EXACT RECONSTRUCTION USING SUPPORT PURSUIT

YOHANN DE CASTRO AND FABRICE GAMBOA

Abstract. In this article we show that measures, with finite support on the real line, are the unique solution to an algorithm, named support pursuit, involving only a finite number of generalized moments (which encompass the standard moments, the Laplace transform, the Stieljes transformation, etc...).

The support pursuit share related geometric properties with basis pursuit of Chen, Donoho and Saunders [CDS98]. As a matter of fact we extend some standard results of compressed sensing (the dual polynomial, the nullspace property) to the signed measure framework.

We express exact reconstruction in terms of a simple interpolation problem. We prove that every nonnegative measure, supported by a set containing $s$ points, can be exactly recovered from only $2s + 1$ generalized moments. This result leads to a new construction of deterministic sensing matrices for compressed sensing. In particular, we prove that one can recover all nonnegative $s$-sparse vectors from only $2s + 1$ linear measurements.

Introduction

In the last decade many emphasis has been put on the exact reconstruction of sparse finite dimensional vectors using the basis pursuit algorithm. The pioneer paper of S. S. Chen, D. L. Donoho and M. A. Saunders [CDS01] has brought this method to the statistic community. Notice that the seminal ideas on the subject appeared in quite older works of D. L. Donoho and P. B. Stark [DS89]. Therein, mainly the discrete Fourier transform is considered. Behind, P. Doukhan, E. Gas- siat and one author of this present paper [DG96, GG96] considered the exact reconstruction of a nonnegative measure. More precisely, they derived results when one only knows the values of a finite number of linear functionals at the target measure.

In this paper, we concern with the measure framework. We show that the exact reconstruction of a signed discrete measure $\sigma$ on a set $I$ is still possible when one only knows a finite number of non-adaptive linear measurements. Surprisingly our method, called support pursuit, appears to uncover exact reconstruction results related to basis pursuit.

Let us explain more precisely what is done here. Consider a signed discrete measure $\sigma$ on a set $I$. Unless otherwise specified, assume that $I := [-1, 1]$. Notice that all our results easily extend to any real bounded set. Consider the Jordan decomposition,

$$\sigma = \sigma^+ - \sigma^-.$$
and denote by $S^+$ (resp. $S^-$) the support of $\sigma^+$ (resp. $\sigma^-$). Let us define the Jordan support of the measure $\sigma$ as the pair $J := (S^+, S^-)$. Assume further that $S := S^+ \cup S^-$ is finite and has cardinality equal to $s$. Moreover suppose that $J$ belongs to a family $\mathcal{Y}$ of pairs of subsets of $I$ (see Definition 1 for more details).

We call $\mathcal{Y}$ a Jordan support family. The measure $\sigma$ can be written as

$$\sigma = \sum_{i=1}^{s} \sigma_i \delta_{x_i},$$

where $S = \{x_1, \ldots, x_s\}$, $\sigma_1, \ldots, \sigma_s$ are nonzero real numbers and $\delta_x$ denotes the Dirac measure at point $x$.

Let $F = \{u_0, u_1, \ldots, u_n\}$ be any family of continuous functions on $I$, where the set $I$ denotes the closure of $I$ (this statement is meant to be general and encompasses the case where $I$ is not closed). Let $\mu$ be a signed measure on $I$, the $k$-th generalized moment of $\mu$ is defined by

$$c_k(\mu) = \int_I u_k \, d\mu,$$

for all the indices $k = 0, 1, \ldots, n$.

Our main issue. In this paper we concern with the reconstruction of the target measure $\sigma$ from the observation of $K_n := (c_0(\sigma), \ldots, c_n(\sigma))$, i.e. its $(n+1)$ first generalized moments. We assume that both the support $S$ and the weights $\sigma_i$ of the target measure $\sigma$ are unknown. We investigate if it is possible to uniquely recover $\sigma$ from the observation of $K_n$. More precisely, does an algorithm fitting $K_n(\sigma)$ among all the signed measures of $I$ recover the measure $\sigma$?

Remark that a finite number of assigned standard moments does not define a unique signed measure. In fact one can check that for all signed measures and for all integers $m$ there exists a distinct signed measure with the same $m$ first standard moments. It seems there is no hope in recovering discrete measures from a finite number of its generalized moments. Surprisingly, we show that every extrema Jordan type measure $\sigma$ (see Definition 1 and the examples that follows) is the unique solution of a total variation minimizing algorithm, the support pursuit.

Support pursuit. In [CDS98] S. S. Chen, D. L. Donoho and M. A. Saunders introduced the basis pursuit. It is defined by

$$x^* = \arg \min_{y \in \mathbb{R}^p} \|y\|_1 \quad \text{s.t.} \quad Ay = Ax_0,$$

where $A \in \mathbb{R}^{n \times p}$ is the design matrix and $x_0 \in \mathbb{R}^p$ is the target vector. This program is one of the original first steps [CRT06a, Don06] of a remarkable theory, the compressed sensing. As a result, this extremum is appropriated to the reconstruction of the sparse vectors (i.e. vectors with a small support, see [Don06]). In this paper we develop a related program that recovers all the measures with enough structured Jordan support (which can be seen as the sparse-related measures).

Definition of support pursuit. Denote by $\mathcal{M}$ the set of the finite signed measures on $I$ and by $\|\cdot\|_{TV}$ the total variation norm. We recall that, for all $\mu \in \mathcal{M}$,

$$\|\mu\|_{TV} = \sup_{\Pi} \sum_{E \in \Pi} |\mu(E)|,$$
where the supremum is taken over all partition $\Pi$ of $I$ into a finite number of disjoint measurable subsets. Given the observation $K_n(\sigma) = (c_0(\sigma), \ldots, c_n(\sigma))$, the support pursuit is

$$
\sigma^* \in \text{Arg min}_{\mu \in \mathcal{M}} \|\mu\|_{TV} \quad \text{s.t.} \quad K_n(\mu) = K_n(\sigma).
$$

On one hand, basis pursuit minimizes the $\ell_1$-norm subject to linear constraints. On the other hand, support pursuit naturally substitutes the TV-norm (the total variation norm) for the $\ell_1$-norm.

Let us emphasize that support pursuit looks for a minimizer among all the signed measures on $I$. Nevertheless, the target measure $\sigma$ is assumed to be of extrema Jordan type.

**Extrema Jordan type measures.** Let us define more precisely what we understand by the Jordan support family $\mathcal{Y}$.

**Definition 1** (Extrema Jordan type measure) — We say that a signed measure $\mu$ is of extrema Jordan type (with respect to a family $\mathcal{F} = \{u_0, u_1, \ldots, u_0\}$) if and only if its Jordan decomposition $\mu = \mu^+ - \mu^-$ satisfies

$$
\text{Supp}(\mu^+) \subset E^+_P \quad \text{and} \quad \text{Supp}(\mu^-) \subset E^-_P,
$$

where $\text{Supp}(v)$ is defined as the support of the measure $v$, and

- $P$ denotes any linear combination of elements of $\mathcal{F}$,
- $P$ is not constant and $\|P\|_{\infty} \leq 1$,
- $E^+_P$ (resp. $E^-_P$) is the set of all points $x_i$ such that $P(x_i) = 1$ (resp. $P(x_i) = -1$).

In the following, we give some examples of extrema Jordan type measures with respect to the family

$$
\mathcal{F}^n_p = \{1, x, x^2, \ldots, x^n\}.
$$

These measures can be seen as "interesting" target measures for support pursuit given the observation of the $n + 1$ first standard moments.

**Examples with respect to the family $\mathcal{F}^n_p$.** For sake of readability, let $n$ be an even integer, $n = 2m$. We present three important examples.

- **Nonnegative measures:** The nonnegative measures, of which support has size $s$ not greater than $n/2$, are extrema Jordan type measures. Indeed, let $\sigma$ be a nonnegative measure and $S = \{x_1, \ldots, x_s\}$ be its support. Set

$$
P = 1 - c \prod_{i=1}^{s} (x - x_i)^2.
$$

Then, for a sufficiently small value of the parameter $c$, the polynomial $P$ has supremum norm not greater than 1. The existence of such polynomial shows that the measure $\sigma$ is an extrema Jordan type measure.

In Section 2 we extend this notion to any homogeneous $M$-system.

- **Chebyshev measures:** The $k$-th Chebyshev polynomial of the first order is defined by

$$
T_k(x) = \cos(k \arccos(x)), \quad \forall x \in [-1, 1].
$$

It is well known that it has supremum norm not greater than 1, and that
Jordan type measures. As a matter of fact, our results extend to others family
In this paper, we give exact reconstruction results for these three kinds of extrema
Roughly, they can be stated as follows:

\[ E^+_{1,1} = \{ \cos(2l\pi/k), l = 0, \ldots, \left\lfloor \frac{k}{2} \right\rfloor \}, \]
\[ E^-_{1,1} = \{ \cos((2l+1)\pi/k), l = 0, \ldots, \left\lfloor \frac{k}{2} \right\rfloor \}, \]
whenever \( k > 0 \). Then, any measure \( \sigma \) such that
\[ \text{Supp}(\sigma^+) \subseteq E^+_{1,1} \quad \text{and} \quad \text{Supp}(\sigma^-) \subseteq E^-_{1,1}, \]
for some \( 0 < k \leq n \), is an extrema Jordan type measure.
Further examples are presented in Section 3.

\( \Delta \)-spaced out type measures: Let \( \Delta \) be a positive real and \( S_\Delta \) be the set of all pairs \((S^+, S^-)\) of subsets of \([-1, 1]\) such that
\[ \forall x, y \in S^+ \cup S^-, x \neq y, \quad |x - y| \geq \Delta. \]
In Lemma 4.2, we prove that, for all \((S^+, S^-) \in S_\Delta\), there exists a polynomial \( P_{(S^+, S^-)} \) such that
\[ \text{\bullet} \quad P_{(S^+, S^-)} \text{ has degree } n \text{ not greater than a bound depending only on } \Delta, \]
\[ \text{\bullet} \quad P_{(S^+, S^-)} \text{ is equal to } 1 \text{ on the set } S^+, \]
\[ \text{\bullet} \quad P_{(S^+, S^-)} \text{ is equal to } -1 \text{ on the set } S^-, \]
\[ \text{\bullet} \quad \text{and } \|P_{(S^+, S^-)}\|_\infty \leq 1. \]
This shows that any measure \( \sigma \) with Jordan support included in \( S_\Delta \) is an extrema Jordan type measure.

In this paper, we give exact reconstruction results for these three kinds of extrema Jordan type measures. As a matter of fact, our results extend to others family \( \mathcal{F} \). Roughly, they can be stated as follows:

Nonnegative measures: Assume that \( \mathcal{F} \) is a homogeneous M-system (see 2.1.3). Theorem 2.1 shows that any nonnegative measure \( \sigma \) is the unique solution of support pursuit given the observation \( K_n(\sigma) \) where \( n \) is not less than twice the size of the support of \( \sigma \).

Generalized Chebyshev measures: Assume that the \( \mathcal{F} \) is a M-system (see definition 2.1.2). Proposition 3.3 shows the following result: Let \( \sigma \) be a signed measure having Jordan support included in \((E^+_{1,1}, E^-_{1,1})\), for some \( 1 \leq k \leq n \), where \( \xi_k \) denotes the \( k \)-th generalized Chebyshev polynomial (see 3.3.1). Then \( \sigma \) is the unique solution to support pursuit (3) given \( K_n(\sigma) \), i.e. its \( (n+1) \) first generalized moments.

\( \Delta \)-interpolation: Considering the standard family \( \mathcal{F}_p^n = \{ 1, x, x^2, \ldots, x^n \} \),
Proposition 4.3 shows that support pursuit exactly recovers any \( \Delta \)-spaced out type measure \( \sigma \) from the observation \( K_n(\sigma) \) where \( n \) is greater than a bound depending only on \( \Delta \).

These results are closely related to standard results of basis pursuit [Don06]. In fact, further analogies with compressed sensing can be emphasized.

Analogy with compressed sensing. Our estimator follows the aura of the recent breakthroughs [CDS98, CRT06a] in compressed sensing.

In the past decade E. J. Candès, J. Romberg, and T. Tao have shown [CRT06b] that it is possible to exactly recover all sparse vectors from few linear measurements. They considered a matrix \( A \in \mathbb{R}^{n \times p} \) with i.i.d entries (centered Gaussian, Bernoulli, random Fourier sampling) and a \( s \)-sparse vector \( x_0 \) (i.e. vector with support of size at most \( s \)). They pointed out that, with very high probability, the
vector $x_0$ is the only point of contact between the $\ell_1$-ball of radius $\|x_0\|_1$ and the affine space $\{y, Ay = Ax_0\}$. This result holds as soon as $n \geq C s \log(p/s)$, where $C > 0$ is a universal constant. In our framework we uncover the same geometric property:

Let $\sigma$ be an extrema Jordan type measure. Then $\sigma$ is a point of contact between the ball of radius $\|\sigma\|_{TV}$ and the affine space $\{\mu \in \mathcal{M}, K_n(\mu) = K_n(\sigma)\}$, where $n$ is greater than a bound depending only on the structure of the Jordan support of $\sigma$. For instance, in the nonnegative measure case, if $\sigma$ has support of size at most $s$, then $n = 2s$ suffice (see Theorem 2.1).

Actually the reader can check that the above property is equivalent to the fact that the measure $\sigma$ is a solution of support pursuit (more details can be found in 1.2). Accordingly, support pursuit (3) minimizes the total variation in order to pursue support of the target measure. Its name is inherited from basis pursuit of S. S. Chen, D. L. Donoho and M. A. Saunders.

**Organization of the paper.** This paper falls into four parts. The next section introduces the generalized dual polynomials and shows that exact recovery can be understood in terms of an interpolation problem. The second section studies the exact reconstruction of the nonnegative measures, and gives explicit construction of design matrices for basis pursuit. The third section focuses on the generalized Chebyshev polynomials and shows that it is possible to reconstruct signed measures from very few generalized moments. The last section uncover a related property to the nullspace property of compressed sensing.

1. The generalized dual polynomials

In this section we introduce the generalized dual polynomial. In particular we concern with a sufficient condition that guarantees the exact reconstruction of the measure $\sigma$. As a matter of fact, this condition relies on an interpolation problem.

1.1. An interpolation problem. An insight into exact reconstruction is given by Lemma 1.1. Roughly, the existence of a generalized dual polynomial is a sufficient condition to the exact reconstruction of a signed measure with finite support.

As usual, the following result holds for any family $F = \{u_0, u_1, \ldots, u_n\}$ of continuous function on $I$. Throughout this paper, $\text{sgn}(x)$ denotes the sign of the real $x$.

**Lemma 1.1 (The generalized dual polynomials) — Let $n$ be a nonzero integer. Let $S = \{x_1, \ldots, x_s\} \subset I$ be a subset of size $s$ and $(\epsilon_1, \ldots, \epsilon_s) \in \{\pm 1\}^s$. If there exists a linear combination $P = \sum_{k=0}^{n} a_k u_k$ such that

(i) the generalized Vandermonde system

\[
\begin{pmatrix}
  u_0(x_1) & u_0(x_2) & \cdots & u_0(x_s) \\
  u_1(x_1) & u_1(x_2) & \cdots & u_1(x_s) \\
  \vdots & \vdots & \ddots & \vdots \\
  u_n(x_1) & u_n(x_2) & \cdots & u_n(x_s)
\end{pmatrix}
\]

has full column rank.

(ii) $P(x_i) = \epsilon_i$, $\forall i = 1, \ldots, s$,
Remark. From a convex optimization point of view, the set \( P = \sigma = \sum_{i=1}^{s} \sigma_i \delta_{x_i} \) is a \textit{generalized dual polynomial}. This naming is from the original article [CRT06a] of E. J. Candès, T. Tao and J. Romberg, and the \textit{dual certificate} named by E. J. Candès and Y. Plan in [CP10].

1.2. Reconstruction of a cone. Given a subset \( \mathcal{S} = \{x_1, \ldots, x_s\} \) and a sign sequence \( (\varepsilon_1, \ldots, \varepsilon_s) \in \{\pm 1\}^s \), Lemma 1.1 shows that if the generalized interpolation problem defined by (i), (ii) and (iii) has a solution, \textit{then} support pursuit recovers exactly all measures \( \sigma \) with support \( \mathcal{S} \) and such that \( \text{sgn}(\sigma_i) = \varepsilon_i \).

Let us emphasize that the result is slightly stronger. Indeed the proof A.1 remains unchanged if some coefficients \( \sigma_i = 0 \). Consequently support pursuit recovers exactly all the measures \( \sigma \) of which support is \textit{included} in \( \mathcal{S} = \{x_1, \ldots, x_s\} \) and such that \( \text{sgn}(\sigma_i) = \varepsilon_i \) for all nonzero \( \sigma_i \).

Let us denote \( \mathcal{C}(x_1, \varepsilon_1, \ldots, x_s, \varepsilon_s) \) this set. It is exactly the \textit{cone} defined by
\[
\mathcal{C}(x_1, \varepsilon_1, \ldots, x_s, \varepsilon_s) = \left\{ \sum_{i=1}^{s} \mu_i \delta_{x_i} \mid \forall \mu_i \neq 0, \text{sgn}(\mu_i) = \varepsilon_i \right\}.
\]

Thus the existence of \( P \) implies the exact reconstruction of \textit{all} measures in this cone. The cone \( \mathcal{C}(x_1, \varepsilon_1, \ldots, x_s, \varepsilon_s) \) is the conic span of an \( (s-1) \)-dimensional face of the TV-unit ball, that is
\[
\mathcal{F}(x_1, \varepsilon_1, \ldots, x_s, \varepsilon_s) = \left\{ \sum_{i=1}^{s} \varepsilon_i \lambda_i \delta_{x_i} \mid \forall i, \lambda_i \geq 0 \text{ and } \sum_{i=1}^{s} \lambda_i = 1 \right\}.
\]

Furthermore, the affine space \( \{\mu, \mathcal{K}_n(\mu) = \mathcal{K}_n(\sigma)\} \) is tangent to the TV-unit ball at any point \( \sigma \in \mathcal{F}(x_1, \varepsilon_1, \ldots, x_s, \varepsilon_s) \), as shown in the following remark.

Remark. From a convex optimization point of view, the \textit{dual certificates} [CP10] and the generalized dual polynomials are deeply related: the existence of a generalized dual polynomial \( P \) implies that, for all \( \sigma \in \mathcal{F}(x_1, \varepsilon_1, \ldots, x_s, \varepsilon_s) \), a subgradient \( \Phi_P \) of the TV-norm at the point \( \sigma \) is perpendicular to the set of the feasible points, that is
\[
\{\mu, \mathcal{K}_n(\mu) = \mathcal{K}_n(\sigma)\} \subset \ker(\Phi_P),
\]
where ker denotes the nullspace. A proof of this remark can be found in A.2.

1.3. On the condition (i) in Lemma 1.1. Obviously, when \( u_k = x^k \) for \( k = 0, 1, \ldots, n \), the conditions (ii) and (iii) imply that \( n \geq s \) and so condition (i). Nevertheless, this implication is not true for a general set of functions \( \{u_0, u_1, \ldots, u_n\} \). Moreover Lemma 1.1 can fail if condition (i) is not satisfied. For example, set \( n = 0 \) and consider a continuous function \( u_0 \) satisfying the two conditions (ii) and (iii). In this case, if the target \( \sigma \) belongs to \( \mathcal{F}(x_1, \varepsilon_1, \ldots, x_s, \varepsilon_s) \) (where \( x_1, \ldots, x_s \) and \( \varepsilon_1, \ldots, \varepsilon_s \) are given by (ii) and (iii)), then \textit{every} measure \( \mu \in \mathcal{C}(x_1, \varepsilon_1, \ldots, x_s, \varepsilon_s) \) is the \textit{exact solution} of the support pursuit given the observation \( \mathcal{K}_n(\sigma) \).

Proof. See A.1. \qed
\( F(x_1, \varepsilon_1, \ldots, x_s, \varepsilon_s) \) is a solution of support pursuit given the observation \( K_0(\sigma) \). Indeed,
\[
\|\mu\|_{TV} = \int_{-1}^{1} u_0 \, d\mu = K_0(\mu),
\]
for all \( \mu \in F(x_1, \varepsilon_1, \ldots, x_s, \varepsilon_s) \). This example shows that the condition (i) is necessary. Reading the proof A.1, the conditions (ii) and (iii) ensure that the solutions to support pursuit belong to the cone \( C(x_1, \varepsilon_1, \ldots, x_s, \varepsilon_s) \), whereas the condition (i) gives uniqueness.

1.4. The extrema Jordan type measures. Lemma 1.1 shows that Definition 1 is well-founded. As a matter of fact, we have the following corollary.

**Corollary** — Let \( \sigma \) be an extrema Jordan type measure. Then the measure \( \sigma \) is a solution to support pursuit given the observation \( K_0(\sigma) \).

Furthermore, if the Vandermonde system given by (i) in Lemma 1.1 has full column rank (where \( S = \{x_1, \ldots, x_s\} \) denotes the support of \( \sigma \)), then the measure \( \sigma \) is the unique solution to support pursuit given the observation \( K_0(\sigma) \).

This corollary shows that the "extrema Jordan type" notion is appropriate to exact reconstruction using support pursuit.

In the next section we focus on nonnegative measure which are extrema Jordan type measure (as mentioned in the introduction).

2. Exact reconstruction of the nonnegative measures

In this section we show that if the underlying family \( F = \{u_0, u_1, \ldots, u_n\} \) is a homogeneous M-system then the support pursuit recovers exactly all nonnegative discrete measures from the observation of few of its generalized moments. We begin with the definition of the homogeneous M-systems.

2.1. The Markov-systems. The Markov-systems were introduced in approximation theory [KN77, BE95, KS66]. They deal with the problem of finding the best approximation, in terms of the \( \ell_\infty \)-norm, of a given continuous function. We begin with the definition of the Chebyshev-systems (the so-called T-system). They can be seen as a natural extension of the algebraic monomials. Thus a finite combination of elements of a T-system is called a generalized polynomial.

2.1.1. The T-systems of order \( k \). Denote \( \{u_0, u_1, \ldots, u_k\} \) a set of continuous real (or complex) functions on \( \mathbb{T} \). This set is a T-system of degree \( k \) if and only if every generalized polynomial
\[
P = \sum_{l=0}^{k} a_l u_l,
\]
where \((a_0, \ldots, a_k) \neq (0, \ldots, 0)\), has at most \( k \) zeros in \( I \).

This definition is equivalent to each of the two following conditions:
- For all \( x_0, x_1, \ldots, x_k \) distinct elements of \( I \) and all \( y_0, y_1, \ldots, y_k \) real (or complex) numbers, there exists a unique generalized polynomial \( P \) (i.e. \( P \in \text{Span}\{u_0, u_1, \ldots, u_k\} \)) such that \( P(x_i) = y_i \), for all \( i = 0, 1, \ldots, k \).
• For all \( x_0, \ldots, x_k \) distinct elements of \( I \), the \emph{generalized Vandermonde system},
\[
\begin{pmatrix}
u(x_0) & \nu(x_1) & \cdots & \nu(x_k) \\
u_1(x_0) & \nu_1(x_1) & \cdots & \nu_1(x_k) \\
\vdots & \vdots & \ddots & \vdots \\
u_k(x_0) & \nu_k(x_1) & \cdots & \nu_k(x_k)
\end{pmatrix}
\]
has full rank.

2.1.2. \emph{The M-systems}. We say that the family \( \mathcal{F} = \{ \nu_0, \nu_1, \ldots, \nu_n \} \) is a \emph{M-system} if and only if it is a \( T \)-system of degree \( k \) for all \( 0 \leq k \leq n \). Actually the M-systems are common objects (see [KN77]), we mention some examples below.

In this paper, we concern with target measures on \( I = [-1,1] \). Usually, the M-systems are defined on general Hausdorff spaces (see [BEZ94] for instance). For sake of readability, we present examples with different values of \( I \). As usual, if not specified, the set \( I \) is assumed to be equal to \([-1,1]\).

\textbf{Real polynomials}: The family \( \mathcal{F}_p = \{ 1, x, x^2, \ldots \} \) is a M-system. The real polynomials give the standard moments.

\textbf{Müntz polynomials}: Let \( 0 < a_1 < a_2 < \cdots \) be any real numbers. The family \( \mathcal{F}_m = \{ 1, x^{a_1}, x^{a_2}, \ldots \} \) is a M-system on \( I = [0, +\infty) \).

\textbf{Trigonometric functions}: The family \( \mathcal{F}_c = \{ 1, \exp(i\pi x), \exp(i2\pi x), \ldots \} \) is a M-system on \( I = [-1,1] \).

\textbf{Characteristic function}: The family \( \mathcal{F}_c = \{ 1, \exp(i\pi x), \exp(i2\pi x), \ldots \} \) is a M-system on \( I = [0, 1] \).

\textbf{Stieltjes transformation}: The family \( \mathcal{F}_s = \{ 1, \frac{1}{-z_1}, \frac{1}{-z_2}, \ldots \} \), where none of the \( z_i \)'s belongs to \([-1,1]\), is a M-system. The corresponding moments are the \emph{Stieltjes transformation} \( S_{\nu}(z_k) \) of \( \sigma \), namely
\[
c_k(\sigma) = \int_{-1}^{1} \frac{d\sigma(t)}{z_k - t} = S_{\nu}(z_k).
\]

\textbf{Laplace transform}: The family \( \mathcal{F}_l = \{ 1, \exp(-x), \exp(-2x), \ldots \} \) is a M-system. The moments are the \emph{Laplace transform} \( \mathcal{L} \sigma \) at the integer points. It holds
\[
c_k(\sigma) = \int_{-1}^{1} \exp(-kt) \, d\sigma(t) = \mathcal{L} \sigma(k).
\]

A broad variety of common families can be considered in our framework. The above list is not meant to be exhaustive.

Consider the family \( \mathcal{F}_s = \{ \frac{1}{z_0}, \frac{1}{z_1}, \ldots \} \). Remark that no linear combination of its elements gives the constant function 1. Thus the constant function 1 is not a generalized polynomial of this system. To avoid such case, we introduce the \emph{homogeneous M-systems}. 

2.1.3. The homogeneous M-systems. We say that a family \( \mathcal{F} = \{u_0, u_1, \ldots, u_n\} \) is a homogeneous M-system if and only if it is a M-system and \( u_0 \) is a constant function. In this case, all the constant functions \( c, \text{ with } c \in \mathbb{R} \text{ (or } \mathbb{C}) \), are generalized polynomial. Hence the field \( \mathbb{R} \text{ (or } \mathbb{C}) \) is naturally embedded in the generalized polynomial. The adjective homogeneous is named after this comment.

From any M-system we can always construct a homogeneous M-system. Indeed, let \( \mathcal{F} = \{u_0, u_1, \ldots, u_n\} \) be a M-system. In particular the family \( \mathcal{F} \) is a T-system of order 0. Thus the continuous function \( u_0 \) does not vanish in \([-1, 1]\).

As a matter of fact the family \( \{1, u_1 u_0, u_2 u_0, \ldots, u_n u_0\} \) is a homogeneous M-system. All the previous examples of M-systems (see 2.1.2) are homogeneous, even the Stieljes transformation considering:

\[
\tilde{\mathcal{F}}_s = \left\{1, \frac{1}{z_1 - x}, \frac{1}{z_2 - x}, \ldots\right\}.
\]

Using homogeneous M-systems, we show that one can exactly recover all nonnegative measures from few generalized moments.

2.2. An important theorem. The following result is one of the main theorem of this paper. It states that the support pursuit (3) recovers all nonnegative measures \( \sigma \), of which size of support is \( s \), from only \( 2s + 1 \) generalized moments.

**Theorem 2.1** — Let \( \mathcal{F} \) be a homogeneous M-system on \( I \). Consider a nonnegative measure \( \sigma \) with finite support included in \( I \). Then the measure \( \sigma \) is the unique solution to support pursuit given the observation \( K_n(\sigma) \) where \( n \) is not less than twice the size of the support of \( \sigma \).

**Proof.** The complete proof can be found in B.1 but some key points from the theory of approximation are presented in 2.2.1. For further insights about the Markov systems, we recommend the fruitful books [KN77, KS66] to the reader. \( \square \)

In addition, this result is sharp in the following sense. Every measure, of which size of support is \( s \), depends on \( 2s \) parameters (\( s \) for its support and \( s \) for its weights). Surprisingly, this information can be recovered from only \( 2s + 1 \) of its generalized moments. Furthermore the program (3) does not use the fact that the target is nonnegative. It recovers \( \sigma \) among all the signed measures with finite support.

2.2.1. Nonnegative interpolation. An important property of the M-systems is the existence of nonnegative generalized polynomial that vanishes exactly at a prescribed set of points \( \{t_1, \ldots, t_m\} \), where \( t_i \in I \) for all \( i = 1, \ldots, m \). Indeed, define the index as

\[
\text{Index}(t_1, \ldots, t_m) = \sum_{j=1}^{m} \chi(t_j),
\]

where \( \chi(t) = 2 \) if \( t \) belongs to \( I \) (the interior of \( I \)) and 1 otherwise. The next lemma guarantees the existence of nonnegative generalized polynomials.

**Lemma 2.2** (Nonnegative generalized polynomial) — Consider a M-system \( \mathcal{F} \) and points \( t_1, \ldots, t_m \) in \( I \). These points are the only zeros of a nonnegative generalized polynomial of degree at most \( n \) if and only if \( \text{Index}(t_1, \ldots, t_m) \leq n \).
The reader can find a proof of this lemma in [KN77]. Notice that this lemma holds for all \( M \)-systems, however our main theorem needs a homogeneous \( M \)-system.

2.2.2. Is \textit{homogeneous} necessary? If one considers non-homogeneous \( M \)-systems then it is possible to give counterexamples that goes against Theorem 2.1 for all \( n \geq 2s \). Indeed, we have the next result.

**Proposition 2.3** — Let \( \sigma \) be a nonnegative measure supported by \( s \) points. Let \( n \) be an integer such that \( n \geq 2s \). Then there exists a \( M \)-system \( F \) and a measure \( \mu \in \mathcal{M} \) such that \( K_n(\sigma) = K_n(\mu) \) and \( \| \mu \|_{TV} < \| \sigma \|_{TV} \).

**Proof.** See B.2. \( \square \)

Theorem 2.1 gives us the opportunity to build a large family of deterministic matrices for the compressed sensing in the case of nonnegative signals.

2.3. Deterministic matrices for the compressed sensing. The heart of this article beats in the next theorem. It gives deterministic matrices for the compressed sensing. We begin with some state-of-the-art results in compressed sensing. In the following, \( p \) denotes the number of predictors (or, from a signal processing viewpoint, the length of the signal).

**Deterministic Design:** As far as we know, for

\[
 n = \mathcal{O}_{p, s \to \infty} \left( s \log \left( \frac{p}{s} \right) \right),
\]

there exists [BGI+08] a deterministic matrix \( A \in \mathbb{R}^{n \times p} \) such that basis pursuit (2) recovers all \( s \)-sparse vectors from the observation \( Ax_0 \).

**Random Design:** If

\[
 n \geq Cs \log \left( \frac{p}{s} \right),
\]

where \( C > 0 \) is a universal constant. Then there exists (with high probability) a random matrix \( A \in \mathbb{R}^{n \times p} \) such that basis pursuit recovers all \( s \)-sparse vectors from the observation \( Ax_0 \).

The deterministic result holds for large values of \( s, n \) and \( p \). For sake of readability, we do not specify the sense of large here. The reader may find an abundant literature in the respective references (see for example [BGI+08, Don06]).

Considering nonnegative sparse vectors (i.e. vectors with nonnegative entries), we drop the bound on \( n \) to

\[
 n \geq 2s + 1.
\]

Unlike the above examples, our result holds for all values of the parameters (as soon as \( n \geq 2s + 1 \)). In addition we give explicit design matrices for basis pursuit. Last but not least, our bound on \( n \) does not depend on \( p \).

**Theorem 2.4** (Deterministic Design Matrices) — Let \( n, p, s \) be integers such that

\[
 s \leq \min(n/2, p).
\]
Let \( \{1, u_1, \ldots, u_n\} \) be a homogeneous M-system on \( I \). Let \( t_1, \ldots, t_p \) be distinct reals of \( I \). Let \( A \) be the generalized Vandermonde system defined by

\[
A = \begin{pmatrix}
1 & 1 & \ldots & 1 \\
u_1(t_1) & u_1(t_2) & \ldots & u_1(t_p) \\
u_2(t_1) & u_2(t_2) & \ldots & u_2(t_p) \\
\vdots & \vdots & \ddots & \vdots \\
u_n(t_1) & u_n(t_2) & \ldots & u_n(t_p)
\end{pmatrix}.
\]

Then basis pursuit (2) exactly recovers all nonnegative \( s \)-sparse vectors \( x_0 \in \mathbb{R}^p \) from the observation \( Ax_0 \).

**Proof.** See B.3.

Although the predictors could be highly correlated, basis pursuit exactly recovers the target vector \( x_0 \). Of course, this result is theoretical. In actual practice, the sensing matrix \( A \) can be very ill-conditioned. In this case, basis pursuit behaves poorly. Nevertheless, in the forthcoming paper "Compression and Support pursuit", Jean-Marc Azaïs and the two authors of this paper [AdCG11] derive an oracle inequality in error prediction using support pursuit given noisy observations.

**Numerical experiments.** Our numerical experiments illustrate Theorem 2.4. They are of the following form:

(a) Choose constants \( s \) (the sparsity), \( n \) (the number of known moments), and \( p \) (the length of the vector). Choose the family \( \mathcal{F} \) (cosine, polynomial, Laplace, Stieltjes,...).

(b) Select the subset \( S \) (of size \( s \)) uniformly at random.

(c) Randomly generate an \( s \)-sparse vector \( x_0 \) of which support is \( S \), and such that its nonzero entries are distributed according to the chi-square distribution with 1 degree of freedom.

(d) Compute the observation \( Ax_0 \).

(e) Solve (2), and compare to the target vector \( x_0 \).

The program (2) can be recasted in a linear program (see [CDS01] for instance). Then we use an interior point method to solve (2).

The entries of the target signal are distributed according to the chi-square distribution with 1 degree of freedom. We chose this distribution to ensure that the entries are nonnegative. Let us emphasized that the actual values of \( x_0 \) can be arbitrary, only its sign matters. The result remains the same if one take the nonzero entries equal to 1, say.

Let us denote \( K : t \mapsto (1, u_1(t), \ldots, u_n(t)) \). The columns of \( A \) are the values of this map at points \( t_1, \ldots, t_p \). For large \( p \), the vectors \( K(t_i) \) can be very correlated. As a matter of fact, the matrix \( A \) can be ill-conditioned to solve the program (2). To avoid such a case, we chose a family such that the map \( K \) has a large derivative function. It appears that the cosine family gives very good numerical results (see Figure 1).

We investigate the reconstruction error between the numerical result \( \hat{x} \) of the program (2) and the target vector \( x_0 \). Our experiment is of the following form:
Figure 1. Consider the family $\mathcal{F}_{\text{cos}} = \{1, \cos(\pi x), \cos(2\pi x), \ldots\}$ on $I = [0, 1]$ and the points $t_k = k/501$, for $k = 1, \ldots, 500$. The blue circles represent the target vector $x_0$ (a 20-sparse vector), while the black crosses represent the solution $x^*$ of (2) from the observation of 41 cosine moments. In this example $s = 20$, $n = 41$, and $p = 500$. More numerical results can be found in the Appendix D.

(a) Choose $p$ (the length of the vector) and $N$ (the number of numerical experiments).
(b) Let $s$ such that $1 \leq s \leq (p - 1)/2$.
(c) Set $n = 2s + 1$ and solve the program (2). Let $\tilde{x}$ be the numerical result.
(d) Compute the $\ell_1$-error $\|\tilde{x} - x_0\|_1 / p$.
(e) Repeat $N$ times the steps (c) and (d), and compute $\text{Err}_s$ the arithmetic mean of the $\ell_1$-errors.
(f) Return $\|\text{Err}_s\|_\infty$, the maximal value of $\text{Err}_s$ when $1 \leq s \leq (p - 1)/2$.

For $p = 100$ and $N = 10$, we find that

$$\|\text{Err}_s\|_\infty \leq 0.05.$$ 

Remark that all the experiments were done for $n = 2s + 1$. This is the smallest value of $n$ such that Theorem 2.3 holds.

3. Exact reconstruction for generalized Chebyshev measures

In this section we give some examples of extremal polynomials $P$ as they appears in Definition 1. Considering $M$-systems, the corollary of Lemma 1.1 shows that every measure with Jordan support included in $(E^+_p, E^-_p)$ is the only solution to support pursuit. Indeed, the condition (i) of Lemma 1.1 is clearly satisfied when the underlying family $\mathcal{F}$ is a $M$-system.

3.1. Trigonometric families. In the context of $M$-system we can exhibit some very particular dual polynomials. The global extrema of these polynomials gives families of support on which results of Lemma 1.1 hold.
The cosine family. To begin with, consider the \((n + 1)\)-dimensional cosine system
\[
\mathcal{F}^n_{\cos} := \{1, \cos(\pi x), \ldots, \cos(n\pi x)\}
\]
on \(I = [0, 1]\). Obviously, the extremal polynomials
\[
P_k(x) = \cos(k\pi x),
\]
for \(k = 1, \ldots, n\), satisfy \(|P_k|_{\infty} \leq 1\) and \(P_k(1/k) = (-1)^l\), for \(l = 0, 1, \ldots, (k - 1)\). According to Definition 1, let us denote
\[
\begin{align*}
E^+_P &: = \{2l/k \mid l = 0, \ldots, \lfloor \frac{k-1}{2} \rfloor \}, \\
E^-_P &: = \{(2l-1)/k \mid l = 1, \ldots, \lfloor \frac{k}{2} \rfloor \}.
\end{align*}
\]
The corollary that follows Lemma 1.1 asserts the following result.

Consider a signed measure \(\sigma\) having Jordan support \((S^+, S^-)\) such that \(S^+ \subset E^+_P\) and \(S^- \subset E^-_P\), for some \(1 \leq k \leq n\). Then the measure \(\sigma\) can be exactly reconstructed from the observation of
\[
\int_0^1 \cos(k\pi t) d\sigma(t), \quad k = 0, 1, \ldots, n.
\]
Moreover, since the family \(\mathcal{F}^n_{\cos}\) is a M-system, the condition \((\ast)\) in Lemma 1.1 is satisfied. Hence, the measure \(\sigma\) is the only solution of support pursuit given the observations \((6)\).

Using the classical mapping,
\[
\Psi : \begin{cases} 
[0, 1] & \rightarrow [-1, 1] \\
x & \mapsto \cos(\pi x)
\end{cases},
\]
the system of function \((1, \cos(\pi x), \ldots, \cos(n\pi x))\) is pushed forward on the system of function \((1, T_1(x), \ldots, T_n(x))\) where \(T_k(x)\) is the so-called Chebyshev polynomial of the first kind of order \(k\), \(k = 1, \ldots, n\) (see 3.2).

The characteristic function. By the same token, consider the complex value M-system defined by
\[
\mathcal{F}^n_c = \{1, \exp(i\pi x), \ldots, \exp(in\pi x)\}
\]
on \(I = [0, 2]\). In this case, one can check that
\[
P_{\alpha,k}(t) = \cos(k\pi(t - \alpha)), \quad \forall t \in [0, 2),
\]
where \(\alpha \in \mathbb{R}\) and \(0 \leq k \leq n/2\), is a generalized polynomial. Following the previous example, we set
\[
\begin{align*}
E^+_P &: = \{\alpha + 2l/k \mod 2 \mid l = 0, \ldots, \lfloor \frac{k-1}{2} \rfloor \}, \\
E^-_P &: = \{\alpha + (2l - 1)/k \mod 2 \mid l = 1, \ldots, \lfloor \frac{k}{2} \rfloor \}.
\end{align*}
\]
Hence Lemma 1.1 can be applied. It yields that:

Any signed measure having Jordan support included in \((E^+_P, E^-_P)\), for some \(\alpha \in \mathbb{R}\) and \(1 \leq k \leq n/2\), is the unique solution of support pursuit given the observation
\[
\int_0^2 \exp(ik\pi t) d\sigma(t) = q_{\sigma}(k\pi), \quad \forall k = 0, \ldots, n,
\]
where \(q_{\sigma}(k\pi)\) has been defined in the previous section (see 2.1.2).
Notice that the study of basis pursuit with this kind of trigonometric moments have been considered in the pioneered work of D.L. Donoho and P.B. Stark [DS89].

3.2. The Chebyshev polynomials. As mentioned in the introduction, the $k$-th Chebyshev polynomial of the first order is defined by

$$T_k(x) = \cos(k \arccos(x)), \quad \forall x \in [-1, 1].$$

We give some well known properties of the Chebyshev polynomials. The $k$-th Chebyshev polynomial satisfies the equioscillation property on $[-1, 1]$. As a matter of fact, there exists $k+1$ points $\zeta_i = \cos(\pi i / k)$ with $1 = \zeta_0 > \zeta_1 > \cdots > \zeta_k = -1$ such that

$$T_k(\zeta_i) = (-1)^i \| T_k \|_\infty = (-1)^i,$$

where the supremum norm is taken over $[-1, 1]$. Moreover, the Chebyshev polynomial $T_k$ satisfies the following extremal property.

**Theorem 3.1** ([Riv90, BE95]) — We have

$$\min_{p \in \mathcal{P}_{k-1}^C} \| x^k - p(x) \|_\infty = \| 2^{1-k} T_k \|_\infty = 2^{1-k},$$

where $\mathcal{P}_{k-1}^C$ denotes the set of the complex polynomials of degree less than $k-1$, and the supremum norm is taken over $[-1, 1]$. Moreover, the minimum is uniquely attained by $p(x) = x^k - 2^{1-k} T_k(x)$.

These two properties, namely the equioscillation property and the extremal property, will be useful to us when defining the generalized Chebyshev polynomial.

As usual, using Lemma 1.1 we uncover an exact reconstruction result. Consider the family

$$\mathcal{F}_p^m = \{1, x, x^2, \ldots, x^n\}$$
on $I = [-1, 1]$. Set

- $E^+_{T_k} = \{ \cos(2l \pi / k), l = 0, \ldots, \lfloor \frac{k}{2} \rfloor \}$,
- $E^-_{T_k} = \{ \cos((2l+1) \pi / k), l = 0, \ldots, \lfloor \frac{k}{2} \rfloor \}.$

The following result holds:

Consider a signed measure $\sigma$ having Jordan support included in $(E^+_{T_k}, E^-_{T_k})$, for some $1 \leq k \leq n$. Then the measure $\sigma$ is the only solution to support pursuit given its $(n+1)$ first standard moments using support pursuit.

As a matter of fact, this result can be extended to any $M$-systems considering the generalized Chebyshev polynomials.

3.3. The generalized Chebyshev polynomials. Following [BE95], we define the generalized Chebyshev polynomials as follows. Let $\mathcal{F} = \{u_0, u_1, \ldots, u_n\}$ be a $M$-system on $I$. 

3.3.1. Definition. The generalized Chebyshev polynomial

\[ \mathcal{T}_k := \mathcal{T}_k\{u_0, u_1, \ldots, u_n; I\}, \]

where \(1 \leq k \leq n\), is defined by the following three properties:

- \(\mathcal{T}_k\) is a generalized polynomial of degree \(k\), i.e. \(\mathcal{T}_k \in \text{Span}\{u_0, u_1, \ldots, u_k\}\),
- there exists an alternation sequence, \(x_0 < x_1 < \cdots < x_k\), that is,
  \[ \text{sgn}(\mathcal{T}_k(x_{i+1})) = -\text{sgn}(\mathcal{T}_k(x_i)) = \pm \|\mathcal{T}_k\|_\infty, \]
  for \(i = 0, 1, \ldots, k-1\),
- and
  \[ \|\mathcal{T}_k\|_\infty = 1 \quad \text{with} \quad \mathcal{T}_k(\text{max } I) > 0. \]

The existence and the uniqueness of such \(\mathcal{T}_k\) is proved in [BE95]. Moreover, the following theorem shows that the extremal property implies the equioscillation property (7).

**Theorem 3.2** ([Riv90, BE95]) — The \(k\)-th generalized Chebyshev polynomial \(\mathcal{T}_k\) exists and can be written as

\[ \mathcal{T}_k = c \left( u_k - \sum_{i=0}^{k-1} a_i u_i \right), \]

where \(a_0, a_1, \ldots, a_{k-1} \in \mathbb{R}\) are chosen to minimize

\[ \left\| u_k - \sum_{i=0}^{k-1} a_i u_i \right\|_\infty, \]

and the normalization constant \(c \in \mathbb{R}\) can be chosen so that \(\mathcal{T}_k\) satisfies property (8).

The generalized Chebyshev polynomials give a new family of extrema Jordan type measure (see Definition 1). The corresponding target measures are named the Chebyshev measures.

3.3.2. Exact reconstruction of the Chebyshev measures. Considering the equioscillation property (7), set

- \(E^+_k\) as the set of the alternation point \(x_i\) such that \(\text{sgn}(\mathcal{T}_k(x_i)) = \|\mathcal{T}_k\|_\infty\),
- \(E^-_k\) as the set of the alternation point \(x_i\) such that \(\text{sgn}(\mathcal{T}_k(x_i)) = -\|\mathcal{T}_k\|_\infty\).

A direct consequence of the last definition is the following proposition.

**Proposition 3.3** — Let \(\sigma\) be a signed measure having Jordan support included in \((E^+_k, E^-_k)\), for some \(1 \leq k \leq n\). Then \(\sigma\) is the unique solution to support pursuit (3) given \(K_n(\sigma)\), i.e. its \((n + 1)\) first generalized moments.

In the special case \(k = n\), Proposition 3.3 shows that support pursuit recovers all signed measures with Jordan support included in \((E^+_n, E^-_n)\) from \((n + 1)\) first generalized moments. Remark that \(E^+_n \cup E^-_n\) has size \(n\). Hence, this proposition shows that, among all the signed measure on \([-1, 1]\), support pursuit can recover a signed measure of which support has size \(n\) from only \((n + 1)\) generalized moments. As a matter of fact, any measure with Jordan support included in \((E^+_n, E^-_n)\) can be uniquely defined by only \((n + 1)\) generalized moments.
As far as we know, it is difficult to give the corresponding generalized Chebyshev polynomials for a given family \( \mathcal{F} = \{ u_0, u_1, \ldots, u_n \} \). Nevertheless, P. Borwein, T. Erdélyi, and J. Zhang [BEZ94] gives the explicit form of \( \mathcal{F}_k \) for the rational spaces (i.e. the Stieljes transformation in our framework). See also [DS89, HSS96] for some applications in optimal design.

3.3.3. Construction of the Chebyshev polynomials for the Stieljes transformation. We consider the case of the Stieljes transformation described in Section 2. In this case, the Chebyshev polynomials \( T_k \) can be precisely described. Consider the homogeneous \( M \)-system on \([-1, 1]\) defined by
\[
\tilde{\mathcal{F}}_n^m = \left\{ \frac{1}{z_1 - x}, \frac{1}{z_2 - x}, \ldots, \frac{1}{z_n - x} \right\},
\]
where \( (z_i)_{i=1}^k \subset \mathbb{C} \setminus [-1, 1] \).

Reproducing [BE95], we can construct the generalized Chebyshev polynomials of the first kind. As a matter of fact, it yields
\[
\mathcal{F}_k(x) = \frac{1}{2}(f_k(z) + f_k(z)^{-1}), \quad \forall x \in [-1, 1],
\]
where \( z \) is uniquely defined by \( x = \frac{1}{2}(z + z^{-1}) \) and \( |z| < 1 \), and \( f_k \) is a known analytic function in a neighborhood of the closed unit disk. Moreover this analytic function can be only expressed in terms of \( (z_i)_{i=1}^k \). We refer to [BE95] for further details.

4. The nullspace property for the measures

In this section we consider any countable family \( \mathcal{F} = \{ u_0, u_1, \ldots, u_n \} \) of continuous functions on \( I \). In particular we do not assume that \( \mathcal{F} \) is a non-homogeneous \( M \)-system. We aim at deriving a sufficient condition for exact reconstruction of signed measures. More precisely, we concern with giving a related property to the nullspace property [CDD09] of compressed sensing.

Remark that the solutions to the program (3) depend only on the \( (n + 1) \) first elements of \( \mathcal{F} \) and on the target measure \( \sigma \). We investigate the condition that must satisfy the family \( \mathcal{F} \) to ensure exact reconstruction. In the meantime, A. Cohen, W. Dahmen and R. DeVore introduced [CDD09] a relevant condition, the nullspace property. Their property binds the geometry of the nullspace of \( A \) and the best \( k \)-term approximation of the target \( x_0 \) given the observation \( Ax_0 \). This well known property can be stated as follows.

4.1. The nullspace property in compressed sensing. Let \( A \in \mathbb{R}^{n \times p} \) be a matrix. We say that \( A \) satisfies the nullspace property of order \( s \) if and only if for all nonzero vectors \( h \) in the nullspace of \( A \), and all subsets of entries \( S \) of size \( s \), it holds
\[
\|h_S\|_1 < \|h_{S'}\|_1,
\]
where \( h_S \) denotes the vector of which \( i \)-th entry is equal to \( h_i \) if \( i \in S \) and 0 otherwise. It is now standard that the basis pursuit (2) exactly recovers all \( s \)-sparse vectors \( x_0 \) (i.e. vectors with at most \( s \) nonzero entries) if and only if the design matrix \( A \) satisfies the nullspace property of order \( s \).
In this section, we show that the same property holds for the support pursuit. According to the compressed sensing literature, we keep the same name for this related property.

4.2. **The nullspace property for the support pursuit.** Consider the linear map $K_n : \mu \mapsto (c_0(\mu), \ldots, c_n(\mu))$ from $M$ to $\mathbb{R}^{n+1}$. We refer to this map as the **generalized moment morphism**. Its nullspace $\ker(K_n)$ is a linear subspace of $M$. The Lebesgue’s decomposition theorem is the precious tool that carves the nullspace property.

4.2.1. **The S-atomic part.** Let $\mu \in M$ and $S = \{x_1, \ldots, x_s\}$ be a finite subset of $I$. Define $\Delta_S = \sum_{i=1}^s \delta_{x_i}$ as the **Dirac comb** with support $S$. The Lebesgue’s decomposition of $\mu$ with respect to $\Delta_S$ gives

\[
\mu = \mu_S + \mu_{S^c},
\]

where $\mu_S$ is a **discrete** measure of which support is included in $S$, and $\mu_{S^c}$ is a measure of which support is included in $S^c := I \setminus S$.

4.2.2. **The nullspace property with respect to a Jordan support family.** To begin with, as in the standard compressed sensing frame [CDD09], we define the nullspace property with respect to a Jordan support family $\Upsilon$. This property is only a sufficient condition for the exact reconstruction of the finite measure, see Proposition 4.1.

**Definition 1 (Nullspace property with respect to a Jordan support family $\Upsilon$)** — We say that the generalized moment morphism $K_n$ satisfies the nullspace property with respect to a Jordan support family $\Upsilon$ if and only if it satisfies the following property. For all nonzero measures $\mu$ in the nullspace of $K_n$, and for all $(S^+, S^-) \in \Upsilon$,

\[
\|\mu_S\|_{TV} < \|\mu_{S^c}\|_{TV},
\]

where $S = S^+ \cup S^-$. The weak nullspace property states as follows: For all nonzero measures $\mu$ in the nullspace of $K_n$, and for all $(S^+, S^-) \in \Upsilon$,

\[
\|\mu_S\|_{TV} \leq \|\mu_{S^c}\|_{TV},
\]

where $S = S^+ \cup S^-$. Given a nonzero measure $\mu$ in the nullspace of $K_n$, this property means that more than half of the total variation of $\mu$ cannot be concentrated on a small subset. The nullspace property is a key to exact reconstruction as shown in the following proposition.

**Proposition 4.1** — Let $\Upsilon$ be a Jordan support family. Let $\sigma$ be a signed measure having a Jordan support in $\Upsilon$. If the generalized moment morphism $K_n$ satisfies the nullspace property with respect to $\Upsilon$. Then, the measure $\sigma$ is the **unique** solution of support pursuit (3) given the observation $K_n(\sigma)$.

— If the generalized moment morphism $K_n$ satisfies the weak nullspace property with respect to $\Upsilon$. Then, the measure $\sigma$ is the **a** solution of support pursuit (3) given the observation $K_n(\sigma)$.

**Proof.** See C.1. □
As far as we know, it is difficult to check the nullspace property. In the following, we give an example such that the weak nullspace property is satisfied.

4.3. The Δ-spaced out interpolation. We recall that \( S_\Delta \) is the set of all pairs \((S^+, S^-)\) of subsets of \( I = [-1, 1] \) such that

\[
\forall x, y \in S^+ \cup S^-, \ x \neq y, \ |x - y| \geq \Delta.
\]

The next lemma shows that if \( \Delta \) is large enough then there exists a polynomial of degree \( n \), with supremum norm not greater than 1, that interpolates 1 on the set \( S^+ \) and \(-1\) on the set \( S^- \).

Lemma 4.2 — For all \((S^+, S^-)\) \( \in S_\Delta \), there exists a polynomial \( P(S^+, S^-) \) such that

- \( P(S^+, S^-) \) has degree \( n \) not greater than \((2/\sqrt{\pi}) (\sqrt{\pi}/\Delta)^{5/2 + 1/\Delta}\),
- \( P(S^+, S^-) \) is equal to 1 on the set \( S^+ \),
- \( P(S^+, S^-) \) is equal to \(-1\) on the set \( S^- \),
- and \( \|P(S^+, S^-)\|_\infty \leq 1 \) over \( I \).

Proof. See C.2. \( \square \)

This upper bound is meant to show that one can interpolate any sign sequence on \( S_\Delta \). Let us emphasize that this result is far from being sharp. Considering \( L_2 \)-minimizing polynomials under fitting constraint, the authors of the present paper think that one can greatly improve the upper bound of Lemma 4.2. As a matter of fact, our numerical experiments is in complete agreement with this comment. Invoking Lemma 1.1, Lemma 4.2 gives the next proposition.

Proposition 4.3 — Let \( \Delta \) be a positive real. If \( n \geq (2/\sqrt{\pi}) (\sqrt{\pi}/\Delta)^{5/2 + 1/\Delta} \) then \( K_n \) satisfies the weak nullspace property with respect to \( S_\Delta \).

Proof. See C.3. \( \square \)

The bound \((2/\sqrt{\pi}) (\sqrt{\pi}/\Delta)^{5/2 + 1/\Delta}\) can be considerably improved in actual practice. The following numerical experiment shows that this bound can be greatly lowered.

Some simulated experiments. Our numerical experiment consists in looking for a generalized polynomial satisfying the assumption of Lemma 1.1. We work here with the cosine system \((1, \cos(\pi x), \cos(2\pi x), \ldots, \cos(n\pi x))\) for various values of the integer \( n \). As explained in Section 3, we can also push this system on the more classical power system \((1, x, x^2, \ldots, x^n)\). So that, our numerical experiments may be interpreted in this last frame. We consider signed measure having a support \( S \) with \(|S| = 10\). We consider Δ-spaced out type measures for various values of \( \Delta \). For each choice of \( \Delta \), we draw uniformly 100 realizations of signed measures. This means that the points of \( S \) are uniformly drawn on \( I^{10} \), where \( I = [0, 1) \) here, with the restriction that the minimal distance between two points is at least \( \Delta \) and that there exists a couple of points that are exactly \( \Delta \) away from each other. Further, we uniformly randomized the signs of the measure on each point of \( S \). As we wish to work with true signed measures, we do not allow the case where all the signs are the same (negative or positive measures). Once we simulated the set \( S^+ \) and \( S^- \), we wish to build an interpolating polynomial \( P \) of degree \( n \)
Figure 2. Consider the family $F_{\cos} = \{1, \cos(\pi x), \cos(2\pi x), \ldots\}$ on $I = [0,1]$. Set $s = 10$ the size of the target support. We are concerned with signed measures with Jordan support in $S_{\Delta}$ (see (11)). The abscissa represents the values of $1/\Delta$ (with $\Delta = 1/15, 1/20, \ldots, 1/55$), and the ordinate represent the values of $n$ (with $n = 20, 30, \ldots, 100$). For each value of $(\Delta, n)$, we draw uniformly 100 realizations of signed measures and the corresponding $L_2$-minimizing polynomial $P$. The gray scale represents the percentage of times that $\|P\|_\infty \leq 1$ occurs. The white color means 100% (support pursuit exactly recovers all the signed measures) while the black color represent 0% (in all our experiments, the polynomial $P$ is such that $\|P\|_\infty > 1$ over $I$).

In our experiments we consider the values $\Delta = 1/15, 1/20, \ldots, 1/55$. According to Proposition 4.3, the corresponding value of $n$ range from $10^{19}$ to $10^{59}$. In our experiments, we find that $n = 80$ suffices.

Acknowledgements: The two authors of the present paper would like to thank Jean Paul Calvi and Viet Hung Pham for fruitful comments. Y.D.C would like to thank Jean Marc Azais and Franck Barthe for unconditional support.

Appendix A. Proofs of Section 1.

A.1. Proof of Lemma 1.1. Assume that a generalized dual polynomial $P$ exists. Let $\sigma$ be such that $\sigma = \sum_{i=1}^s \sigma_i \delta_{x_i}$ with $\text{sgn}(\sigma_i) = \epsilon_i$. Let $\sigma^*$ be a solution of the support pursuit (3) then $\int P \, d\sigma = \int P \, d\sigma^*$. The equality (ii) yields $\|\sigma\|_{TV} =$...
\[ \int P \, d\sigma. \] Combining the two previous equalities,
\[ \|\sigma\|_{TV} = \int P \, d\sigma = \int P \, d\sigma^* = \sum_{i=1}^{s} \varepsilon_i \sigma_i^* + \int P \, d\sigma_S^*, \]
where \( \varepsilon_i = \text{sgn}(\sigma_i) \) and
\[ \sigma^* = \sum_{i=1}^{s} \sigma_i^* \delta_{x_i} + \sigma_S^*, \]
according to the Lebesgue decomposition (9). Since \( \|P\|_{\infty} = 1 \), it holds
\[ \sum_{i=1}^{s} \varepsilon_i \sigma_i^* + \int P \, d\sigma_S^* \leq \|\sigma_S^*\|_{TV} + \|\sigma_S^*\|_{TV} = \|\sigma^*\|_{TV}. \]
Observe \( \sigma^* \) is a solution of the support pursuit, it follows that \( \|\sigma\|_{TV} = \|\sigma^*\|_{TV} \)
and the above inequality is an equality. It yields \( \int P \, d\sigma_S^* = \|\sigma_S^*\|_{TV}. \) Moreover
we have the following result.

**Lemma A.1** — Let \( \nu \in \mathcal{M} \) with its support included in \( S^c \). If \( \int P \, dv = \|\nu\|_{TV} \)
then \( \nu = 0. \)

**Proof.** Consider the compact set
\[ \Omega_k = I \setminus \bigcup_{i=1}^{s} \left[ x_i - \frac{1}{k}, x_i + \frac{1}{k} \right], \quad \forall k > 0, \]
Suppose that there exists \( k > 0 \) such that \( \|\nu_{\Omega_k}\|_{TV} \neq 0. \) Then the inequality (iii)
leads to \( \int_{\Omega_k} P \, dv < \|\nu_{\Omega_k}\|_{TV}. \) It yields
\[ \|\nu\|_{TV} = \int P \, dv = \int_{\Omega_k} P \, dv + \int_{\Omega_k^c} P \, dv < \|\nu_{\Omega_k}\|_{TV} + \|\nu_{\Omega_k^c}\|_{TV} = \|\nu\|_{TV}, \]
which is a contradiction. We deduce that \( \|\nu_{\Omega_k}\|_{TV} = 0, \) for all \( k > 0. \) The equality
\( \nu = 0 \) follows with \( S^c = \bigcup_{k>0} \Omega_k. \)
This lemma shows that \( \sigma^* \) is a discrete measure with its support included in \( S. \)
In this case, the moment constraint \( \mathcal{K}_n(\sigma^* - \sigma) = 0 \) can be written as a generalized
Vandermonde system,
\[
\begin{pmatrix}
  u_0(x_1) & u_0(x_2) & \cdots & u_0(x_s) \\
  u_1(x_1) & u_1(x_2) & \cdots & u_1(x_s) \\
  \vdots & \vdots & \ddots & \vdots \\
  u_n(x_1) & u_n(x_2) & \cdots & u_n(x_s)
\end{pmatrix}
\begin{pmatrix}
  \sigma_1^* - \sigma_1 \\
  \sigma_2^* - \sigma_2 \\
  \vdots \\
  \sigma_s^* - \sigma_s
\end{pmatrix}
= 0.
\]
From the condition (i), we deduce that the above generalized Vandermonde system is injective. This concludes the proof.

**A.2. Proof of the remark in Section 1.2.** Let \( \sigma \) belong to \( \mathcal{F}(x_1, \varepsilon_1, \ldots, x_s, \varepsilon_s) \).
Consider the linear functional,
\[ \Phi_f : \mu \mapsto \int f \, d\mu, \]
where \( f \) denotes a continuous bounded function. By definition, any subgradient \( \Phi_f \) of the \( TV \)-norm at point \( \sigma \) satisfies that, for all measures \( \mu \in \mathcal{M}, \)
\[ \|\mu\|_{TV} - \|\sigma\|_{TV} \geq \Phi_f(\mu - \sigma). \]
So that, one can easily check that $f$ is equal to 1 (resp. $-1$) on supp$(\sigma^+)$ (resp. supp$(\sigma^-)$) and that $\|f\|_{\infty} = 1$. Conversely, any function $f$ satisfying the latter condition leads to a subgradient $\Phi_f$. Therefore, when it exists, the generalized dual polynomial $P$ is such that $\Phi_P$ is a subgradient of the TV-norm at point $\sigma$. Furthermore, let $\mu$ be a feasible point (i.e. $\mathcal{K}_n(\mu) = \mathcal{K}_n(\sigma)$). Since $P$ a generalized polynomial of order $n$, we deduce that $\Phi_P(\mu - \sigma) = 0$. Hence, the subgradient $\Phi_P$ is perpendicular to the set of the feasible points.

**Appendix B. Proofs of Section 2**

**B.1. Proof of Theorem 2.1.** The proof essentially relies on Lemma 1.1. Let $s$ be an integer. Let $\sigma$ be a nonnegative measure. Let $S = \{x_1, \ldots, x_s\} \subset I$ be its support.

The next lemma shows the existence of a generalized dual polynomial.

**Lemma B.1 (Dual polynomial)** — Let $s$ be an integer and $n$ be such that $n = 2s$. Let $\mathcal{F}$ be an homogeneous $M$-system on $I$. Let $(x_1, \ldots, x_s)$ be such that Index$(x_1, \ldots, x_s) \leq n$. Then there exists a generalized polynomial $P$ of degree $d$ such that

(i) $s \leq d \leq n$,
(ii) $P(x_i) = 1$, $\forall i = 1, \ldots, s$,
(iii) and $|P(x)| < 1$ for all $x \notin \{x_1, \ldots, x_s\}$.

We recall that Index is defined by (5). Notice that these polynomials are presented in the first example of Definition 1.

**Proof of Lemma B.1.** Let $(x_1, \ldots, x_s)$ be such that Index$(x_1, \ldots, x_s) \leq n$. From Lemma 2.2, there exists a nonnegative polynomial $Q$ of degree $d$ that vanishes exactly at the points $x_i$. Moreover, its degree $d$ satisfies (i).

Since $Q$ is continuous on the compact set $I$ then it is bounded. There exists a real $c$ such that $\|Q\|_{\infty} < 1/c$. The generalized polynomial

$$P = 1 - cQ,$$

is the expected generalized polynomial. This concludes the proof.\[\square\]

Observe that

- Using Lemma B.1, it yields that there exists a generalized dual polynomial, of degree at most $n = 2s$, which interpolates the value 1 at points $\{x_1, \ldots, x_s\}$.
- Since $\mathcal{F} = \{u_0, u_1, \ldots, u_n\}$ is a $T$-system, the Vandermonde system given by (i) in Lemma 1.1 has full column rank.

Lemma 1.1 concludes the proof.

**Remark.** Since $\mathcal{F}$ is a homogeneous $M$-system, the constant function 1 is a generalized polynomial. Remark the linear combination $P = 1 - cQ$ is a generalized polynomial because 1 is a generalized polynomial. This assumption is essential (see 2.2.2).

**B.2. Proof of Proposition 2.3.** Let $\sigma = \sum_{i=1}^s \sigma_i \delta_{x_i}$ be a nonnegative measure. Denote $S = \{x_1, \ldots, x_s\}$ its support. Let $n$ be an integer such that $n \geq 2s$. 


Step 1: Let $F_{h} = \{1, u_{1}, u_{2}, \ldots \}$ be an homogeneous $M$-system (the standard polynomials for instance). Let $t_{1}, \ldots, t_{n+1} \in I \setminus \mathcal{S}$ be distinct points. It follows that the Vandermonde system \[
abla \begin{pmatrix}
1 & \ldots & 1 \\
u_{1}(t_{1}) & \ldots & u_{1}(t_{n+1}) \\
u_{2}(t_{1}) & \ldots & u_{2}(t_{n+1}) \\
\vdots & \ddots & \vdots \\
u_{n}(t_{1}) & \ldots & u_{n}(t_{n+1})
\end{pmatrix}\] has full rank. It yields that we may choose $(v_{1}, \ldots, v_{n+1}) \in \mathbb{R}^{n+1}$ such that
\begin{itemize}
\item $v = \sum_{i=1}^{n+1} v_{i} \delta_{t_{i}}$,
\item and for all $k = 0, \ldots, n$, \[\int_{I} u_{k} v = \int_{I} u_{k} d\sigma.\]
\end{itemize}

Step 2: Set \[r = \frac{\|\sigma\|_{TV}}{\|v\|_{TV} + 1}.\] Consider a positive continuous functions $u_{0}$ such that
\begin{itemize}
\item $u_{0}(x_{i}) = r$, for $i = 1, \ldots, s$,
\item $u_{0}(t_{i}) = 1$, for $i = 1, \ldots, n+1$,
\item the function $u_{0}$ is not constant.
\end{itemize}
Set $F = \{u_{0}, u_{0} u_{1}, u_{0} u_{2}, \ldots \}$. Obviously, $F$ is a non-homogeneous $M$-system. As usual, denote $\mathcal{K}_{n}$ the generalized moment morphism of order $n$ derived from the family $F$.

Last step: Set $\mu = r v$. An easy calculation gives $\mathcal{K}_{n}(\sigma) = \mathcal{K}_{n}(\mu)$. Remark that
\begin{equation}
\|\mu\|_{TV} = \sum_{i=1}^{n+1} r |v_{i}| = \frac{\sum_{i=1}^{n+1} |v_{i}|}{\|v\|_{TV} + 1} \|\sigma\|_{TV} < \|\sigma\|_{TV},
\end{equation}
this concludes the proof.

B.3. Proof of Theorem 2.4. Set $T = \{t_{1}, \ldots, t_{p}\}$. Let us denote $\mathcal{M}_{T}$ the set of all finite measure of which support is included in $T$. Let $\Theta_{T}$ be the linear map defined by
\begin{equation}
\Theta_{T} : \begin{cases}
(\mathbb{R}^{p}, \ell_{1}) & \rightarrow (\mathcal{M}_{T}, \|\cdot\|_{TV}) \\
(x_{1}, \ldots, x_{p}) & \mapsto \sum_{i=1}^{p} x_{i} \delta_{t_{i}}
\end{cases}
\end{equation}
One can check that $\Theta_{T}$ is a bijective isometry. Moreover, it holds
\begin{equation}
\forall y \in \mathbb{R}^{p}, \quad \mathcal{K}_{n}(\Theta_{T}(y)) = Ay,
\end{equation}
where $A$ is the generalized vandermonde system defined by
\[
A = \begin{pmatrix}
1 & 1 & \ldots & 1 \\
u_{1}(t_{1}) & u_{1}(t_{2}) & \ldots & u_{1}(t_{p}) \\
u_{2}(t_{1}) & u_{2}(t_{2}) & \ldots & u_{2}(t_{p}) \\
\vdots & \ddots & \vdots & \vdots \\
u_{n}(t_{1}) & u_{n}(t_{2}) & \ldots & u_{n}(t_{p})
\end{pmatrix}.
\]
In the meantime, let $x_{0}$ be a nonnegative $s$-sparse vector. Set $\sigma = \Theta_{T}(x_{0})$ then the size support of $\sigma$ is at most $s$. Consequently, Theorem 2.1 shows that $\sigma$ is the
unique solution to support pursuit. Since \( \sigma \in \mathcal{M}_T \), it yields that \( \sigma \) is the unique solution to the following program
\[
\sigma = \text{Arg min}_{\mu \in \mathcal{M}_T} \| \mu \|_{TV} \quad \text{s.t.} \quad \mathcal{K}_n(\mu) = \mathcal{K}_n(\sigma).
\]
Using (12) and the isometry \( \Theta_T \), it gives that \( x_0 \) is the unique solution to the following program
\[
x_0 = \text{Arg min}_{y \in \mathbb{R}^p} \| y \|_1 \quad \text{s.t.} \quad Ay = Ax_0.
\]
This concludes the proof.

Appendix C. Proofs of Section 4

C.1. Proof of Proposition 4.1. Let \( \mathcal{K}_n \) be a generalized moment morphism that satisfies the nullspace property with respect to a Jordan support family \( Y \). Let \( \sigma \) be a signed measure of which Jordan support belongs to \( Y \). Let \( \sigma^* \) be a solution of the support pursuit (3), it follows that \( \| \sigma^* \|_{TV} \leq \| \sigma \|_{TV} \). Denote \( \mu = \sigma^* - \sigma \) and remark that \( \mu \in \ker(\mathcal{K}_n) \). It holds
\[
\| \sigma^* \|_{TV} = \| \sigma^* S \|_{TV} + \| \sigma^* S^c \|_{TV},
\]
\[
= \| \sigma + \mu S \|_{TV} + \| \mu S^c \|_{TV},
\]
\[
\geq \| \sigma \|_{TV} - \| \mu S \|_{TV} + \| \mu S^c \|_{TV},
\]
where \( S \) denotes the support of \( \sigma \). Suppose that \( \mu \neq 0 \). The nullspace property yields that the measure \( \mu \) satisfies the inequality (10). We deduce \( \| \sigma^* \|_{TV} > \| \sigma \|_{TV} \) which is a contradiction. Thus \( \mu = 0 \) and \( \sigma^* = \sigma \). This concludes the proof.

C.2. Proof of Lemma 4.2. For sake of readability, we present the sketch of the proof here. Let \( (S^+, S^-) \in S_\Delta \). Set \( S = S^+ \cup S^- = \{x_1, \ldots, x_s\} \). Consider the Lagrange interpolation polynomials
\[
l_k(x) = \frac{\prod_{i \neq k}(x - x_i)}{\prod_{i \neq k}(x_k - x_i)},
\]
for \( 1 \leq k \leq s \). One can bound the supremum norm of \( l_k \) over \([0, 1]\) by
\[
\| l_k \|_\infty \leq L(\Delta),
\]
where \( L(\Delta) \) is an upper bound that depends only on \( \Delta \). Consider the \( m \)-th Chebyshev polynomial of the first order \( T_m(x) = \cos(m \arccos(x)) \), for all \( x \in [-1, 1] \). For a sufficient large value of \( m \), there exists \( 2s \) extrema \( \xi_i \) of \( T_m \) such that \( |\xi_i| \leq 1/(sL(\Delta)) \). Interpolating values \( \xi_i \) at point \( x_k \), we build the expected polynomial \( P \). We find that the polynomial \( P \) has degree not greater than
\[
C \left( \sqrt{\varepsilon}/\Delta \right)^{5/2+1/\Delta},
\]
where \( C = 2/\sqrt{\pi} \).
C.3. **Proof of Proposition 4.3.** Let $\mu$ be a nonzero measure in the nullspace of $\mathcal{K}_n$ and $(A, B)$ be in $S_\Delta$. Let $S$ be equal to $A \cup B$. Set $S^+$ (resp. $S^-$) the set of points $x$ in $S$ such that the $\mu$-weight at point $x$ is nonnegative (resp. negative). Observe that $S = S^+ \cup S^-$ and $(S^+, S^-) \in S_\Delta$. From Lemma 4.2, there exists $P_{(S^+, S^-)}$ of degree not greater than $n$ such that $P_{(S^+, S^-)}$ is equal to 1 on $S^+$, $-1$ on $S^-$, and $\|P_{(S^+, S^-)}\|_\infty \leq 1$. It yields

$$\int P_{(S^+, S^-)} \, d\mu = \|\mu_S\|_{TV} + \int_{S^c} P_{(S^+, S^-)} \, d\mu \geq \|\mu_S\|_{TV} - \|\mu_{S^c}\|_{TV}.$$ 

Since $\mu \in \ker(\mathcal{K}_n)$, it follows that $\int P_{(S^+, S^-)} \, d\mu = 0$. This concludes the proof.
Figure 3. These numerical experiments illustrate Theorem 2.4. We consider the family $\mathcal{F}_{\cos} = \{1, \cos(\pi x), \cos(2\pi x), \ldots\}$ and the points $t_k = k/(p + 1)$, for $k = 1, \ldots, p$. The blue circles represent the target vector $x_0$, while the black crosses represent the solution $x^*$ of (2). The respective values are $s = 10, n = 21, p = 500$; $s = 50, n = 101, p = 500$; and $s = 150, n = 301, p = 500$. 

Appendix D. Numerical Experiments
References


Institut de Mathématiques de Toulouse (CNRS UMR 5219). Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse, France.

E-mail address: yohann.decastro@math.univ-toulouse.fr
E-mail address: fabrice.gamboa@math.univ-toulouse.fr