LINEAR INVERSE PROBLEMS WITH NONNEGATIVITY CONSTRAINTS THROUGH DIVERGENCES: SPARSITY OF OPTIMISERS

CAMILLE POUCHOL AND OLIVIER VERDIER

Abstract. We pass to continuum in optimisation problems associated to linear inverse problems $y = Ax$ with non-negativity constraint $x \geq 0$. We focus on the case where the noise model leads to maximum likelihood estimation through general divergences, which cover a wide range of common noise statistics such as Gaussian and Poisson. Considering $x$ as a Radon measure over the domain on which the reconstruction is taking place, we show a general sparsity result. In the high noise regime corresponding to $y \notin \{ Ax \mid x \geq 0 \}$ and under a key assumption on the divergence as well as on the operator $A$, any optimiser has a singular part. We hence provide an explanation as to why any possible algorithm successfully solving the optimisation problem will lead to undesirably spiky-looking images when the image resolution gets finer, a phenomenon well documented in the literature. We illustrate these results with several numerical examples inspired by medical imaging.

1. Introduction

We consider linear inverse problems of the type $Ax = y$, where $A$ is a linear operator, and $y$ lies in a suitable linear space. These problems are often associated with natural constraints. One of the most common such constraints is nonnegativity of the unknown, i.e., $x \geq 0$. This happens in various applications, in particular in medical imaging. In this setting, $x \in \mathbb{R}^r$, $y \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times r}$, where $m$ is the number of data points, $r$ is the number of voxels. One important example is that of Positron Emission Tomography (PET) where the unknown image is the activity, which must be nonnegative [23]. Deconvolution problems often also incorporate such constraints [14].

Depending on the noise model, the corresponding (negative) log-likelihood problem typically writes

$$\min_{x \geq 0} D(y, Ax),$$

where $D$ is some kind of divergence functional. If the noise model is Gaussian, for instance, then $D$ is simply the Euclidean distance, whereas if the noise model is Poisson, $D$ is the Kullback–Leibler divergence.

We analyse the effect of an ever increasing resolution, which leads us to regard the unknown $x$ as a function (henceforth denoted $\mu$) in some functional space $X$, and the operator $A$ is now a linear mapping from the space $X$ to $\mathbb{R}^m$. This leads to the optimisation problem

$$\min_{\mu \geq 0} D(y, A\mu).$$

The non-negativity constraint has been proved to cause sparsity in various contexts in optimisation and optimal control [8]. As a result, the right functional
space $X$ to be considered appears to be that of Radon measures, where discrete measures are regarded as sparse, with the operator $A$ defined for all $\mu \in X, \mu \geq 0$ by

$$(A\mu)_i = \int a_i \, d\mu, \quad a_i \geq 0.$$  

Examples of this phenomenon can be seen in optimal control [9, 16], and the same goes for the optimisation problem (1) when $D$ is the Kullback–Leibler divergence [19]. In fact, such a framework may be inherent in the physical model, when built around Poisson point processes as in PET [23, 20].

In the examples above, sparsity (which can arise in the form of Dirac masses) is undesirable as the sought-for image is expected to be at least piecewise smooth. In other contexts, sparsity of the signal must be enforced, as is the case for instance in sparse super-resolution [11].

In the latter case, the situation is completely different as the unknown signal is known to be a sum of (nonnegative) Dirac masses and the aim is to make sure that the optimisation problem has at least one minimiser which is a sum of Dirac masses, thereby hoping to recover the number and support of the unknown discrete measure. A very general treatment of the existence of such sparse optimisers has recently been given in [2, 3].

The goal of the present work is to establish general conditions which the operator $A$ and the divergence $D$ must satisfy so that, when the measured data $y$ is not in $\{A\mu \mid \mu \geq 0\}$, all minimisers will be sparse either in the form of Dirac masses, or, in a weaker sense, as measures having a singular part, in the spirit of [19]. That last paper established that optimal measures are Dirac masses, in the specific case where

(i) the functions $a_i$ are linearly independent and (real) analytic,

(ii) $D$ is the Kullback-Leibler divergence.

The first assumption (i) is restrictive since many practical problems are such that all the functions $a_i$ are compactly supported inside $K$, making analyticity impossible in view of $a_i \geq 0$. The second assumption (ii) also restricts the possible noise models to the single, albeit important, case of Poisson measurements.

In fact, the so-called $\beta$-divergences have attracted interest recently in non-negative matrix factorisation [12], and are now also advocated for in some medical imaging contexts, such as in PET [5]. This is a family of divergences depending on a real parameter $\beta$. This family has the appealing property of interpolating between three common divergences, namely the Itakura–Saito divergence ($\beta = 0$), the Kullback–Leibler divergence ($\beta = 1$) and the Euclidean distance ($\beta = 2$) [7], which correspond to different noise models (see §2.4).

In this article, we first generalise the above result (when (i) holds) to very general divergences. Then, we proceed to treating more general $a_i$ functions, which requires restricting the class of divergences. Informally at this stage, we hence consider

(C) functions $a_i$ which are locally linearly independent and locally analytic,

(H) divergences which are sufficiently averse to zero values whenever the data is non-zero (see the precise statement in equation (H) in §2.2).

Most functions we are aware of are expected to satisfy (C), such as those defining PET (although local linear independence might be difficult to prove for each particular case). Condition (H) limits the class of divergences, as it for instance excludes the Euclidean distance. It still covers a wide class of $\beta$-divergences, f-divergences and Bregman divergences, in particular those which arise in contexts featuring nonnegativity constraints.

Our main result may then informally be stated as follows (see Theorem 3.6 for a precise statement).
**Theorem 1.1.** If \( y \not\in \{ A\mu, \mu \geq 0 \} \) with \( A \) satisfying (C) and \( D \) satisfying (H), then all optimal solutions \( \mu^* \) to (1) have a singular part.

In other words, if the data is not in the image of \( \{ \mu \geq 0 \} \) under the operator \( A \), any optimiser \( \mu^* \) will have a singular part. The condition that the data \( y \) does not belong to the image cone \( \{ A\mu \mid \mu \geq 0 \} \) should be interpreted as a condition on the level of noise: the more noise there is, the more likely it is that this condition be fulfilled. Undesirable sparsity will hence arise in the high-noise regime.

Our results show that the optimisation problem itself leads to sparse results. Thus, any algorithm successfully solving (1) will inevitably lead to undesirably spiky-looking images as one keeps iterating. In the context of medical imaging, this has been observed when using the Maximum-Likelihood-Expectation-Maximisation (ML-EM, also called the Richardson–Lucy algorithm) for solving (1), and has been referred to in the literature as the “night-sky” or the “draughtsboard” effect [25].

Consequently, the same kind of artefacts will be observed for other likelihoods, hinting at the necessity of either early stopping when solving (1) (see [21], [18]), or adding appropriate regularisation terms in the form
\[
\min_{\mu \geq 0} D(y, A\mu) + \lambda R(\mu).
\]
with \( \lambda \) large enough in order to alleviate the issue [5].

In this work, we also provide examples where, solving (1) for some commonly used divergences, reconstructions exhibit the night-sky effect for inverse problems with sufficient noise. In agreement with our theoretical results, this is the case with enough iterations of either a convergent algorithm for solving (1), or a convergent algorithm for solving (1) with \( \lambda \) small.

**Outline of the paper.** The paper is organised as follows. In §2, we define the inverse problem by setting the functional analytic framework as well as the family of divergences being considered, leading to the corresponding maximum likelihood problem. Then, in §3, we analyse the resulting optimisation problem and prove our main result. Section §4 is devoted to discussing assumption (C) in more detail, in the case of a representative toy example and the PET operator for the regular polygon. We finally illustrate our results about sparsity by numerical simulations presented in §5. They feature different contexts taken from medical imaging, with and without regularisation.

**2. Inverse Problem Setup**

**2.1. Linear inverse problem.** We aim at reconstructing an image \( \mu \) defined on a compact \( K \subset \mathbb{R}^p \), \( p \geq 1 \). The unknown \( \mu \) is an element of the space of Radon measures, denoted \( \mathcal{M}(K) \), which is the topological dual space of continuous functions over the compact, denoted \( \mathcal{C}(K) \). We endow \( \mathcal{M}(K) \) with the weak-* topology, making \( \mathcal{C}(K) \) its dual space. We also recall that the bounded sets of \( \mathcal{M}(K) \) are compact in the weak-* topology, by virtue of the Banach–Alaoglu Theorem.

Finally, we denote the dual pairing between a function \( \mu \in \mathcal{M}(K) \) and a function \( f \in \mathcal{C}(K) \) by \( \langle \mu, f \rangle \), and \( \mathcal{M}_+(K) \) stands for the set of nonnegative Radon measures. Whenever the context is clear, we shall drop the reference to the compact \( K \) and write \( \mathcal{M}, \mathcal{M}_+ \) and \( \mathcal{C} \).

A data point \( y \) is a vector of \( m \) scalar nonnegative measurements, that is, \( y \in \mathbb{R}_+^m \).

This vector itself typically is the realisation of some random variable with mean \( A\mu \), where \( \mu \in \mathcal{M}_+(K) \) is the image to be reconstructed, and \( A \) is a linear mapping \( A: \mathcal{M}(K) \to \mathbb{R}^m \).
The only assumption we make on \( A \) is that it is continuous in the weak-* topology. From [4, Proposition 3.14], this forces \( A \) to be of the form
\[
(A\mu)_i = \langle \mu, a_i \rangle, \quad i = 1, \ldots, m,
\]
where the \( a_i \) are elements of \( \mathcal{C}(K) \). This covers the case of PET [17, 19] and more generally the setting of kernel operators: if the underlying operator in infinite dimension is of the form
\[
\mu \mapsto \int_K k(\cdot, y) \, d\mu(y),
\]
for some smooth kernel \( k \in \mathcal{C}(K \times K) \), the operator \( A \) typically is obtained from sampling for \( m \) points \( x_i \in K \) or integrating the kernel over some subdomains \( \Omega_i \subset K \), namely
\[
a_i = k(x_i, \cdot), \quad \text{or} \quad a_i = \int_{\Omega_i} k(x, \cdot) \, dx, \quad i = 1, \ldots, m.
\]
We will make the assumption that \( A \) maps \( M^+_m(K) \) into the set of (componentwise) nonnegative vectors denoted \( \mathbb{R}^m_+ \) which of course is equivalent to
\[
a_i \geq 0, \quad i = 1, \ldots, m.
\]
Finally, we assume
\[
(3) \quad \sum_{i=1}^m a_i > 0.
\]
Indeed, at any point \( x \) such that \( \sum_{i=1}^m a_i(x) = 0 \), we would have no information on the unknown \( \mu \).

Another consequence of the simple continuity assumption on \( A \) is that its adjoint \( A^*: \mathbb{R}^m \to \mathcal{C}(K) \) is simply defined as
\[
A^*\lambda = \sum_{i=1}^m \lambda_i a_i, \quad \lambda \in \mathbb{R}^m.
\]
We shall sometimes need to know when \( A^* \) is injective. This is of course equivalent to the linear independence of the family \( (a_i)_{i=1, \ldots, m} \). Note that since the codomain of \( A \) is finite dimensional, we have
\[
A^* \text{ is injective } \iff A \text{ is surjective.}
\]
In order to solve the inverse problem (which, without the nonnegativity constraint, would reduce to \( A\mu = y \)), we aim at solving the optimisation problem
\[
(4) \quad \min_{\mu \in M^+_m(K)} D(y, A\mu).
\]
Here, \( D \) stands for some divergence over \( \mathbb{R}^m_+ \), and we will use the notation
\[
\ell(\mu) := D(y, A\mu).
\]
Finally, let us define the notion of support for the various relevant cases (all these cases can be covered in one single definition, but we prefer separating them for clarity):
\begin{itemize}
  \item for a vector \( w \in \mathbb{R}^m_+ \),
    \[ \text{supp}(w) = \{ i \in \{1, \ldots, m\} \mid w_i > 0 \} \]
  \item for a nonnegative function \( f \in \mathcal{C}(K) \),
    \[ \text{supp}(f) = \{ x \in K \mid f(x) > 0 \} \]
    where \( \overline{F} \) stands for the closure of a set \( F \subset K \).
\end{itemize}
• for a nonnegative measure $\mu \in \mathcal{M}_+(K)$, 
  \[ \text{supp}(\mu) := \{ x \in K \mid \mu(N) > 0, \forall N \in N(x) \}. \]
where $N(x)$ is the set of all open neighbourhoods of $x$. Notice that with these notations in place, we may now rewrite assumption (2.1) as 
  \[ \bigcup_{i=1}^m \text{supp}(a_i) = K. \]

2.2. Divergences. We shall assume that $D$ is a separable divergence, i.e.,
  \[ D(u, v) = \sum_{i=1}^m d(u_i, v_i) \quad u, v \in \mathbb{R}_+^m, \]
where $d: \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+ \cup \{+\infty\}$ is a (scalar) divergence. When needed, we will implicitly extend the function $d$ to the whole $\mathbb{R}^2$ by $+\infty$ for $u < 0$ or $v < 0$. Throughout, we will assume that $d$ satisfies the following basic properties

**Separation:** $d(u, v) = 0 \iff u = v$.

**Convexity:** for $u \in \mathbb{R}_+$, $v \mapsto d(u, v)$ is convex on $\mathbb{R}_+$.

**Regularity:** for $u \in \mathbb{R}_+$, the mapping $v \mapsto d(u, v)$ is lower-semicontinuous on $\mathbb{R}_+$.

**Coercivity:** for $u \in \mathbb{R}_+$, the mapping $v \mapsto d(u, v)$ is coercive, i.e.,
  \[ \forall u \in \mathbb{R}_+, \lim_{v \to +\infty} d(u, v) = +\infty. \]

![Figure 1](image.png)

**Figure 1.** An example of map $v \mapsto d(1, v)$ which fulfills the assumptions above. In particular, the derivative tends to $-\infty$ when $v$ approaches zero, and the divergence is zero only at the point $v = 1$. Note that this turns out to be a plot of the function $d_\beta(1, v)$ (see §2.3.1) for $\beta = 1.2$.

The above assumptions ensure that, for a fixed $u \in \mathbb{R}_+$, $v \mapsto d(u, v)$ is subdifferentiable over $\mathbb{R}_+^*$. For our results about sparsity, we will also sometimes consider a subclass of such divergences for which we must specify (non-)subdifferentiability at $v = 0$ as follows:

**H** $v \mapsto d(u, v)$ is not subdifferentiable at 0 if $u > 0$.

From standard convex analysis, this is equivalent to the directional derivative of being $-\infty$ at 0, i.e.,

**H** \[ \iff \forall u > 0, \lim_{t \to 0, t > 0} \frac{d(u, t) - d(u, 0)}{t} = -\infty \]
Finally, we will need one very weak assumption involving both the operator $A$ and the divergence $D$:

\begin{equation}
\text{(5)} \quad \text{there exists } \mu \in \mathcal{M}_+ \text{ such that } D(y, \cdot) \text{ is finite and continuous at } A\mu.
\end{equation}

One sufficient and simpler assumption for the above to hold true is the existence of some $x \in K$ such that the divergences $d(y_i, \cdot)$ are continuous at $a_i(x)$ for all $i \in \{1, \ldots, m\}$, which follows from choosing $\mu = \delta_x$.

2.3. Examples.

2.3.1. Beta Divergences. For $u > 0$, $v > 0$ scalar variables, $\beta \in (1, 2]$, we define

\[ d_\beta(u, v) := \frac{1}{\beta(\beta - 1)} \left( u^\beta + (\beta - 1)v^\beta - \beta uv^{\beta - 1} \right), \]

which for $\beta = 2$ gives the Euclidean distance

\[ d_2(u, v) = \frac{1}{2}(u - v)^2, \]

and by continuity for $\beta = 1$ the Kullback–Leibler divergence

\[ d_1(u, v) = u \log \left( \frac{u}{v} \right) - u + v. \]

The corresponding divergence over $\mathbb{R}_m^+$ will be denoted $D_\beta$.

More precisely, using the convention $0/0 = 0$, $0 \log 0 = 0$, $d_1$ is defined for nonnegative scalars $u \geq 0$, $v \geq 0$ as follows.

\[ d_1(u, v) = \begin{cases} v & \text{if } u = 0, v \geq 0 \\ +\infty & \text{if } u > 0, v = 0 \\ u \log \left( \frac{u}{v} \right) - u + v & \text{if } u > 0, v > 0 \end{cases} \]

Note that the $\beta$-divergences $d_\beta$ satisfy Hypothesis (H) for all $\beta \in [1, 2)$, but $d_2$ does not.

2.3.2. Reverse $f$- and Bregman divergences. Given a convex function $F: \mathbb{R}_+ \to \mathbb{R} \cup \{+\infty\}$ with $F(1) = 0$, one can define the following “reverse $f$-divergence”

\[ d_F^f(u, v) := F(v/u)u \quad u > 0 \quad v \geq 0 \]

Similarly, with a convex function $F: \mathbb{R}_+ \to \mathbb{R} \cup \{+\infty\}$ such that $F$ is differentiable, we define the “reverse Bregman-divergence”

\[ d_F^B(u, v) := F(v) - F(u) - F'(u)(v - u) \quad u > 0 \quad v \geq 0 \]

Both fulfill our assumptions as long as

- $F'(0) = -\infty$
- They are suitably extended to the case $u = 0$ such that $d(0, \cdot)$ is a convex function with finite values on $\mathbb{R}_+$.

Note that $\beta$-divergences are not reverse $f$-divergences nor reverse Bregman-divergences for $\beta \in (1, 2)$, $d_1$ is a reverse $f$-divergence but not a reverse Bregman-divergence, and $d_2$ is a reverse Bregman-divergence but not a reverse $f$-divergence.
2.4. The noise model. In general, $y$ is drawn according to a distribution parameterised by $w = A\mu$ as well as some additional dispersion parameter $\phi > 0$ controlling the noise level. For a better understanding of why $y \notin A(M_+)$ may be considered a situation typical of high noise, let us review the underlying statistical model in the specific case of $\beta$-divergences.

In this setting, the minimisation problem \((2.1)\) is (up to constants) the corresponding (negative-log) likelihood maximum problem. We write the statistical model for $y$ and $w$ as scalar variables, the full statistical model is defined componentwise by independent draws of the same form.

A general way to write the noise model giving rise to $\beta$-divergences is to use the so-called Tweedie distributions. The $\beta$-divergences are a special case of such distributions, as the corresponding Tweedie distribution is given by

$$y \mapsto h(x, \phi) \exp\left(-\frac{1}{\phi}d_\beta(y, w)\right),$$

so that minimising the negative log-likelihood problem indeed is equivalent to minimising $w \mapsto d_\beta(y, w)$. We refer to [24] for more details.

The underlying density is not always tractable (this is the case if $1 < \beta < 2$), making the noise model unclear. In other cases, the noise model can be further identified as follows.

**Case $\beta = 2$:** the noise model is Gaussian, i.e.

$$y \sim \mathcal{N}(w, \phi).$$

**Case $1 < \beta < 2$:** up to our knowledge, no explicit model is known.

**Case $\beta = 1$:** the noise model is Poisson, i.e.

$$y \sim \phi \mathcal{P}(\frac{1}{\phi}w).$$

Note that the first example $\beta = 2$ is the only one which does not (necessarily) lead to nonnegative data $y \in \mathbb{R}^m_+$. Yet, we will always make this assumption throughout.

A common property of these noise models is that the random variable $y$ tends to $w$ as the dispersion parameter $\phi$ goes to zero:

$$y \to w \text{ a.s., as } \phi \to 0.$$

Hence, the less noise there is, the more likely it is that $y \in A(M_+)$. For a more quantitative version of this statement in the case of $d = d_1$, i.e., Poisson distributed measurements, see [19].

3. Optimisation problem

We now investigate the optimisation problem \((2.1)\), starting with a related optimisation problem and its dual.

3.1. Related optimisation problem and its dual. We consider the cone

$$A(M_+) = \{ A\mu \mid \mu \in M_+ \},$$

which is clearly convex. We notice that it is also closed.

**Lemma 3.1.** The cone $A(M_+)$ is closed.

**Proof.** Pick a sequence $w_n = A\mu_n$ in $A(M_+)$ converging to some $w \in \mathbb{R}^m_+$. Appealing to \((2.1)\), we choose $c > 0$ such that $\sum_{i=1}^m a_i \geq c$ and write

$$\mu_n(K) \leq \frac{1}{c} \int_K \left(\sum_{i=1}^m a_i\right) d\mu_n = \frac{1}{c} \sum_{i=1}^m (w_n)_i,$$
where the last quantity is bounded since \((w_n)\) converges. Hence, the sequence \((\mu_n)\) is bounded, it has a weak-* converging subsequence, say to some \(\mu \in \mathcal{M}_+\). Along the subsequence, \((A\mu_n)\) converges to \(A\mu\) by weak-* continuity of \(A\), which proves that \(w = A\mu \in A(\mathcal{M}_+)\).

The original optimisation problem (2.1) is related to the following optimisation problem:

\[
\min_{w \in A(\mathcal{M}_+)} D(y, w).
\]

More precisely, for any \(w^*\) optimal for the above problem, any measure \(\mu^* \in \mathcal{M}_+\) such that \(A\mu^* = w^*\) is optimal for the original problem.

**Lemma 3.2.** The minimum in (2.1) is attained.

**Proof.** We show that the optimal value in the minimisation problem (3.1) is attained. The function \(w \mapsto D(y, w)\) is coercive and lower semi-continuous. Since \(A(\mathcal{M}_+)\) is closed (Lemma 3.1), there exists an optimal \(w^* \in A(\mathcal{M}_+)\) for the problem (3.1). Any \(\mu^* \in \mathcal{M}_+\) such that \(A\mu^* = w^*\) then provides a minimiser for the original problem (2.1). □

We now compute the (Lagrange) dual problem to the problem (3.1). These computations turn out to be crucial in analysing results from numerical simulations, determining whether we should expect sparsity or not for some given \(y \in \mathbb{R}^m_+\) from appropriately defined sparsity certificates.

We define the cone \(A(\mathcal{M}_+)^* := \{ \lambda \in \mathbb{R}^m \mid \langle \lambda, w \rangle \geq 0, \forall w \in A(\mathcal{M}_+) \}\) dual to \(A(\mathcal{M}_+)\), which can be characterised as in [13] by

\[
A(\mathcal{M}_+)^* = \{ \lambda \in \mathbb{R}^m \mid A^* \lambda \geq 0 \text{ on } K \}.
\]

The dual problem writes

\[
\max_{\lambda \in A(\mathcal{M}_+)^*} g(\lambda),
\]

where the function \(g: \mathbb{R}^m \to \mathbb{R}\) is defined for \(\lambda \in A(\mathcal{M}_+)^*\) by

\[
g(\lambda) := \min_{w \in \mathbb{R}^m_+} D(y, w) - \langle \lambda, w \rangle.
\]

Note that strong duality holds since the problem is convex and Slater’s condition is obviously satisfied since the minimisation occurs over a cone in finite dimension.

The point of taking the Lagrange dual is that, since \(D\) decomposes, so does \(g\) and we find

\[
g(\lambda) = \sum_{i=1}^m \min_{w_i \geq 0} (d(y_i, w_i) - \lambda_i w_i).
\]

As a result, defining

\[
h(y, \lambda) := \min_{w \geq 0} (d(y, w) - \lambda w)
\]

for a scalar \(y \geq 0\), the resulting dual function will take the form

\[
g(\lambda) = \sum_{i=1}^m h(y_i, \lambda_i).
\]

The explicit computation of the functions \(h\) in (3.1) is carried out for the case of \(\beta\)-divergences in Appendix A.
3.2. Optimality conditions. We now use Fenchel duality to compute the optimality conditions for problem (2.1).

**Proposition 3.3.** Let \( \mu^\ast \) be an optimal measure for (2.1). Then, there exists \( \lambda^\ast \in \partial D(y,\cdot)(A\mu^\ast) \) such that
\[
A^\ast \lambda^\ast \geq 0 \text{ on } K, \quad A^\ast \lambda^\ast = 0 \text{ on } \text{supp}(\mu^\ast).
\]

*Proof.* We use the Fenchel–Rockafellar Theorem [22], with the optimisation problem rewritten as
\[
\min_{\mu \in \mathcal{M}} D(y, \mu) + \delta_{\mathcal{M}_+}(\mu).
\]
In this context, we need *paired spaces*. The natural choice is \( \mathcal{C} \) (endowed with its strong topology) and \( \mathcal{M} \) (endowed with its weak-* topology). We may apply the theorem thanks to the hypothesis (2.2) and we obtain
\[
\mu^\ast \text{ is optimal } \iff 0 \in A^\ast \partial D(y,\cdot)(A\mu^\ast) + N_{\mathcal{M}_+}(\mu^\ast),
\]
where \( N_{\mathcal{M}_+}(\mu) \) is the normal cone of \( \mathcal{M}_+ \) at \( \mu \), defined by
\[
N_{\mathcal{M}_+}(\mu) := \{ f \in \mathcal{C} \mid \forall \nu \in \mathcal{M}_+, \langle \mu - \nu, f \rangle \geq 0 \}.
\]
The normal cone can be identified as
\[
N_{\mathcal{M}_+}(\mu) = \{ f \in \mathcal{C} \mid f \leq 0 \text{ on } K, f = 0 \text{ on } \text{supp}(\mu) \},
\]
see [19]. Picking \( \lambda^\ast \in \partial D(y,\cdot)(A\mu^\ast) \), we exactly obtain (3.3).

Note that for \( \mu^\ast \) optimal and with the notations of the above result, the separable form of \( D \) ensures
\[
\lambda^\ast \in \partial D(y,\cdot)(A\mu^\ast) \iff \forall i \in \{1,\ldots,m\}, \lambda^\ast_i \in \partial d(y_i,\cdot)((A\mu^\ast)_i).
\]

**Remark 3.4.** Let us then emphasise the following important (yet straightforward) property: if \( \mu^\ast \) is optimal, then any \( \lambda^\ast \in \partial D(y,\cdot)(A\mu^\ast) \) is such that
\[
y_i \neq (A\mu^\ast)_i \implies \lambda^\ast_i \neq 0.
\]
Indeed, assume \( \lambda^\ast_i = 0 \). Then the inclusion \( \lambda^\ast_i = 0 \in \partial d(y_i,\cdot)((A\mu^\ast)_i) \) is equivalent to \((A\mu^\ast)_i\), minimising \( d(y_i,\cdot) \), which by the separation property enforces \( y_i = (A\mu^\ast)_i \).

3.3. Sparsity theorems. We now address the following question

If the data \( y \) is not in the image cone \( A(\mathcal{M}_+) \), when do the optimality conditions (3.3) lead to sparse measures?

By sparse, we mean measures that have a singular part (but may have an absolutely continuous part) with respect to the Lebesgue measure.

We start by a result ensuring that all optimal measures are sums of Dirac masses, but that hold only under the restrictive assumption that the functions \( a_i \) be analytic.

**Proposition 3.5.** Assume that the functions \( (a_i)_{i=1,\ldots,m} \) are linearly independent in \( \mathcal{C}(K) \) and are analytic in \( K \) with \( K \) connected.

Then, if \( y \notin A(\mathcal{M}_+) \), any optimal measure is a sum of Dirac masses.

*Proof.* Let us pick some optimal measure \( \mu^\ast \). The optimality conditions (3.3) provide \( \lambda^\ast \) such that \( \lambda^\ast \in \partial D(y,\cdot)(A\mu^\ast) \) and \( A^\ast \lambda^\ast = 0 \) on \( \text{supp}(\mu^\ast) \). In other words, \( \text{supp}(\mu^\ast) \) is contained in the zeros of the continuous function \( \varphi^\ast := A^\ast \lambda^\ast = \sum_{i=1}^m \lambda^\ast_i a_i \).

Since \( y \notin A(\mathcal{M}_+) \), there exists \( i \) such that \( y_i \neq (A\mu^\ast)_i \). Using Remark 3.4, this implies \( \lambda^\ast_i \neq 0 \). Now, the linear independence of the functions \( a_i \) implies that \( \varphi^\ast \) is not identically zero. Since the function is by assumption analytic, its zeros must be isolated in the connected set \( K \). Hence, the support of \( \mu^\ast \) is reduced to a set of singletons, which concludes the proof. \( \square \)
The main reason why this theorem is not satisfactory is that many applications feature functions \(a_i\) that are compactly supported inside \(K\), although they may be analytic (or piecewise analytic) on their support. This is why we now consider the case where the functions \(a_i\) may only be piecewise analytic. Dealing with this more complicated case also requires restricting our attention to divergences which satisfy the property C.

We make some assumptions on the functions \((a_i)\), which require some notations. First, for \(i \in \{1, \ldots, m\}\), we let

\[
J_i := \{ j \in \{1, \ldots, m\}, \ supp(a_j) \cap supp(a_i) \neq \emptyset \},
\]

be the set storing indices of functions \(a_j\) which are active on \(supp(a_i)\); in particular, we obviously have \(i \in J_i\).

We consider the following condition, which combines local linear independence and local analyticity:

(C) there exists a partition of \(K = \bigcup_{k=1}^{\nu} \Omega_k\) in smooth domains such that
- the functions \((a_i)\), \(i \in \{1, \ldots, m\}\) are piecewise analytic on the partition,
- for all \(i \in \{1, \ldots, m\}\), the family \((a_j)_{j \in J_i}\) is linearly independent in \(C(\Omega_k)\) for any \(k\) such that \(\Omega_k \cap supp(a_i) \neq \emptyset\).

Then, the following holds.

**Theorem 3.6.** Assume that the divergence \(D\) satisfies \(\text{C}\), and that the functions \((a_i)_{i=1, \ldots, m}\) satisfy \(\text{C}\). Then, if \(y \notin A(M)\), any \(\mu^*\) optimal has a singular part.

More precisely, for any \(i_0 \in \{1, \ldots, m\}\) such that \(y_{i_0} \neq (A\mu^*)_i\), \(\mu^*|_{supp(a_{i_0})}\) is a singular measure.

**Proof.** Let \(\mu^*\) be optimal. We may write the optimality conditions (3.3), namely

\[
A^*\lambda^* \geq 0 \text{ on } K, \quad A^*\lambda^* = 0 \text{ on } supp(\mu^*),
\]

where \(\lambda^*_i \in \partial d(y_{i_*})(w^*_i)\) for all \(i\). We shall prove that the following key property holds:

\[(P) \quad \forall i \in \{1, \ldots, m\}, \quad y_i \neq (A\mu^*)_i \quad \Rightarrow \quad (A\mu^*)_i > 0.
\]

Indeed, let us pick \(i_0\) such that \((A\mu^*)_{i_0} = 0\), and we need to show that \(y_{i_0} = (A\mu^*)_{i_0} = 0\). If we had \(y_{i_0} > 0\), we would find

\[
\lambda_{i_0} \in \partial d(y_{i_*})((A\mu^*)_{i_0}) = \partial d(y_{i_*})(0),
\]

contradicting the emptiness of the subdifferential as given by \(\text{C}\), and hence proving \(P\).

Continuing with \(i_0 \in \{1, \ldots, m\}\) such that \(y_{i_0} \neq (A\mu^*)_{i_0}\) and hence such that \(\lambda_{i_0}^* \neq 0\) as explained in Remark 3.4 Property \(P\) entails

\[
\text{(10)} \quad supp(\mu^*) \cap supp(a_{i_0}) \neq \emptyset.
\]

The optimality of \(\mu^*\) also ensures

\[
supp(\mu^*) \subset \left\{ x \in K \left| \sum_{j=1}^{m} \lambda_{i}^* a_j(x) = 0 \right. \right\}.
\]

Hence, we have

\[
supp(\mu^*) \cap supp(a_{i_0}) \subset \left\{ x \in supp(a_{i_0}) \left| \sum_{j \in J_{i_0}} \lambda_{i}^* a_j(x) = 0 \right. \right\},
\]

and the set on the right-hand side cannot be empty owing to (3.3). We denote it \(K_{i_0}\).

Now let us analyse this set more precisely, by letting \(k\) such that \(\Omega_k\) intersects \(supp(a_{i_0})\). \(K_{i_0}\) may contain parts of the edges \(\partial \Omega_k\). Other than that, it cannot
have any limit point in the interior of $\Omega_k$. It it were the case, analyticity would imply $\sum_{j \in J_{0}} \lambda_{j}^{*}a_{j} = 0$ identically on $\Omega_k$. Then, linear independence would enforce $\lambda_{j}^{*} = 0$ for all $j \in J_{0}$, contradicting $\lambda_{i_{0}}^{*} \neq 0$. Hence, the support of $\mu|_{\text{supp}(a_{i_{0}})}$ is included in a union of singletons (lying in the interior of the partition pieces) and of $(d - 1)$-dimensional edges (lying at their boundary), which concludes the proof. □

Remark 3.7. We emphasise that Theorem 3.6 does not use the additional condition $\text{supp}(\mu^{*}) \subset \arg \min \left( \sum_{i = 1}^{m} \lambda_{i}^{*}a_{i} \right)$, which, if anything, can only make measures “more sparse”. For instance, one has for all $k$

$$\text{supp}(\mu^{*}) \cap \Omega_k \subset \left\{ x \in \Omega_k \mid \sum_{j = 1}^{m} \lambda_{j}^{*} \nabla a_{j}(x) = 0 \right\}.$$

4. Functions $a_{i}$ that lead to sparsity

The purpose of this section is to explain why, in some sense, functions encountered in practice are expected to fulﬁl condition [C].

4.1. Shifted polynomials in dimension 1. Let $\mathbb{R}_k[X]$ denote the vector space of real polynomials of degree $k$ over $\mathbb{R}$.

Lemma 4.1. Let $P \in \mathbb{R}_k[X]$ and $z_1, \ldots, z_r$ be distinct in $\mathbb{R}$. Then the family $(P(\cdot + z_1), \ldots, P(\cdot + z_r))$ is linearly independent in $\mathbb{R}_k[X]$ if $r \leq k + 1$.

Proof. Since $P$ is in $\mathbb{R}_k[X]$, the family $(D^\alpha P)_{0 \leq \alpha \leq d}$ is a basis of $\mathbb{R}_k[X]$. Now, using the Taylor formula for polynomials

$$\forall z \in \mathbb{R}, \quad P(X + z) = \sum_{\alpha = 0}^{k} z^\alpha \frac{D^\alpha P}{\alpha!}(X),$$

the idea is to notice that the matrix of the family $(P(\cdot + z_1), \ldots, P(\cdot + z_r))$ in that basis is nothing but the Vandermonde matrix given by

$$\begin{pmatrix}
1 & 1 & \ldots & 1 \\
z_1 & z_2 & \ldots & z_r \\
z_1^2 & z_2^2 & \ldots & z_r^2 \\
\vdots & \vdots & \ddots & \vdots \\
z_1^k & z_2^k & \ldots & z_r^k
\end{pmatrix}.$$  

Hence, since the shifts $z_1, \ldots, z_r$ are distinct, this matrix is of rank $r$ as soon as $r \leq k + 1$, which concludes the proof. □

Now, consider some univariate polynomial $P \in \mathbb{R}_k[X]$ such that $P > 0$ on some interval $(a, b)$, with $P(a) = P(b) = 0$. Let $a_0 = P\mathbbm{1}_{(a,b)}$ and define a countable family of shifted functions by

$$a_0(\cdot - lh), \quad l \in \mathbb{Z},$$

where $h > 0$ is the grid spacing. Suppose $K$ is some compact interval in $\mathbb{R}$, and $m$ stands for the number the above functions whose support intersects the interior of $K$, which we denote $a_1, \ldots, a_m$. For an example, see Figure 2.

Corollary 4.2. Let $r := \left\lfloor \frac{b - a}{h} \right\rfloor$ and assume $r \leq k + 1$. Then the functions $(a_i)_{1 \leq i \leq m}$ satisfy Assumption [C].
Figure 2. Example of shifts with $P(x) := 1 - x^4$, $a = -1$, $b = 1$, $h = \frac{2}{5}$, shown over the compact $K = [0, 3]$. The resulting functions satisfy Assumption (C), as per Corollary 4.2.

**Proof.** The functions $(a_i)_{1 \leq i \leq m}$ are clearly piecewise analytic. Let us tackle the issue of local linear independence in more detail. We let $\Omega \subset K$ be any open interval such that all functions $a_i$ are analytic on $\Omega$. It is easily checked that there are at most $r$ functions active on each such set $\Omega$. Hence, if $i$ is such that $\text{supp}(a_i)$ intersects $\Omega$, a linear combination of $(a_j)_{j \in J_i}$ writes

$$\sum_{\lambda_j \in J_i} \lambda_j P(\cdot + z_j),$$

for some distinct shifts $z_j$ (which are multiples of $h$). This polynomial vanishes on $\Omega$, and hence is the zero polynomial. We may then use Lemma 4.1 to infer that $\lambda_j = 0$ for all $j \in J_i$, hence the local linear independence. □

4.2. PET functions on a regular polygon. In view of illustrating why PET functions are expected to satisfy Assumption (C), let us give the 2D example of PET where the detectors are regularly placed on the unit circle.

Let $\theta = \frac{2\pi}{N}$ for some $N \in \mathbb{N}^*$. Working with complex notations, we denote $z_k = e^{ik\theta}$, $k = 0, \ldots, N - 1$. Define $K$ to be the regular polygon associated to the points $z_k$, which lies inside the unit disk.

For convenience, we change notation and use two indices by letting $a_{j,k}$ be the function associated to the pair of detectors formed by the line segments $[z_j, z_{j+1}]$ and $[z_k, z_{k+1}]$ for $0 \leq j < k \leq N - 1$. There are $m = \frac{N(N - 1)}{2}$ such functions.

First assume that $k = j + 1$, so that the two segments have the point $z_{j+1} = z_k$ in common. Then the function $a_{j,j+1}$ is zero of outside of the triangle defined by the convex hull of $z_j, z_{j+1}, z_{j+2}$. Inside, it is given in complex notations by

$$a_{j,j+1}(z) = \frac{1}{\pi} \text{Arg}\left(\frac{z_{j+2} - z}{z - z_j}\right).$$

Now, assume that $k - j \geq 2$. The lines $(z_{j+1}, z_k)$ and $(z_j, z_{k+1})$ are easily shown to be parallel. Again, $a_{j,k}$ is zero outside of the trapezium defined by the convex hull of $z_j, z_{j+1}, z_k, z_{k+1}$.

With $c$ denoting the point at which the trapezium diagonals intersect, we let $T_{r,s}$ be the triangle defined by the points $z_r, z_s$ and $c$. Then we have:
• If \( z \in T_{j,j+1} \),
  \[
  a_{j,k}(z) = \frac{1}{\pi} \text{Arg} \left( \frac{z - z_{k+1}}{z - z_k} \right).
  \]
• If \( z \in T_{j+1,k} \),
  \[
  a_{j,k}(z) = \frac{1}{\pi} \text{Arg} \left( \frac{z_k - z}{z - z_{j+1}} \right).
  \]
• If \( z \in T_{k,k+1} \),
  \[
  a_{j,k}(z) = \frac{1}{\pi} \text{Arg} \left( \frac{z - z_j}{z - z_{j+1}} \right).
  \]
• If \( z \in T_{k+1,j} \),
  \[
  a_{j,k}(z) = \frac{1}{\pi} \text{Arg} \left( \frac{z - z_j}{z_{k+1} - z} \right).
  \]

The situation is illustrated by Figure 3.

Note that we obviously have the relation.
\[
\forall z \in K, \quad \sum_{0 \leq j < k \leq N-1} a_{j,k}(z) = 1,
\]
By symmetry, we also have for any \( j < k, j' < k' \) such that \( k' - j' = k - j = r \)
\[
\forall z \in K, \quad a_{j',k'}(z) = a_{j,k}(e^{-ir\theta}z),
\]
meaning that all functions can be obtained from \( a_{0,j}, j = 1, \ldots, N - 1 \), up to appropriate rotations, which is reminiscent of the toy example above with functions all equal up to translation. Here, there is not 1 but \( N \) "base" functions from which all the others are deduced by rotation.

It is clear that the functions \( a_i \) are locally analytic in the sense defined by Assumption (C). We conjecture that they also are locally linearly independent.

**Figure 3.** Example of PET for the regular polygon, with \( N = 12 \). In pale blue, the support of \( a_{7,8} \). In orange, the support of \( a_{0,4} \), a trapezium whose diagonals intersect at the point denoted \( c \).
5. Numerical experiments

We here present some simulations of algorithms solving (2.1) in different contexts, where the results exhibit sparsity as expected from the theoretical results. All simulations are run using Python.

5.1. Sparsity Certificates. As evidenced by our results, under the Assumptions [H] and Assumption [C], the relevant criterion for sparsity of optimisers of (2.1) is independent of the specific divergence $D$, as the question reduces to:

$$\text{do we have } y \in A(M_+) ?$$

In practice, as one wants to solve (2.1) (or possibly a regularised version thereof) in the form of some iterative algorithm defined by iterates of the form

$$\mu_{k+1} = G_k(\mu_k),$$

we are looking for methods allowing us to guarantee that $y \notin A(M_+)$. Our strategy is to devise a method to prove that $y \notin A(M_+)$ which writes as a function of $\mu_k$ and that gets better as $k \to +\infty$.

One approach towards this is to make use of duality: by weak duality (and recalling the definition of $g$ in (3.1)) we always have

$$\forall \mu \in M_+, \forall \lambda \in A(M_+)^*, \quad \ell(\mu) \geq g(\lambda).$$

Since $y \in A(M_+)$ $\iff$ $\ell(\mu) = 0$, this entails the following straightforward result:

$$\exists \lambda \in A(M_+)^*, \quad g(\lambda) > 0 \implies y \notin A(M_+).$$

We will call a vector $\lambda \in A(M_+)^*$ such that $g(\lambda) > 0$ a dual certificate of sparsity.

This provides a natural method when it comes to establishing sparsity: assume we have a convergent algorithm for solving (2.1), in the sense that all subsequences of $(\mu_k)_{k \in \mathbb{N}}$ converge (in the weak-* sense) to some minimiser $\mu^*$ of (2.1).

Oftentimes, one can derive an explicit link between primal and dual variables from the relation

$$\lambda^* \in \partial D(y, \cdot)(w^*),$$

which we know hold for optimal $w^*$ under the hypotheses of Theorem 3.6. For instance, in the case of $\beta$-divergences, an optimal dual variable $\lambda^* \in A(M_+)^*$ is related to a primal variable $w^*$ by $\lambda^* := (w^*)^{\beta-2}(w^* - y)$. Then, a candidate of choice for a dual certificate is given by

$$\lambda_k := (A\mu_k)^{\beta-2}(A\mu_k - y).$$

In particular, $\lambda_k$ will converge to $\lambda^*$ along subsequences, hence the convergence of $g(\lambda_k)$ to $\max_{\lambda \in A(M_+)^*} g(\lambda) = \min_{\mu \geq 0} \ell(\mu)$, the last equality being valid since strong duality obtains. Hence, if $y \notin A(M_+)$, we should have $\lim g(\lambda_k) > 0$ as $k \to +\infty$.

A caveat with our choice is that we should only have dual admissibility $\lambda_k \in A(M_+)^*$ at the limit $k \to +\infty$, and not for a fixed iteration number $k$. In practice, we take $k$ large and if $\lambda_k \notin A(M_+)^*$, we set $\tilde{\lambda}_k := \lambda_k + c$ where $c > 0$ is the smallest constant restoring dual admissibility. In other words, we choose the minimal $c > 0$ such that $A^*\tilde{\lambda}_k \geq 0$, i.e., $\tilde{\lambda}_k \in A(M_+)^*$.

5.2. Emission Tomography Example. We first look at an example from PET, where the aim is to solve (2.1). A common way to do so is to use the following iterates, called multiplicative [15] [12]. Starting from some $\mu_0 \in M_+$ (typically a positive constant over the domain), the iterates write

$$\mu_{k+1} = \mu_k \frac{A^*((A\mu_k)^{\beta-2}y)}{A^*((A\mu_k)^{\beta-1})}.$$
These iterates have the ML-EM algorithm ($\beta = 1$) and the Iterative Image Space Reconstruction ($\beta = 2$) as particular cases \cite{10}, and proofs of convergence for these algorithms with any $\beta \in [1, 2]$ can be found in \cite{26}, in the finite-dimensional case. One advantage of these algorithms is the decrease of the functional $\ell$ along iterates, a result we may generalise to the infinite-dimensional setting, see Appendix B adapting a proof of \cite{12}.

We now present the results of applying the algorithm in the case of a 2D PET operator $A$ with 90 views and 64 tangential positions (hence, $m = 90 \times 64 = 5760$). Simulations are run using Python with the Operator Discretization Library \cite{1}. The sought-for image is taken to be the Derenzo phantom, denoted $\mu_r$. The data is obtained by (re-scaled) Poisson draws, with a time-variable (or dose-variable) $t$ which accounts for the level of noise. In other words, $y \sim 1/t \mathcal{P}(tA\mu_r)$, and the higher $t$, the lower the noise. In order to approach the infinite-dimensional setting of our work, we take a high resolution image space discretisation of $512 \times 512$ pixels.

Finally, we take $\beta = 1/2$, on purpose not quite matching the noise statistics, as Poisson noise should lead one to take $\beta = 1$. We hence mimic the situation of not knowing the exact noise statistics.

In Figure 4 and Figure 5, we display the evolution of the loss function $\ell$ along iterates, i.e, $k \mapsto \ell(\mu_k)$, starting from $\mu_0 = 1$, for noise levels $t = 1$ and $t = 10^{-1}$ respectively. As theoretically expected, the function decreases. We also plot the maximum attained for each reconstruction, namely $k \mapsto \max(\mu_k)$, which tends to increase. Finally, we show the reconstruction after $k = 100$ and $k = 1000$ iterates.

In the noisier case $t = 1$, some pixels clearly take over as one keeps iterating. Moreover, we can guarantee that the we should indeed expect sparsity, as we may provide a dual certificates proving that the data $y$ is not in the image cone $A(M_+)$. This is also suggested by the fact that $k \mapsto \ell(\mu_k)$ seems to converge to a positive value rather than to zero.

In the less noisy case $t = 10^{-1}$, it seems that the loss $\ell(\mu_k)$ is not converging to zero, which may be a hint that we should also expect sparsity. However, in this case we are not able to guarantee it with dual certificates. Note that if sparsity were to be true in this case as well, the reason could be that many more iterates are needed to ascertain its presence.

5.3. **Examples with $\beta = 2$ and Regularisation.**

5.3.1. **Toy Example.** We illustrate sparsity with a toy example. There will just be two detectors $a_0$ and $a_1$, so $m = 2$. We choose specifically

$$K = [0, 1], \quad a_0 = 1, \quad a_1(x) = x.$$  

With such analytic $a_i$ functions, \cite[Proposition 3.5]{3} applies.

We also choose $\beta = 2$. Recall from \cite[§2.3.1]{3} that the divergence reduces to the Euclidean distance in this case. We can now compute the sparse solutions explicitly depending on the parameter $y \in \mathbb{R}^2$, as shown in Figure 6. For $y \in \mathbb{R}^2_+$ outside of the cone

$$A(M_+) = \{(y_1, y_2) \in \mathbb{R}^2, \ 0 \leq y_1 \leq y_2\},$$

the solution is of the form $\xi \delta_1$ with $\xi \geq 0$ a varying parameter. For completeness, we also display the solution for $y \notin \mathbb{R}^2_+$ (which is either 0 or $\xi \delta_0$ with $\xi \geq 0$ a varying parameter).

We also look at the effect of regularisation. In this case, we use total variation regularisation, that is, we solve

$$\min_{\mu \geq 0} \ell(\mu) + \rho \text{TV}(\mu),$$
where $\text{TV}(\mu)$ is the total variation of the derivative of the measure $\mu$ and $\rho > 0$ is a regularisation parameter. In this discretised, one-dimensional setting, this is simply $\text{TV}(\mu) = \sum_{i} |\mu_{i+1} - \mu_{i}|$, where $\mu_{i}$ is the value of the discretised measure at pixel $i$.

We then compute the minimum using a primal-dual hybrid gradient method [6]. We plot the resulting minima for various values of the regularisation parameter $\rho$ in Figure 7, in the case $y = (0, 1)$. As $\rho$ goes to 0, the solution approaches the expected sparse limit $\frac{1}{2}\delta_{1}$.

5.3.2. **Tomography example.** We finally look at a more realistic 2D example taken from tomography, where the unknown $\mu$ equals the usual Shepp-Logan phantom used as a benchmark in CT tomography. The example of Figure 8 features an image resolution of $127 \times 127$, and there are 285 angles and 183 tangential coordinates.

We consider the case of the Euclidean distance $d_{2}$. The data $y$ is obtained by Gaussian draws with negative values clipped to 0, i.e., $y \sim \max(\mathcal{N}(A\mu, \sigma^{2}), 0)$, which ensures $y \in \mathbb{R}^{m}_{+}$. We then solve the corresponding TV-regularised problem

$$
\min_{\mu \geq 0} d_{2}(y, A\mu) + \rho \text{TV}(\mu),
$$

with $\rho > 0$, by a primal-dual hybrid gradient method.

As Figure 8 shows, the maximum of the reconstruction gets bigger as $\rho$ tends to 0. In fact, the reconstruction for small $\rho$ clearly exhibits the night-sky effect. The
Figure 5. Case $t = 10^{-1}$. (A) Divergence along iterates. (B) Maximum of reconstruction along iterates. (C) Reconstruction after 100 iterates. (D) Reconstruction after 1000 iterates.

The fact that $\rho = 0$ does not lead to higher maximum values is due to the resolution, which acts as a regulariser for the expected sparse measures.

References

[1] Adler, J., Kohr, H., and Öktem, O. ODL—a Python framework for rapid prototyping in inverse problems. Royal Institute of Technology (2017).

[2] Boyer, C., Chambolle, A., Castro, Y. D., Duval, V., De Gournay, F., and Weiss, P. On representer theorems and convex regularization. SIAM Journal on Optimization 29, 2 (2019), 1260–1281.

[3] Bredies, K., and Carioni, M. Sparsity of solutions for variational inverse problems with finite-dimensional data. Calculus of Variations and Partial Differential Equations 59, 1 (2020), 1–26.

[4] Brezis, H. Functional analysis, Sobolev spaces and partial differential equations. Springer Science & Business Media, 2010.

[5] Cavallanti, Y. C., Oberlin, T., Dobigeon, N., Févotte, C., Stute, S., Ribeiro, M.-J., and Tauber, C. Factor analysis of dynamic PET images: beyond Gaussian noise. IEEE transactions on medical imaging (2010).

[6] Chambolle, A., and Pock, T. A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision 40, 1 (2010), 120–145.

[7] Cichocki, A., and Amari, S.-I. Families of alpha-beta-and gamma-divergences: Flexible and robust measures of similarities. Entropy 12, 6 (2010), 1532–1568.
Figure 6. The sparse solutions of the problem (2.1) with two detectors, $a_0 = 1$, $a_1(x) = x$ on the interval $K = [0, 1]$. There are three distinct regions outside the cone $\mathcal{A}M_+$, but only one which intersects the first quadrant $y \in \mathbb{R}^2_+$, where the optimal solution is a Dirac, given by $\xi \delta_1$ for some $\xi \geq 0$.

Figure 7. The solutions for $y = (0, 1)$ of the problem in Figure 6 with total variation regularisation. Each curve corresponds to a different regularisation parameter, labelled by its base-10 logarithm. We see that when the regularisation parameter goes to zero, the computed solution converges to the expected exact solution, which, we see from Figure 6, is $\mu = .5\delta_1$, depicted by a blue circle.

[8] Clason, C., Kaltenbacher, B., and Resmerita, E. Regularization of ill-posed problems with non-negative solutions. In Splitting Algorithms, Modern Operator Theory, and Applications. Springer, 2019, pp. 113–135.
[9] Clason, C., and Schiela, A. Optimal control of elliptic equations with positive measures. ESAIM: Control, Optimisation and Calculus of Variations 23, 1 (2017), 217–240.
[10] De Pierro, A. R. On the relation between the ISRA and the EM algorithm for positron emission tomography. IEEE transactions on Medical Imaging 12, 2 (1993), 328–333.
[11] Denoyelle, Q., Duval, V., and Peyré, G. Support recovery for sparse super-resolution of positive measures. Journal of Fourier Analysis and Applications 23, 5 (2017), 1153–1194.
[12] Févotte, C., and Idier, J. Algorithms for nonnegative matrix factorization with the $\beta$-divergence. Neural computation 23, 9 (2011), 2421–2456.
[13] Georgiou, T. T. Solution of the general moment problem via a one-parameter imbedding. IEEE transactions on automatic control 50, 6 (2005), 811–826.
[14] Henrot, S.,oussein, C., and Brie, D. Fast positive deconvolution of hyperspectral images. IEEE Transactions on Image Processing 22, 2 (2012), 828–833.
Figure 8. Relation between regularisation parameter and the reconstruction of the Shepp–Logan phantom $\mu$. Here, the data is given by $y_{\text{max}}(\sim \mathcal{N}(A\mu, \sigma^2), 0)$, with $\sigma = 10$. 

[15] Lee, D. D., and Seung, H. S. Algorithms for non-negative matrix factorization. In *Advances in neural information processing systems* (2001), pp. 556–562.

[16] Lohéac, J., Trélat, E., and Zuazua, E. Minimal controllability time for the heat equation under unilateral state or control constraints. *Mathematical Models and Methods in Applied Sciences* 27, 09 (2017), 1587–1644.

[17] Mair, B., Rao, M., and Anderson, J. Positron emission tomography, Borel measures and weak convergence. *Inverse Problems* 12, 6 (1996), 965.

[18] Öktem, O., Pouchol, C., and Verdier, O. Spatiotemporal PET reconstruction using ML-EM with learned diffeomorphic deformation. In *International Workshop on Machine Learning for Medical Image Reconstruction* (2019), Springer, pp. 151–162.

[19] Pouchol, C., and Verdier, O. The ML-EM algorithm in continuum: sparse measure solutions. *Inverse Problems* 36, 3 (2020).

[20] Pouchol, C., and Verdier, O. Statistical model and ML-EM algorithm for emission tomography with known movement. *Journal of Mathematical Imaging and Vision* (2021), 1–14.

[21] Resmerita, E., Engl, H. W., and Iusem, A. N. The expectation-maximization algorithm for ill-posed integral equations: a convergence analysis. *Inverse Problems* 23, 6 (2007), 2575.

[22] Rockafellar, R. T. Extension of Fenchel’s duality theorem for convex functions. *Duke Mathematical Journal* 33 (1966), 81–89.
Let us denote
\[ \psi_y(w) := d_\beta(y, w) - \lambda w. \]

**A.1. Case \( \beta = 1 \)**: We have \( \psi_y(w) \sim (1 - \lambda)w \) as \( w \to +\infty \), hence \( h(y, \lambda) = -\infty \) if \( \lambda > 1 \). If \( \lambda = 1 \) and \( y > 0 \), \( \psi_y(w) \sim -y \ln(w) \) as \( w \to 0 \): the function tends to \(-\infty \) as \( w \to 0 \) and \( h(y, \lambda) = -\infty \). If \( y = 0 \), the function equals 0 identically and its minimum is 0.

We now focus on the case \( \lambda < 1 \). We still have \( \psi_y(w) \sim (1 - \lambda)w \) as \( w \to +\infty \).

If \( y > 0 \), \( \psi_y(w) \sim -y \ln(w) \) as \( w \to 0 \), thus the function tends to \(+\infty \) at both ends. Since \( \psi_y \) is strictly convex in this case, it has a unique minimum for \( w > 0 \), which we again denote \( w(y, \lambda) \), solving
\[ w(y, \lambda)^{-1}(w(y, \lambda) - y) = \lambda \iff w(y, \lambda) = \frac{y}{1 - \lambda}. \]

If \( y = 0 \), it is easily seen that the function \( \psi_y(w) = \frac{1}{\beta}w^\beta - \lambda w \) is minimised at \( w = 0 \), with value 0.

We may also gather the cases \( y = 0 \) and \( y > 0 \) whenever \( \lambda \leq 1 \), since the formula for \( w(y, \lambda) \) shows that it vanishes with \( y \).

Summing up, we find
\[ h(y, \lambda) = \begin{cases} -\infty & \text{if } \lambda > 1 \\ d_\beta(y, w(y, \lambda)) - \lambda w(y, \lambda) & \text{if } \lambda \leq 1 \end{cases} \]

In the last case, further computations lead to
\[ d_\beta(y, w(y, \lambda)) - \lambda w(y, \lambda) = y \ln(1 - \lambda). \]

**A.2. Case \( 1 < \beta < 2 \)**: If \( y > 0 \), \( \psi_y(w) \sim \frac{w^\beta}{\beta} \) as \( w \to +\infty \), and \( \psi_y(w) \sim -\frac{w}{\beta - 1}w^{\beta - 1} \) as \( w \to 0 \). The derivative \( \psi'_y(w) \) satisfies \( \psi'_y(w) \sim -yw^{\beta - 2} \) as \( w \to 0 \), and \( \psi'_y(w) \sim w^{\beta - 1} \) as \( w \to +\infty \). Thus the function \( \psi'_y \) increases (as \( \psi \) is convex) from \(-\infty \) to \(+\infty \). As a consequence, it has a unique minimum \( w > 0 \), which we again denote \( w(y, \lambda) \), solving
\[ w(y, \lambda)^{\beta - 2}(w(y, \lambda) - y) = \lambda. \]

If \( y = 0 \) and \( \lambda \leq 0 \), it is easily seen that the function \( \psi_y(w) = \frac{1}{\beta}w^\beta - \lambda w \) is minimised at \( w = 0 \), with value 0, whereas if \( \lambda > 0 \), it has a unique minimum (also defined by the equation for \( w(y, \lambda) \)). Summing up, we find
\[ h(y, \lambda) = \begin{cases} 0 & \text{if } y = 0, \lambda \leq 0 \\ d_\beta(y, w(y, \lambda)) - \lambda w(y, \lambda) & \text{otherwise.} \end{cases} \]

**A.3. Case \( \beta = 2 \)**: In this case, \( \psi_y \) has a unique minimum \( w \) given by \( w = \lambda + y \), which gives the explicit formula
\[ h(y, \lambda) = -\frac{1}{2}(\lambda + y)^2 + \frac{1}{2}y^2. \]

**Appendix A. Computation of the dual**
Appendix B. Monotonicity along multiplicative updates

We let

\[ \tilde{K} := \bigcup_{i \in \text{supp}(y)} \text{supp}(a_i). \]

Proposition B.1. Assume that \( \beta \in [1, 2] \). For any \( \mu_0 \in \mathcal{M}_+ \) with \( \tilde{K} \subset \text{supp}(\mu_0) \), the iterates

\[ \mu_{k+1} = \mu_k \frac{A^*((A\mu_k)^{\beta-2}y)}{A^*((A\mu_k)^{\beta-1})}, \]

are well-defined and they make \( \ell \) decrease, i.e.,

\[ \forall k \in \mathbb{N}, \quad D_\beta(y, A\mu_{k+1}) \leq D_\beta(y, A\mu_k). \]

Proof. Let us check that the assumption \( \tilde{K} \subset \text{supp}(\mu_0) \) implies \( \text{supp}(\mu_k) = \tilde{K} \) along iterates \( k \geq 1 \), making all divisions well-defined (with the convention \( 0/0 = 0 \)). All we have to prove is that for all \( k \in \mathbb{N}^* \), there holds

- \( \text{supp}(y) \subset \text{supp}(A\mu_k) \),
- \( \text{supp}(\mu_k) = \tilde{K} \).

Then, the product \( (A\mu_k)^{\beta-2}y \) is indeed well-defined with the convention \( 0/0 = 0 \), and whenever the denominator \( A^*((A\mu_k)^{\beta-1}) \) vanishes, so does the numerator \( A^*((A\mu_k)^{\beta-2}y) \).

We proceed recursively. The second point is clear since we have \( \text{supp}(\mu_{k+1}) \subset \text{supp}(\mu_k) \), and \( \text{supp}(A^*((A\mu_k)^{\beta-2}y)) = \tilde{K} \). For the first point, we fix \( i \in \text{supp}(y) \) and prove that \( (A\mu_{k+1})_i > 0 \). Notice that the denominator \( A^*((A\mu_k)^{\beta-1}) \) is bounded from above by some positive constant \( M_k \). We may write

\[ (A_{k+1})_i \geq y_i (A\mu_k)_i^{\beta-2} \int_{\tilde{K}} a^2 \frac{a_i}{\mu_k(K)} y \, d\mu_k \geq \frac{y_i (A\mu_k)_i^{\beta-2}}{M_k} \int_{\tilde{K}} a^2 \, d\mu_k \]

\[ \geq \frac{y_i (A\mu_k)_i^{\beta-2}}{M_k \mu_k(K)} (A\mu_k)_i^\beta = \frac{y_i (A\mu_k)_i^\beta}{M_k \mu_k(K)} > 0, \]

where we have used the Cauchy–Schwarz inequality \( \int_{\tilde{K}} a^2 \, d\mu_k \geq \frac{(A\mu_k)_i^2}{\mu_k(K)} \).

For \( k \in \mathbb{N} \) and with the notation \( w_k := A\mu_k \), consider the surrogate function, defined for \( \mu \in \mathcal{M}_+ \) absolutely continuous with respect to \( \mu_k \) by

\[ G_k(\mu) = \sum_{i=1}^{m} \int_{K} d_\beta(y, (w_k)_i) \frac{d\mu}{d\mu_k} \frac{a_i}{(w_k)_i} \, d\mu_k. \]

Step 1. Let us prove that \( G_k \) is above \( \ell \) and coincides with it at \( \mu_k \). We compute

\[ G_k(\mu_k) = \sum_{i=1}^{m} \int_{K} d_\beta(y, (w_k)_i) \frac{a_i}{(w_k)_i} \, d\mu_k = \sum_{i=1}^{m} \int_{K} (w_k)_i \, d\mu_k = D_\beta(y, \mu_k) = \ell(\mu_k) \]

Then, for a given \( \mu \in \mathcal{M}_+ \) absolutely continuous with respect to \( \mu_k \) and by the convexity of \( v \mapsto d_\beta(u, v) \) with the probability measure \( \nu_k := \frac{a_i}{(w_k)_i} \, d\mu_k \),

\[ \int_{K} d_\beta(y, (w_k)_i) \frac{d\mu}{d\mu_k} \, d\nu_k \geq d_\beta(y, \int_{K} (w_k)_i \frac{d\mu}{d\mu_k} \, d\nu_k) = d_\beta(y, (A\mu)_i). \]

Summing over \( i \) yields the result.

Step 2. We now prove

\[ G_k(\mu_{k+1}) \leq G_k(\mu_k). \]
From the convexity of $v \mapsto d_\beta(v, u)$ and for $i \in \text{supp}(y)$,
\[
d_\beta \left( y_i, (w_k), \frac{d\mu_{k+1}}{d\mu_k} \right) \geq d_\beta \left( y_i, (w_k), \frac{d\mu_{k+1}}{d\mu_k} \right)
+ \partial_v d_\beta \left( y_i, (w_k), \frac{d\mu_{k+1}}{d\mu_k} \right) (w_{k+1}, \frac{d\mu_{k+1}}{d\mu_k}),
\]
which yields
\[
G_k(\mu_k) \geq G(\mu_{k+1}) + \sum_{i=1}^m \int_K \partial_v d_\beta \left( y_i, (w_k), \frac{d\mu_{k+1}}{d\mu_k} \right) (w_{k+1}, \frac{d\mu_{k+1}}{d\mu_k}) \frac{a_i}{(w_k)_i} \, d\mu_k
= G(\mu_{k+1}) + \int_K \sum_{i=1}^m a_i \partial_v d_\beta \left( y_i, (w_k), \frac{d\mu_{k+1}}{d\mu_k} \right) d(\mu_k - \mu_{k+1}).
\]
We are done if we prove that
\[
\sum_{i=1}^m a_i \partial_v d_\beta \left( y_i, (w_k), \frac{d\mu_{k+1}}{d\mu_k} \right) = A^* \partial_v d_\beta(y, w_k) \frac{d\mu_{k+1}}{d\mu_k}
\]
on $K$. Since $\frac{d\mu_{k+1}}{d\mu_k} = \frac{A^*(w_{k+1}^{\beta-2} y)}{A^*(w_k^{\beta-2} y)} = f_{k+1}$, and $\partial_v d(u, v) = v^{\beta-2} (v - u)$, we compute
\[
A^* \partial_v d_\beta(y, w_k f_{k+1}) = f_{k+1}^{\beta-2} A^* (w_k^{\beta-1} f_{k+1} - w_k^{\beta-2} y)
= f_{k+1}^{\beta-2} \left( f_{k+1} A^* w_k^{\beta-1} - A^* (w_k^{\beta-1} y) \right) = 0
\]
by the definition of $f_{k+1}$. We may now conclude as the combination of the two results entails for all $k \in \mathbb{N}$
\[
\ell(\mu_{k+1}) \leq G_k(\mu_{k+1}) \leq G_k(\mu_k) = \ell(\mu_k).
\]

\[
\Box
\]

Université de Paris, FP2M, CNRS FR 2036, MAP5 UMR 8145, F-75006 Paris, France.
Email address: camille.pouchol@u-paris.fr

Department of Mathematics, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden.
Email address: olivierv@kth.se

Department of Computing, Electrical Engineering and Mathematical Sciences, Western Norway University of Applied Sciences, Bergen, Norway.
Email address: olivier.verdier@hvl.no