathbf{l})}(r).$$

(iii) If $\eta_\infty^{(3)}(x_\nu) \neq 0$, then there exists $n_0 \in \mathbb{N}$ such that for each $n \geqslant n_0$, $S_{n,\nu}^{(\mathbf{r})}(r) = \emptyset$ or $S_{n,\nu}^{(\mathbf{l})}(r) = \emptyset$.

(iv) If $\eta_\infty^{(4)}(x_\nu) \neq 0$, the set of $n \in \mathbb{N}$ such that $S_{n,\nu}^{(\mathbf{r})}(r) = \{(i_n - 1)h_n, i_n h_n\}$ and $S_{n,\nu}^{(\mathbf{l})}(r) = \{i_n h_n, (i_n + 1)h_n\}$ (with the same $i_n \in [\![0, G_n - 1]\!]$) is finite.

Proof.   (i) For all $\tilde{r} \in (0, r)$, by compactness, $\sup \{\eta_\infty(t) \;;\; t \in \mathbb{T} \setminus \bigcup(x_\nu - \tilde{r}, x_\nu + \tilde{r})\} < 1$. Thus by uniform convergence there exists $n_0 \in \mathbb{N}$ such that for all $n \geqslant n_0$, $(\eta_n \pm \frac{h_n}{2}\eta_n') < 1$ on $\mathbb{T} \setminus \bigcup_{\nu=1}^N (x_\nu - \tilde{r}, x_\nu + \tilde{r})$, and the first claim is proved.

(ii) If moreover $\eta_\infty''(x_\nu) \neq 0$, it is in fact negative. Choosing $\tilde{r} \in (0, r)$ small enough and then n large enough, we may assume that $\eta_n'' < -k_0$ in $(x_\nu - \tilde{r}, x_\nu + \tilde{r})$, for some $k_0 > 0$, and by (D.1) that $S_{n,\nu}^{(\mathbf{r})}(r) \cup S_{n,\nu}^{(\mathbf{l})}(r) \subset (x_\nu - \tilde{r}, x_\nu + \tilde{r})$. By uniform convergence, $\eta_n'' + \frac{h_n}{2}|\eta_n^{(3)}| < -\frac{k_0}{2}$ for n large enough, so that both functions $\eta_n + \frac{h_n}{2}\eta_n'$ and $\eta_n - \frac{h_n}{2}\eta_n'$ are strictly concave in $(x_\nu - \tilde{r}, x_\nu + \tilde{r})$. This implies that $S_{n,\nu}^{(\mathbf{r})}(r)$ (resp. $S_{n,\nu}^{(\mathbf{l})}(r)$) is of the form $\emptyset$, $\{ih_n\}$, or $\{ih_n, (i+1)h_n\}$.
Observe also that $\eta_n + \frac{h_n}{2}\eta_n' - (\eta_n - \frac{h_n}{2}\eta_n') = h_n\eta_n'$. Since the function η_n' is strictly decreasing in $(x_\nu - \tilde{r}, x_\nu + \tilde{r})$, it vanishes at most once. If $S_{n,\nu}^{(\mathbf{r})}(r) \neq \emptyset$ and $S_{n,\nu}^{(\mathbf{l})}(r) \neq \emptyset$, it must change sign in $(x_\nu - \tilde{r}, x_\nu + \tilde{r})$ and thus it vanishes exactly once, at some $\xi \in (x_\nu - \tilde{r}, x_\nu + \tilde{r})$. Then for $t \in (x_\nu - \tilde{r}, \xi)$,

$$(\eta_n - \frac{h_n}{2}\eta_n')(t) = (\eta_n + \frac{h_n}{2}\eta_n')(t) - h_n\eta_n'(t) \leqslant 1 - h_n\eta_n'(t) < 1$$

so that $\min S_{n,\nu}^{(\mathbf{l})}(r) \geqslant \xi$. Similarly $\max S_{n,\nu}^{(\mathbf{r})}(r) \leqslant \xi$.

(iii) By contradiction, assume that the set of $n' \in \mathbb{N}$ such that $S_{n',\nu}^{(\mathbf{r})}(r) \neq \emptyset$ and $S_{n',\nu}^{(\mathbf{l})}(r) \neq \emptyset$ is infinite. We may extract a subsequence $n = n'(m)$ such that there exists $i_n, j_n \in [\![0, G_n - 1]\!]$ (denoted hereafter i, j) with $ih_n \in S_{n,\nu}^{(\mathbf{r})}(r_m)$, $jh_n \in S_{n,\nu}^{(\mathbf{l})}$. Combining the Taylor expansions of η_n and $(\eta_n)'$ around ih_n (resp. jh_n), we get

$$1 \geqslant \eta_n((i+1)h_n) - \frac{h_n}{2}\eta_n'((i+1)h_n)$$

$$= \underbrace{\eta_n(ih_n) + h_n\eta_n'(ih_n)(1 - \frac{1}{2})}_{=1} + h_n{}^2\eta_n''(ih_n)\underbrace{\left(\frac{1}{2!} - \frac{1}{2}\right)}_{=0} + h_n{}^3\eta_n^{(3)}(ih_n)\alpha_3$$

$$+ h_n{}^4 \int_0^1 \eta_n^{(4)}(ih_n + th_n)\left(\frac{(1-t)^3}{3!} - \frac{(1-t)^2}{2! \times 2}\right) \mathrm{d}t, \text{ and}$$

$$1 \geqslant \eta_n((j-1)h_n) + \frac{h_n}{2}\eta_n'((j-1)h_n)$$

$$= \underbrace{\eta_n(jh_n) - h_n\eta_n'(jh_n)(1 - \frac{1}{2})}_{=1} + h_n{}^2\eta_n''(jh_n)\underbrace{\left(\frac{1}{2!} - \frac{1}{2}\right)}_{=0} - h_n{}^3\eta_n^{(3)}(jh_n)\alpha_3$$

$$+ h_n{}^4 \int_0^1 \eta_n^{(4)}(jh_n - th_n)\left(\frac{(1-t)^3}{3!} - \frac{(1-t)^2}{2! \times 2}\right) \mathrm{d}t$$

where α_k is defined in (D.11). Now, let $n \to +\infty$. By (D.1), $ih_n \to x_\nu$ and $jh_n \to x_\nu$, and using the uniform convergence of $\eta_n^{(k)}$ towards $\eta_\infty^{(k)}$, dividing by $h_n{}^3$, we obtain respectively $0 \geqslant -\eta_\infty^{(3)}(x_\nu) \times \frac{1}{12}$ and $0 \geqslant \eta_\infty^{(3)}(x_\nu) \times \frac{1}{12}$, thus $\eta_\infty^{(3)}(x_\nu) = 0$.

(iv) Assume, by contradiction, that the mentioned set is infinite. For such n, a Taylor

expansion at ih_n yields (we write i for i_n):

$$1 = \eta_n((i+1)h_n) - \frac{h_n}{2}\eta_n{}'((i+1)h_n)$$

$$= \underbrace{\eta_n(ih_n) + \frac{h_n}{2}\eta_n{}'(ih_n)}_{=1} + \gamma_3 h_n{}^3 \eta_n{}^{(3)}(ih_n) + \gamma_4 h_n{}^4 \eta_n{}^{(4)}(ih_n)$$

$$+ h_n{}^5 \int_0^1 \eta_n{}^{(5)}(ih_n + th_n)\left(\frac{(1-t)^4}{4!} - \frac{(1-t)^3}{3! \times 2}\right) dt, \text{ and}$$

$$1 = \eta_n((i-1)h_n) + \frac{h_n}{2}\eta_n{}'((i-1)h_n)$$

$$= \underbrace{\eta_n(ih_n) - \frac{h_n}{2}\eta_n{}'(ih_n)}_{=1} - \gamma_3 h_n{}^3 \eta_n{}^{(3)}(ih_n) + \gamma_4 h_n{}^4 \eta_n{}^{(4)}(ih_n)$$

$$+ h_n{}^5 \int_0^1 \eta_n{}^{(5)}(ih_n + th_n)\left(\frac{(1-t)^4}{4!} - \frac{(1-t)^3}{3! \times 2}\right) dt,$$

with $\gamma_k = \frac{1}{k!} - \frac{1}{(k-1)! \times 2}$. Summing both equalities, dividing by $h_n{}^4$ and taking the limit $n \to +\infty$ yields $(\eta_\infty)^{(4)}(x_\nu) = 0$, a contradiction.

$\square$

This other lemma focusses on the limit of the sets D^n defined in (33).

Lemma 5. *As $n \to +\infty$, the sets D^n converge towards D^∞ defined in (35) (in the sense of set convergence).*

Proof. We observe that $E^n \subset D^n \subset F^n$, where

$$E^n \overset{\text{def.}}{=} \left\{p \in \mathcal{H} \; ; \; \max_{t \in \mathbb{T}}(\Phi^* p)(t) + \frac{h_n}{2}|(\Phi^* p)'(t)| \leqslant 1\right\},$$

$$F^n \overset{\text{def.}}{=} \left\{p \in \mathcal{H} \; ; \; \max_{k \in [\![0, G_n - 1]\!]} \Phi^* p(kh_n) \leqslant 1\right\}$$

so that it suffices to prove that E^n and F^n converge towards D^∞. On the one hand, it is clear that $D^\infty = \bigcap_{n \in \mathbb{N}} F^n$, and the sequence F^n is non-increasing. On the other hand, it is possible to check that $D^\infty = \overline{\bigcup_{n \in \mathbb{N}} E^n}$, and the sequence E^n is non-decreasing. As a consequence, the claimed set convergences hold (see [26, Ex. 4.3]). $\square$

Let us recall that the dual problem to $(\mathcal{Q}_\lambda^n(y))$ is the projection onto the closed convex set

$$D^n \overset{\text{def.}}{=} \left\{p \in \mathcal{H} \; ; \; (\Phi^* p)(ih_n) + \frac{h_n}{2}|(\Phi^* p)'(ih_n)| \leqslant 1\right\}.$$

Since the set convergence of D^n (see Lemma 5 above) implies the convergence of the projections onto D^n (see [26], or [16] for a direct proof in a similar context), we obtain:

Lemma 6. *Let $p_{\lambda,n}$ (resp. $p_{\lambda,\infty}$) be a solution of $(\mathcal{E}_\lambda^n(y))$ (resp. $(\mathcal{E}_\lambda^\infty(y))$), and $\eta_{\lambda,n} = \Phi^* p_{\lambda,n}$ (resp. $\eta_{\lambda,\infty} = \Phi^* p_{\lambda,\infty}$). Then*

$$\lim_{n\to+\infty} p_{\lambda,n} = p_{\lambda,\infty} \text{ strongly in } \mathcal{H},$$

$$\lim_{n\to+\infty} \eta_{\lambda,n}^{(k)} = \eta_{\lambda,\infty}^{(k)} \text{ in the sense of the uniform convergence, for all } k \in \mathbb{N} \text{ up to the regularity of } \varphi.$$

We are now in position to prove Proposition 4.

Proof of Proposition 4. By Lemma 6, we know that the dual certificates $\eta_{\lambda,n}$ converge towards $\eta_{\lambda,\infty}$. By Lemma 4 (i) and the optimality conditions, we have thus $\limsup_{n\to+\infty}(\mathrm{supp}(m_{\lambda,n})) \subset \{x_1,\ldots,x_N\}$.

If $m_{\lambda,\infty}$ is the unique solution, assume by contradiction that $\liminf(\mathrm{supp}(m_{\lambda,n})) \subsetneq \{x_1,\ldots,x_N\}$. Then there is some ν, some $\varepsilon > 0$ such that (up to a subsequence) $(\mathrm{supp}(m_{\lambda,n})) \cap (x_\nu - \varepsilon, x_\nu + \varepsilon) = \emptyset$. This contradicts the Γ-convergence result (Prop. 3) which ensures that $m_{\lambda,n}$ converges towards $m_{\lambda,\infty}$ in the weak* topology. As a result $\lim_{n\to+\infty}(\mathrm{supp}(m_{\lambda,n})) = \{x_1,\ldots,x_N\}$.

If $\mu_{\lambda,\infty}''(x_\nu) \neq 0$, Lemma 4 ensures that the sets $S_{n,\nu}^{(\mathbf{r})}(r)$ and $S_{n,\nu}^{(\mathbf{l})}(r)$ are of the form $\emptyset$, $\{ih_n\}$, or $\{ih_n, (i+1)h_n\}$. Moreover, since $\lim_{n\to+\infty}(\mathrm{supp}\, m_{\lambda,n}) = \{x_1,\ldots,x_N\}$ we must have $S_{n,\nu}^{(\mathbf{r})}(r) \neq \emptyset$ or $S_{n,\nu}^{(\mathbf{l})}(r) \neq \emptyset$. Using the fact that $\max S_{n,\nu}^{(\mathbf{r})}(r) \leqslant \min S_{n,\nu}^{(\mathbf{l})}(r)$, one may check that the only possible saturation points of $\eta_{\lambda,n} + \frac{h_n}{2}\eta_{\lambda,n}'$ and $\eta_{\lambda,n} - \frac{h_n}{2}\eta_{\lambda,n}'$ are given in Table 1. The optimality conditions of Proposition 1 imply that $m_{\lambda,n}$ is at most a sum of Dirac masses at those locations.

If $\eta_{\lambda,\infty}^{(3)}(x_\nu) \neq 0$, Lemma 4 (iii) implies that for n large enough, $S_{n,\nu}^{(\mathbf{r})}(r) = \emptyset$ or $S_{n,\nu}^{(\mathbf{r})}(r) = \emptyset$ (but not both). Hence there are at most two (successive) saturations, produced either by $\eta_{\lambda,n} + \frac{h_n}{2}\eta_{\lambda,n}'$ or by $\eta_{\lambda,n} - \frac{h_n}{2}\eta_{\lambda,n}'$. $\qquad\square$

Appendix D.1. Sufficient Condition for Asymptotic Solutions

The following property ensures that m_0 is a solution to $(\mathcal{Q}_0^n(y_0))$ for each n large enough.

Lemma 7. *Assume that there exists a function $\eta \in \mathrm{Im}\,\Phi^*$, such that for all $t \in \mathbb{T} \setminus \{x_{0,1},\ldots,x_{0,N}\}$, $\eta(t) < 1$ and*

$$\forall \nu \in \{1,\ldots,N\}, \ \eta(x_{0,\nu}) = 1, \ \eta''(x_{0,\nu}) \neq 0, \ \eta^{(3)}(x_{0,\nu}) = 0, \ \eta^{(4)}(x_{0,\nu}) > 0.$$

Then, for all n large enough, η is a dual certificate for $m_0 = \sum_{\nu=1}^N \alpha_{0,\nu}\delta_{x_{0,\nu}}$ for $(\mathcal{Q}_0^n(y_0))$, and m_0 is a solution to $(\mathcal{Q}_0^n(y_0))$. Moreover, if Γ_{x_0} has full rank, this solution is unique.

Remark 3. The condition $\eta^{(3)}(x_{0,\nu}) = 0$ is natural since our aim is to build a certificate which is valid for all n, hence Lemma 4 applies with $S_{n,\nu}^{(\mathbf{r})}(r) \neq \emptyset$ and $S_{n,\nu}^{(\mathbf{l})}(r) \neq \emptyset$.

Proof. Let $\nu \in \{1,\ldots,N\}$ and $r_\nu \in (0,r)$ such that $\eta''(t) < 0$, and $\eta^{(4)}(t) > 0$ in $(x_{0,\nu} - r_\nu, x_{0,\nu} + r_\nu)$. We shall prove that $\eta(kh_n) + \frac{h_n}{2}|\eta'(kh_n)| < 1$ for all k such that
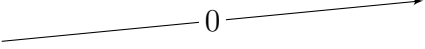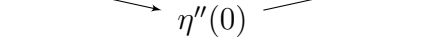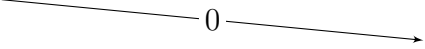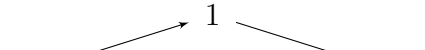
t	$-\tilde{r}$	0	$\tilde{r}$
$\eta^{(4)}$		$+$	
$\eta^{(3)}$		0	
η''	$\eta''(-\tilde{r}) < 0$	$\eta''(0)$	$\eta''(\tilde{r}) < 0$
η'		0	
η		1	

Table D1: Variations of η and its derivatives.

$kh_n \in (x_{0,\nu} - r_\nu, x_{0,\nu} + r_\nu) \setminus \{x_{0,\nu}\}$. To simplify the notation, we assume without loss of generality that $x_{0,\nu} = 0$ and we write $\tilde{r} = r_\nu$. The variations of η and its derivatives are given in Table D1.

Let us observe that the function $\theta : t \mapsto \eta(t) - \frac{t}{2}\eta'(t)$ is (strictly) decreasing in $[0, \tilde{r})$, since

$$\forall t \in (0, \tilde{r}), \ \theta'(t) = \frac{1}{2}\left(\eta'(t) - t\eta''(t)\right) = \frac{1}{2}\int_0^t \underbrace{\left(\eta''(u) - \eta''(t)\right)}_{<0} du < 0. \tag{D.2}$$

Hence, for all k such that $kh_n \in (0, \tilde{r})$,

$$\eta(kh_n) - \frac{h_n}{2}\eta'(kh_n) = \underbrace{\eta(kh_n) - \frac{kh_n}{2}\eta'(kh_n)}_{=\theta(kh_n)<\theta(0)=1} + \underbrace{\frac{(k-1)h_n}{2}\eta'(kh_n)}_{<0} < 1. \tag{D.3}$$

On the other hand, θ is (strictly) increasing on $(-\tilde{r}, 0]$ since

$$\forall t \in (-\tilde{r}, 0), \ \theta'(t) = \frac{1}{2}\left(\eta'(t) - t\eta''(t)\right) = \frac{1}{2}\int_0^t \underbrace{\left(\eta''(u) - \eta''(t)\right)}_{<0} du > 0. \tag{D.4}$$

As a consequence, for all k such that $kh_n \in (-\tilde{r}, 0)$,

$$\eta(kh_n) + \frac{h_n}{2}\eta'(kh_n) = \underbrace{\eta(kh_n) - \frac{kh_n}{2}\eta'(kh_n)}_{=\theta(kh_n)<\theta(0)=1} + \underbrace{\frac{(k+1)h_n}{2}\eta'(kh_n)}_{\leqslant 0} < 1. \tag{D.5}$$

Thus we see that $\eta(kh_n) + \frac{h_n}{2}|\eta'(kh_n)| < 1$ for all $kh_n \in (-\tilde{r}, \tilde{r}) \setminus \{0\}$, and we proceed similarly on all the intervals of the form $(x_{0,\nu} - r_\nu, x_{0,\nu} + r_\nu)$. By a compactness argument, there exists a constant $\beta < 1$ such that $\eta(t) \leqslant \beta$ for all $t \in \mathbb{T} \setminus \bigcup_{\nu=1}^N (x_{0,\nu} - r_\nu, x_{0,\nu} + r_\nu)$.

For n large enough, the inequality $\frac{h_n}{2}\left(\sup_{t\in\mathbb{T}}|\eta'(t)|\right) < 1 - \beta$ holds, and we see that $\eta(kh_n) + \frac{h_n}{2}|\eta'(kh_n)| < 1$ for all $t \in \mathbb{T} \setminus \bigcup_{\nu=1}^{N}(x_{0,\nu} - r_\nu, x_{0,\nu} + r_\nu)$.

As a conclusion, we see that η is a valid certificate for $(a_0, 0)$ (see the optimality conditions of Proposition 1), thus $(a_0, 0)$ is a solution of $(\mathcal{Q}_0^n(y_0))$. $\qquad\square$

Appendix D.2. Asymptotics of the Minimal Norm Certificates

Observe that if the Twice Non-Degenerate Source Condition (Definition 5) holds, the hypotheses of Lemma 7 are satisfied and m_0 is a solution to $(\mathcal{Q}_0^n(y_0))$ for n large enough. In fact the associated minimal norm certificates (which thus exist) converge towards η_T.

Proposition 5. *Let $m_0 \in \mathcal{M}(\mathbb{T})$ satisfy the Twice Non-Degenerate Source Condition (and η_T the corresponding Third derivative (pre)certificate). Let $p_{0,n}$ be the minimal norm solution of $(\mathcal{E}_0^n(y_0))$, and $\eta_{0,n} = \Phi^* p_{0,n}$. Then,*

$$\lim_{n\to+\infty} p_{0,n} = p_T \text{ for the } \mathcal{H} \text{ strong topology,} \tag{D.6}$$

$$\lim_{n\to+\infty} \eta_{0,n}^{(k)} = \eta_T^{(k)} \text{ in the sense of the uniform convergence, for all } k \in \mathbb{N} \text{ up to the regularity of } \varphi. \tag{D.7}$$

Proof. As mentioned above, the Twice Non-Degenerate Source Condition implies that Lemma 7 applies to the function η_T, hence η_T is a certificate for $(\mathcal{Q}_0^n(y_0))$. As a result, $\|p_{0,n}\| \leqslant \|p_T\|$ and the sequence $(p_{0,n})_{n\in\mathbb{N}}$ is bounded in $\mathcal{H}$. We may extract a subsequence $p_{0,n'}$ which weakly converges towards some $\tilde{p} \in \mathcal{H}$, and then $\|\tilde{p}\| \leqslant \liminf_{n'\to+\infty} \|p_{0,n}\| \leqslant \|p_T\|$. Since Φ^* and $\Phi^{(k),*}$ are compact (see [17, Lemma 1]), we obtain that $\eta_{0,n'}^{(k)} = (\Phi^* p_{0,n'})^{(k)}$ converges toward $\tilde{\eta}^{(k)} \overset{\text{def.}}{=} (\Phi^* \tilde{p})^{(k)}$ in the (strong) topology of uniform convergence. We immediately obtain that $\tilde{\eta}(t) \leqslant 1$ for all $t \in \mathbb{T}$, and $\tilde{\eta}(x_{0,\nu}) = 1$, $\tilde{\eta}(x_{0,\nu}) = 0$ for all $\nu \in \{1, \ldots, N\}$.

Moreover, applying Lemma 4 to $\Phi^* p_{0,n}$ (observing that $x_\nu \in S_{n,\nu}^{(\mathbf{r})}(r) \cap S_{n,\nu}^{(\mathbf{l})}(r)$), we get $\tilde{\eta}^{(3)}(x_{0,\nu}) = 0$. As a result, $\tilde{p}$ is admissible for (40), hence $\|p_T\| \leqslant \|\tilde{p}\|$. Thus in fact $\|p_T\| = \|\tilde{p}\|$ and $p_T = \tilde{p}$. Since the limit of the extracted subsequence does not depend on the choice of the subsequence, in fact the whole sequence converges. Moreover, the convergence is strong in $\mathcal{H}$ since $\lim_{n\to+\infty} \|p_{0,n}\| = \|p_T\|$. $\qquad\square$

As a consequence of the above convergence result, the third derivative precertificate controls the extended support on thin grids.

Proposition 6. *Let $m_0 \in \mathcal{M}(\mathbb{T})$ (with $\{x_{0,1}, \ldots x_{0,N}\} \subset \mathcal{G}_n$) such that the Twice Non Degenerate Source Condition holds. Then, for n large enough, m_0 is a solution to $(\mathcal{Q}_0^n(y_0))$ and its extended support is given by:*

$$\text{ext}_n^{(\mathbf{r})}(m_0) = \bigcup_{\nu=1}^{N} S_{n,\nu}^{(\mathbf{r})}(r), \quad \text{and} \quad \text{ext}_n^{(\mathbf{l})}(m_0) = \bigcup_{\nu=1}^{N} S_{n,\nu}^{(\mathbf{l})}(r), \tag{D.8}$$

where

- $S_{n,\nu}^{(\mathbf{r})}(r)$ *is equal to* $\{x_{0,\nu}\}$ *or* $\{x_{0,\nu} - h_n, x_{0,\nu}\}$,
- $S_{n,\nu}^{(\mathbf{l})}(r)$ *is equal to* $\{x_{0,\nu}\}$ *or* $\{x_{0,\nu}, x_{0,\nu} + h_n\}$.

Moreover, one cannot have simultaneously $S_{n,\nu}^{(\mathbf{r})}(r) = \{x_{0,\nu} - h_n, x_{0,\nu}\}$ *and* $S_{n,\nu}^{(\mathbf{l})}(r) = \{x_{0,\nu}, x_{0,\nu} + h_n\}$.

Proof. By Lemma 7, m_0 is a solution to $(\mathcal{Q}_0^n(y_0))$ and η_T is a solution to $(\mathcal{E}_0^n(y_0))$. Applying Lemma 4 to $\eta_{0,n}$, η_T, we see that $S_{n,\nu}^{(\mathbf{r})}(r)$ is of the form $\emptyset$, $\{ih_n\}$ or $\{(i-1)h_n, ih_n\}$, and that $S_{n,\nu}^{(\mathbf{l})}(r)$ is of the form $\emptyset$, $\{jh_n\}$ or $\{jh_n, (j+1)h_n\}$, with $i \leqslant j$. On the other hand, by the extremality relations between $\eta_{0,n}$ (solution of $(\mathcal{E}_0^n(y_0))$) and $m_{0,n}$ (solution of $(\mathcal{Q}_0^n(y_0))$), $x_{0,\nu} \in S_{n,\nu}^{(\mathbf{r})}(r)$ and $x_{0,\nu} \in S_{n,\nu}^{(\mathbf{l})}(r)$. As a consequence $S_{n,\nu}^{(\mathbf{r})}(r)$ is equal to $\{x_{0,\nu}\}$ or $\{x_{0,\nu} - h_n, x_{0,\nu}\}$, and $S_{n,\nu}^{(\mathbf{l})}(r)$ is equal to $\{x_{0,\nu}\}$ or $\{x_{0,\nu}, x_{0,\nu} + h_n\}$.

Now, since $\eta_T^4(0) \neq 0$, the fourth point of Lemma 4 ensures that for n large enough, one cannot have simultaneously $S_{n,\nu}^{(\mathbf{r})}(r) = \{x_{0,\nu} - h_n, x_{0,\nu}\}$ and $S_{n,\nu}^{(\mathbf{l})}(r) = \{x_{0,\nu}, x_{0,\nu} + h_n\}$. $\qquad\square$

Remark 4. As Proposition 6 shows, for each original spike, at most one pair of spikes appears at low noise: the original spike slightly shifted and either the immediate left neighbor shifted by $+h_n/2$ or the immediate right neighbor shifted by $-h_n/2$.

Appendix D.3. Proof of Theorem 2

We proceed by building a good candidate for $\eta_{0,n}$, making the ansatz that for all $\nu \in \{1, \dots, N\}$, its saturation points satisfy

$$\text{if } \rho_\nu > 0, \text{ then } S_{n,\nu}^{(\mathbf{r})}(r) = \{x_{0,\nu} - h_n, x_{0,\nu}\}, \quad \text{and} \quad S_{n,\nu}^{(\mathbf{l})}(r) = \{x_{0,\nu}\}, \tag{D.9}$$

$$\text{if } \rho_\nu < 0, \text{ then } S_{n,\nu}^{(\mathbf{l})}(r) = \{x_{0,\nu}\}, \quad \text{and} \quad S_{n,\nu}^{(\mathbf{l})}(r) = \{x_{0,\nu}, x_{0,\nu} + h_n\}, \tag{D.10}$$

and then, using Lemma 3, we prove that this candidate is indeed the minimal norm certificate.

To comply with the notations of Lemma 3, let us write $\sum_{\nu=1}^N \alpha_{0,i}\delta_{x_{0,\nu}} = \sum_{k=0}^{G_n-1} a_{0,k}\delta_{kh_n}$, and $I^{(\mathbf{r})} \overset{\text{def.}}{=} I^{(\mathbf{l})} \overset{\text{def.}}{=} I \overset{\text{def.}}{=} \{i \in [\![0, G_n - 1]\!] \; ; \; a_{0,i} \neq 0\}$.

For any choice of shift $(\varepsilon_i)_{i\in I} \in \{-1, +1\}^N$, we set $J^{(\mathbf{r})} \overset{\text{def.}}{=} I^{(\mathbf{r})} \cup \{i + \varepsilon_i \; ; \; i \in I \quad \text{and} \quad \varepsilon_i = -1\}$ and $J^{(\mathbf{l})} \overset{\text{def.}}{=} I^{(\mathbf{l})} \cup \{i + \varepsilon_i \; ; \; i \in I \quad \text{and} \quad \varepsilon_i = +1\}$. Since $|x_{0,\nu} - x_{0,\nu'}| > 2h_n$ for $\nu' \neq \nu$ and n large enough, we have $\operatorname{Card} J^{(\mathbf{r})} + \operatorname{Card} J^{(\mathbf{l})} = 3 \times \operatorname{Card} I = 3N$. The idea is to find a choice of ε such that $u_j > 0$ for all $j \in J^{(\mathbf{r})} \setminus I$, and $v_j > 0$ for all $j \in J^{(\mathbf{l})} \setminus J$, where

$$\begin{pmatrix} u_{J^{(\mathbf{r})}} \\ v_{J^{(\mathbf{l})}} \end{pmatrix} \overset{\text{def.}}{=} -(\hat{\mathcal{A}}_h^* \hat{\mathcal{A}}_h)^{-1} \begin{pmatrix} \mathbb{1}_{J^{(\mathbf{r})}} \\ \mathbb{1}_{J^{(\mathbf{l})}} \end{pmatrix}, \quad \hat{\mathcal{A}}_h \overset{\text{def.}}{=} \left((A + \tfrac{h}{2}B)_{J^{(\mathbf{r})}} \quad (A - \tfrac{h}{2}B)_{J^{(\mathbf{l})}} \right)$$

$$A \overset{\text{def.}}{=} \Phi_{\mathcal{G}_n} \quad \text{and} \quad B \overset{\text{def.}}{=} \Phi'_{\mathcal{G}_n}.$$

In this particular case where $I^{(\mathbf{r})} = I^{(\mathbf{l})} = I$, all j in $(J^{(\mathbf{r})} \setminus I) \cup (J^{(\mathbf{l})} \setminus I)$ may be uniquely written as $j = i + \varepsilon_i$ for some $i \in I$, where $\varepsilon_i \in \{-1, +1\}$. We may swap the

columns of $\hat{\mathcal{A}}_h$ so as to reformulate the condition $\begin{pmatrix} u_{J^{(\mathbf{r})}} \\ v_{J^{(\mathbf{l})}} \end{pmatrix} = -(\hat{\mathcal{A}}_h^* \hat{\mathcal{A}}_h)^{-1} \mathbb{1}_{3N}$ into

$$
\begin{pmatrix} \tilde{u}_I \\ \tilde{v}_I \\ \tilde{t}_I \end{pmatrix} = -(\bar{\mathcal{A}}_h^* \bar{\mathcal{A}}_h)^{-1} \begin{pmatrix} \mathbb{1}_N \\ \mathbb{1}_N \\ \mathbb{1}_N \end{pmatrix},
$$

where $\bar{\mathcal{A}}_h \overset{\text{def.}}{=} \left(A_I + \frac{h_n}{2} B_I \operatorname{diag}(\varepsilon) \quad A_I - \frac{h_n}{2} B_I \operatorname{diag}(\varepsilon) \quad A_{I+\varepsilon} - \frac{h_n}{2} B_{I+\varepsilon} \operatorname{diag}(\varepsilon) \right)$ and $\tilde{t}_i > 0$ for all $i \in I$. But a Taylor expansion yields

$$
A_{I+\varepsilon} - \frac{h_n}{2} B_{I+\varepsilon} \operatorname{diag}(\varepsilon) = \underbrace{\Phi_{x_0}}_{=A_I} + \frac{h_n}{2} \underbrace{\Phi'_{x_0} \operatorname{diag}(\varepsilon)}_{=B_I \operatorname{diag}(\varepsilon)} + (h_n)^3 \gamma_3 \Phi_{x_0}^{(3)} \operatorname{diag}(\varepsilon) + o(h_n{}^3),
$$

where we defined

$$
\gamma_k \overset{\text{def.}}{=} \frac{1}{k!} - \frac{1}{(k-1)! \times 2}. \tag{D.11}
$$

Hence, we may apply Lemma 2 to Φ_{x_0}, $\Phi'_{x_0} \operatorname{diag}(\varepsilon)$ and $\gamma_3 \Phi_{x_0}^{(3)} \operatorname{diag}(\varepsilon)$ so as to obtain

$$
\tilde{t}_I = -\frac{1}{\gamma_3 h_n{}^3} \operatorname{diag}(\varepsilon) \rho + o\left(\frac{1}{h_n{}^3} \right).
$$

Therefore it is sufficient to choose $\varepsilon = -\operatorname{sign}(\rho)$ to make all the components of $\tilde{t}_I$ nonnegative.

With that choice of ε, it remains to prove that

$$
\max \left[\begin{pmatrix} (A + \frac{h}{2} B)_{J^{(\mathbf{r})c}}^* \\ (A - \frac{h}{2} B)_{J^{(\mathbf{l})c}}^* \end{pmatrix} \hat{\mathcal{A}}_h \begin{pmatrix} u_{J^{(\mathbf{r})}} \\ v_{J^{(\mathbf{l})}} \end{pmatrix} \right] < 1.
$$

Let us write $\tilde{p}_n \overset{\text{def.}}{=} \hat{\mathcal{A}}_h \begin{pmatrix} u_{J^{(\mathbf{r})}} \\ v_{J^{(\mathbf{l})}} \end{pmatrix}$. Since $\begin{pmatrix} u_{J^{(\mathbf{r})}} \\ v_{J^{(\mathbf{l})}} \end{pmatrix} = -(\hat{\mathcal{A}}_h^* \hat{\mathcal{A}}_h)^{-1} \mathbb{1}_{3N}$, we get $\tilde{p}_n = \hat{\mathcal{A}}_h^{+,*} \mathbb{1}_{3N}$, and applying Lemma 2 to Φ_{x_0}, $\Phi'_{x_0} \operatorname{diag}(\varepsilon)$ and $\gamma_3 \Phi_{x_0}^{(3)} \operatorname{diag}(\varepsilon)$, we see that $\tilde{p}_n$ converges towards p_T (using (41)).

By construction of $\tilde{p}_n$,

$$
\forall j \in J^{(\mathbf{r})} \setminus I, \ \left(\Phi^* \tilde{p}_n + \frac{h_n}{2} (\Phi^* \tilde{p}_n)' \right)(j h_n) = 1,
$$

$$
\text{and} \quad \forall j \in J^{(\mathbf{l})} \setminus I, \ \left(\Phi^* \tilde{p}_n - \frac{h_n}{2} (\Phi^* \tilde{p}_n)' \right)(j h_n) = 1, \tag{D.12}
$$

which may be summarized as

$$
\forall i \in I, \left(\Phi^* \tilde{p}_n - \varepsilon_i \frac{h_n}{2} (\Phi^* \tilde{p}_n)' \right)((i + \varepsilon_i) h_n) = 1.
$$

Arguing as in the proof of point (iv) in Lemma 4 (replacing "$1 = \ldots$" with "$1 \geqslant \ldots$" and using that $\eta_T^{(4)}(x_{0,\nu}) > 0$), we may prove that for n large enough, $\left(\Phi^* \tilde{p}_n + \varepsilon_i \frac{h_n}{2} (\Phi^* \tilde{p}_n)' \right)((i - \varepsilon_i) h_n) < 1$.

Then, by the same argument of compactness and local concavity as in point (ii) of Lemma 4, we observe that

$$\left\{ k \in [\![0,\, G_n - 1]\!] \; ; \; (\Phi^*\tilde{p}_n + \frac{h_n}{2}(\Phi^*\tilde{p}_n)')(kh_n) \geqslant 1 \right\} \subset J^{(\mathbf{r})},$$

$$\left\{ k \in [\![0,\, G_n - 1]\!] \; ; \; (\Phi^*\tilde{p}_n - \frac{h_n}{2}(\Phi^*\tilde{p}_n)')(kh_n) \geqslant 1 \right\} \subset J^{(\mathbf{l})},$$

and those inclusions are in fact equalities. That precisely means that

$$\max \left[\begin{pmatrix} (A + \frac{h}{2}B)^*_{J^{(\mathbf{r})c}} \\ (A - \frac{h}{2}B)^*_{J^{(\mathbf{l})c}} \end{pmatrix} \tilde{p}_n \right] < 1.$$

Hence, by Lemma 3, $\Phi^*\tilde{p}_n$ is the minimal norm certificate $\eta_{0,n}$ and $(J^{(\mathbf{r})}h_n, J^{(\mathbf{l})}h_n)$ is the extended support. This concludes the proof.

Appendix D.4. Proof of Corollary 1

The result is a consequence of Theorem 1 with the extended support provided by Theorem 2. With the notations of [17, Appendix C.2], the constants are given by

$$C_n^{(1)} \stackrel{\text{def.}}{=} \left(\frac{c_{1,n}c_{5,n}}{c_{4,n}} + c_{2,n} \right)^{-1} \quad \text{and} \quad C_n^{(2)} \stackrel{\text{def.}}{=} \min\left(c_{3,n}, \frac{c_{5,n}}{c_{4,n}} \right).$$

Replacing the constants $c_1, \ldots, c_3$ with the expressions corresponding to the continuous framework, and using Lemma 2 we get

$$c_{1,n} = \left\| R_{I^{(\mathbf{r})} \cup I^{(\mathbf{l})}} \hat{\mathcal{A}}_h^+ \right\| \sim \frac{1}{(h_n)^3} \left\| \begin{pmatrix} (\Phi_{x_0}^{(3),*}\tilde{\Pi}\Phi_{x_0}^{(3)})^{-1}\Phi_{x_0}^{(3),*}\tilde{\Pi} \\ 0 \end{pmatrix} \right\|_{\infty,\mathcal{H}} \tag{D.13}$$

$$c_{2,n} = \left\| \begin{pmatrix} \tilde{u}_I \\ \tilde{v}_I \end{pmatrix} \right\| \sim \frac{1}{(h_n)^3} \left\| \begin{pmatrix} \rho \\ 0 \end{pmatrix} \right\|_\infty \tag{D.14}$$

$$c_{3,n} = \left(\left\| R_{(J^{(\mathbf{r})} \setminus I^{(\mathbf{r})}) \cup (J^{(\mathbf{l})} \setminus I^{(\mathbf{l})})} \hat{\mathcal{A}}_h^+ \right\|_{\infty,\mathcal{H}} \right)^{-1} \left(\min_{i \in I} \tilde{t}_i \right) \sim \frac{\min_i \left| \frac{1}{\gamma_3} \rho_i \right|}{\left\| (\Phi_{x_0}^{(3),*}\tilde{\Pi}\Phi_{x_0}^{(3)})^{-1}\Phi_{x_0}^{(3),*}\tilde{\Pi} \right\|_{\infty,\mathcal{H}}} \tag{D.15}$$

where γ_k is defined in (D.11).

As for $c_{4,n}$ and c_{5_n}, the expressions given in [17, Appendix C.2] yield a pessimistic bound for the low noise regime, and we are led to make finer majorizations. Using the reformulation $(\tilde{\mathcal{Q}}_\lambda(y))$ of the C-BP as a (positive) Lasso, we have to ensure that $p_\lambda \stackrel{\text{def.}}{=} \frac{1}{\lambda}(y - Aa - Bb)$ satisfies

$$\max \left[\begin{pmatrix} (A^* + \frac{h_n}{2}B^*)_{(J^{(\mathbf{r})})c} \\ (A^* - \frac{h_n}{2}B^*)_{(J^{(\mathbf{l})})c} \end{pmatrix} p_\lambda \right] < 1$$

where $A \overset{\text{def.}}{=} \Phi_{\mathcal{G}_n}$, $B \overset{\text{def.}}{=} \Phi'_{\mathcal{G}_n}$. Let $\hat{\mathcal{A}}_h \overset{\text{def.}}{=} \left((A + \frac{h}{2}B)_{J^{(\mathbf{r})}} \quad (A - \frac{h}{2}B)_{J^{(\mathbf{l})}} \right)$, $\tilde{\Pi}$ be the orthogonal projector onto $\ker \hat{\mathcal{A}}_h^* = (\operatorname{Im} \hat{\mathcal{A}}_h)^\perp$, and $\omega \overset{\text{def.}}{=} \Phi^* \tilde{\Pi} w$. Since

$$
y - Aa - Bb = w - \hat{\mathcal{A}}_h (\hat{\mathcal{A}}_h^* \hat{\mathcal{A}}_h)^{-1} \hat{\mathcal{A}}_h^* w + \lambda \hat{\mathcal{A}}_h (\hat{\mathcal{A}}_h^* \hat{\mathcal{A}}_h)^{-1} \mathbb{1}_{3N} = \tilde{\Pi} w + \lambda \hat{\mathcal{A}}_h^{+,*} \mathbb{1}_{3N},
$$

we are led to check that

$$
(\omega + \lambda \eta_{0,n})(jh_n) + \frac{h_n}{2}(\omega + \lambda \eta_{0,n})'(jh_n) < \lambda \quad \text{for all } j \in (J^{(\mathbf{r})})^C, \tag{D.16}
$$

$$
(\omega + \lambda \eta_{0,n})(jh_n) - \frac{h_n}{2}(\omega - \lambda \eta_{0,n})'(jh_n) < \lambda \quad \text{for all } j \in (J^{(\mathbf{l})})^C, \tag{D.17}
$$

where $\eta_{0,n} \overset{\text{def.}}{=} \Phi^* (\hat{\mathcal{A}}_h^* \hat{\mathcal{A}}_h)^{-1}) \mathbb{1}_{3N}$ yields the minimal norm certificate

$$
\begin{pmatrix} (\eta_{0,n} + \frac{h_n}{2} \eta_{0,n})(\mathcal{G}_n) \\ (\eta_{0,n} - \frac{h_n}{2} \eta_{0,n})(\mathcal{G}_n) \end{pmatrix} = \begin{pmatrix} (A + \frac{h}{2}B)^* \\ (A - \frac{h}{2}B)^* \end{pmatrix} (\hat{\mathcal{A}}_h^* \hat{\mathcal{A}}_h)^{-1} \mathbb{1}_{3N}.
$$

Given $0 < r < \frac{1}{2} \min_{\nu \neq \nu'} |x_{0,\nu} - x_{0,\nu'}|$, let $N(r) \overset{\text{def.}}{=} \bigcup_\nu (x_{0,\nu} - r, x_{0,\nu} + r)$ be a neighborhood of the $x_{0,\nu}$'s. By the Twice Non-Degenerate Source condition, we may choose $r > 0$, such that

$$
-\tilde{k}_1 \overset{\text{def.}}{=} \sup_{t \in N(r)} \eta_T''(t) < 0, \quad \text{and} \quad \tilde{k}_2 \overset{\text{def.}}{=} \inf_{t \in N(r)} \eta_T^{(4)}(t) > 0.
$$

By compactness, $\tilde{k}_3 \overset{\text{def.}}{=} \sup_{t \in \mathbb{T} \setminus N(r)} \eta_T(t) < 1$.

Let us recall that $\eta_{0,n} \to \eta_T$ in the sense of the uniform convergence (and similarly for the derivatives). As a result, for $n \in \mathbb{N}$ large enough,

$$
\sup_{t \in N(r)} (\eta_{0,n})''(t) < -\frac{\tilde{k}_1}{2} < 0, \quad \inf_{t \in N(r)} (\eta_{0,n})^{(4)}(t) > \frac{\tilde{k}_2}{2} > 0, \quad \sup_{t \in \mathbb{T} \setminus N(r)} \eta_{0,n}(t) < \frac{1 + \tilde{k}_3}{2} < 1,
$$

$$
\tag{D.18}
$$

$$
\frac{h_n}{2} \left\| (\eta_{0,n})^{(3)} \right\|_\infty \leqslant \frac{\tilde{k}_1}{8}, \quad \text{and} \quad \frac{h_n}{2} \left\| (\eta_{0,n})' \right\|_\infty \leqslant \frac{1 - \tilde{k}_3}{6}.
$$

Now, we assume that $\frac{\|w\|_{\mathcal{H}}}{\lambda}$ is small enough, so that

$$
\left\| (\Phi^{(k)})^* \right\|_{\infty, \mathcal{H}} \frac{\|w\|_{\mathcal{H}}}{\lambda} < \frac{\tilde{k}_1}{8}, \text{ for } k \in \{2, 3\}, \quad \left\| (\Phi^{(4)})^* \right\|_{\infty, \mathcal{H}} \frac{\|w\|_{\mathcal{H}}}{\lambda} < \frac{\tilde{k}_2}{4},
$$

$$
\text{and} \quad \left\| (\Phi^{(k)})^* \right\|_{\infty, \mathcal{H}} \frac{\|w\|_{\mathcal{H}}}{\lambda} < \frac{1 - \tilde{k}_3}{6}, \text{ for } k \in \{0, 1\}, \tag{D.19}
$$

Then, using the fact that and $|\omega^{(k)}|(t) \leqslant \left\| (\Phi^{(k)})^* \right\|_{\infty, \mathcal{H}} \|w\|_{\mathcal{H}}$ and $h_n \leqslant 1$, we obtain

$$
\sup_{t \in \mathbb{T} \setminus N(r)} \left(\frac{\omega}{\lambda} + \eta_{0,n} + \frac{h_n}{2} \left| \left(\frac{\omega}{\lambda} + \eta_{0,n} \right)' \right| \right)(t) < 1.
$$

Thus it remains to prove that for each $\nu \in \{1, \ldots, N\}$,

$$\left(\frac{\omega}{\lambda} + \eta_{0,n} + \frac{h_n}{2}(\frac{\omega}{\lambda} + \eta_{0,n})'\right)(t) < 1 \text{ for } t \in (x_{0,\nu} - r, x_{0,\nu} + r) \setminus S_{n,\nu}^{(\mathbf{r})}(r),$$

(D.20)

$$\text{and} \quad \left(\frac{\omega}{\lambda} + \eta_{0,n} - \frac{h_n}{2}(\frac{\omega}{\lambda} + \eta_{0,n})'\right)(t) < 1 \text{ for } t \in (x_{0,\nu} - r, x_{0,\nu} + r) \setminus S_{n,\nu}^{(\mathbf{l})}(r).$$

(D.21)

We only deal with the case $S_{n,\nu}^{(\mathbf{r})}(r) = \{x_{0,\nu}\}$, $S_{n,\nu}^{(\mathbf{l})}(r) = \{x_{0,\nu}, x_{0,\nu} + h_n\}$, the symmetric case being similar. Let $f \stackrel{\text{def.}}{=} \frac{1}{\lambda}\omega(\cdot - x_{0,\nu}) + \eta_{0,n}(\cdot - x_{0,\nu})$. By definition of $\tilde{\Pi}$, $\omega(x_{0,\nu}) = \omega'(x_{0,\nu}) = \omega(x_{0,\nu} + h_n) - \frac{h_n}{2}\omega(x_{0,\nu} + h_n) = 0$, so that

$$f(0) = 1, \quad f'(0) = 1, \quad \text{and} \quad f(h_n) - \frac{h_n}{2}f'(h_n) = 1.$$

(D.22)

Moreover, from Eq. (D.18) to (D.19), and letting $k_1 = \frac{\tilde{k}_1}{8}$, $k_2 = \frac{\tilde{k}_2}{4}$, we deduce that

$$\forall t \in (-r, r), \quad f''(t) + \frac{h}{2}|f^{(3)}(t)| < -k_1 < 0, \quad \text{and} \quad f^{(4)}(t) > k_2 > 0,$$

(D.23)

so that the strict concavity of $f - \frac{h_n}{2}f'$ implies that $(f - \frac{h_n}{2}f')(t) < 1$ for $t \in (-r, -h_n) \cup (0, r)$.

It remains to prove that $(f + \frac{h_n}{2}f')(t) < 1$ for $t \in (-r, r) \setminus (-h_n, 0]$. A Taylor expansion of f and f' yields (writing as usual $\gamma_k = \frac{1}{k!} - \frac{1}{(k-1)!\times 2}$)

$$\underbrace{1 - (f(h) - \frac{h_n}{2}f(h_n))}_{=0} = \underbrace{1 - f(0) - \frac{h_n}{2}f'(0)}_{=0} - h_n{}^3\gamma_3 f^{(3)}(0) - h_n{}^4\gamma_4 f^{(4)}(0) + R_1(h_n),$$

$$1 - (f(-h_n) + \frac{h_n}{2}f'(-h_n)) = \underbrace{1 - f(0) + \frac{h_n}{2}f'(0)}_{=0} + h_n{}^3\gamma_3 f^{(3)}(0) - h_n{}^4\gamma_4 f^{(4)}(0) + R_2(h_n).$$

Adding both equations we get

$$1 - (f(-h_n) + \frac{h_n}{2}f'(-h_n)) = -2h_n{}^4\gamma_4 f^{(4)}(0) + (R_1 + R_2)(h_n)$$

(D.24)

$$\geqslant 2(-\gamma_4)k_2 h_n{}^4 + (R_1 + R_2)(h_n),$$

(D.25)

where

$$R_1 + R_2(h) = h_n{}^5 \int_0^1 (-f^{(5)}(sh_n) + f^{(5)}(-sh_n))\left(\frac{(1-s)^4}{4!} - \frac{(1-s)^3}{2 \times 3!}\right) ds,$$

(D.26)

$$\text{with} \quad \|f^{(5)}\|_\infty \leqslant \left\|\frac{1}{\lambda}\omega^{(5)}\right\|_\infty + \|(\eta_{0,n})^{(5)}\|_\infty = O(1).$$

(D.27)

Hence,

$$1 - (f(-h_n) + \frac{h_n}{2} f'(-h_n)) \geqslant \underbrace{2(-\gamma_4) k_2}_{>0} {h_n}^4 + O({h_n}^5) > 0. \qquad \text{(D.28)}$$

Moreover, by the strict concavity of $f + \frac{h_n}{2} f'$, we also deduce that $(f + \frac{h_n}{2} f')(t) < 1$ for $t \in (-r, -h_n] \cup (0, r)$, thus we get the local inequalities (D.20) and (D.21), hence the global inequalities (D.16) and (D.17).

To conclude, the constants in the condition on $\frac{\|w\|_{\mathcal{H}}}{\lambda}$ are $O(1)$, and gathering the asympotics for $c_{1,n}, c_{2,n}, c_{3,n}$ we obtain $C_n^{(1)} = O({h_n}^3)$, $C_n^{(2)} = O(1)$.

Eventually, the proof of (44) follows from (D.13) and (D.14).

References

[1] J-M. Azais, Y. De Castro, and F. Gamboa. *Spike detection from inaccurate samplings*. Tech. rep. Preprint hal-00780808, Jan. 2013.

[2] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. New York: Springer-Verlag, 2011.

[3] B.N. Bhaskar and B. Recht. "Atomic norm denoising with applications to line spectral estimation". In: *2011 49th Annual Allerton Conference on Communication, Control, and Computing*. 2011, pp. 261–268.

[4] T. Blu et al. "Sparse Sampling of Signal Innovations". In: *IEEE Signal Processing Magazine* 25.2 (2008), pp. 31–40.

[5] K. Bredies and H.K. Pikkarainen. "Inverse problems in spaces of measures". In: *ESAIM: Control, Optimisation and Calculus of Variations* 19.1 (2013), pp. 190–218.

[6] E. J. Candès and C. Fernandez-Granda. "Super-Resolution from Noisy Data". In: *Journal of Fourier Analysis and Applications* 19.6 (2013), pp. 1229–1254.

[7] E. J. Candès and C. Fernandez-Granda. "Towards a Mathematical Theory of Super-Resolution". In: *Communications on Pure and Applied Mathematics* 67.6 (2013), pp. 906–956.

[8] S.S. Chen, D.L. Donoho, and M.A. Saunders. "Atomic decomposition by basis pursuit". In: *SIAM journal on scientific computing* 20.1 (1999), pp. 33–61.

[9] J. F. Claerbout and F. Muir. "Robust modeling with erratic data". In: *Geophysics* 38.5 (1973), pp. 826–844.

[10] G. Dal Maso. *An Introduction to Γ-convergence*. Vol. 8. Progress in nonlinear differential equations and their applications. Boston, MA: Birkhauser, 1993.

[11] Y. de Castro and F. Gamboa. "Exact reconstruction using Beurling minimal extrapolation". In: *Journal of Mathematical Analysis and Applications* 395.1 (2012), pp. 336–354.

[12] Q. Denoyelle, V. Duval, and G. Peyré. "Support Recovery for Sparse Deconvolution of Positive Measures". In: 1506.08264 (2015). URL: http://arxiv.org/abs/1506.08264.

[13] D. L. Donoho. "Superresolution via sparsity constraints". In: *SIAM J. Math. Anal.* 23.5 (1992), pp. 1309–1331.

[14] C. Dossal. "A necessary and sufficient condition for exact recovery by ℓ^1 minimization". In: *Comptes Rendus Mathématique* 350.1 (2011), pp. 117–120.

[15] C. Dossal and S. Mallat. "Sparse spike deconvolution with minimum scale". In: *Proceedings of SPARS.* Nov. 2005, pp. 123–126.

[16] V. Duval and G. Peyré. "Exact Support Recovery for Sparse Spikes Deconvolution". In: *to appear in Foundation of Computational Mathematics* (2015).

[17] V. Duval and G. Peyré. *Sparse Spikes Retrieval on Thin Grids I: the LASSO.* Tech. rep. 2016.

[18] C. Ekanadham, D. Tranchina, and E.P. Simoncelli. "A unified framework and method for automatic neural spike identification". In: *Journal of Neuroscience Methods* 222 (2014), pp. 47–55.

[19] C. Ekanadham, D. Tranchina, and E.P. Simoncelli. "Recovery of Sparse Translation-Invariant Signals With Continuous Basis Pursuit". In: *Signal Processing, IEEE Transactions on* 59.10 (2011), pp. 4735–4744.

[20] C. Fernandez-Granda. "Support detection in super-resolution". In: *Proc. Proceedings of the 10th International Conference on Sampling Theory and Applications* (2013), pp. 145–148.

[21] A. Florescu et al. *A Constrained Optimization Approach For Complex Sparse Perturbed Models.* Tech. rep. Preprint hal-00783298, 2013.

[22] J.J. Fuchs. "On sparse representations in arbitrary redundant bases". In: *IEEE Transactions on Information Theory* 50.6 (2004), pp. 1341–1344.

[23] P. Heins. In: *PhD thesis* (2014).

[24] J. Min et al. "FALCON: Fast and Unbiased Reconstruction of High-Density Super-Resolution Microscopy Data". In: *Scientific Reports* 4577 (2014), pp. 1–9.

[25] J. W. Odendaal, E. Barnard, and C. W. I. Pistorius. "Two-dimensional superresolution radar imaging using the MUSIC algorithm". In: *IEEE Transactions on Antennas and Propagation* 42 (Oct. 1994), pp. 1386–1391.

[26] R. T. Rockafellar, R. J-B. Wets, and M. Wets. *Variational analysis.* Grundlehren der mathematischen Wissenschaften. Berlin, Heidelberg, New York: Springer, 1998.

[27] G. Still. "Discretization in semi-infinite programming: the rate of convergence". In: *Mathematical Programming* 91.1 (2001), pp. 53–69.

[28] G. Tang, B. Narayan Bhaskar, and B. Recht. "Near Minimax Line Spectral Estimation". In: *CoRR* abs/1303.4348 (2013).

[29] G. Tang, B. Narayan Bhaskar, and B. Recht. "Sparse recovery over continuous dictionaries-just discretize". In: *Proc. ACSSC*. IEEE, 2013, pp. 1043–1047.

[30] R. Tibshirani. "Regression shrinkage and selection via the Lasso". In: *Journal of the Royal Statistical Society. Series B. Methodological* 58.1 (1996), pp. 267–288.