RANDOM PROJECTIONS, MARGINALS, AND MOMENTS

DRAFT EXPOSITORY NOTES BY DJALIL CHAFAÏ

Abstract. Just linear problems with positivity constraints in infinite dimensions...

Contents

1. Introduction

These notes deal with the characterization or the reconstruction of probability distributions from their lower dimensional projections.

We denote by $\langle x, y\rangle$ the Euclidean scalar product of $\mathbb{R}^d$, by $\|x\| = \sqrt{\langle x, x\rangle}$ the associated norm, by $\mathbb{S}(\mathbb{R}^d) = \{x \in \mathbb{R}^d; \|x\| = 1\}$ the centered unit radius sphere, and by λ the uniform distribution on this sphere. For any probability distribution P on $\mathbb{R}^d$ and any vector $x \in \mathbb{R}^d$, we denote by $P_{\langle x\rangle}$ the law of the projection in the direction x. In other words, if $X \sim P$ then $\langle X, x\rangle \sim P_{\langle x\rangle}$. For any couple P, Q of probability distributions on $\mathbb{R}^d$, we define

$$\mathcal{E}(P, Q) = \{x \in \mathbb{R}^d; P_{\langle x\rangle} = Q_{\langle x\rangle}\} \quad \text{and} \quad \mathcal{E}_1(P, Q) = \{x \in \mathbb{S}(\mathbb{R}^d); P_{\langle x\rangle} = Q_{\langle x\rangle}\}.$$

Notice the identity $\mathcal{E}_1(P, Q) = \mathcal{E}(P, Q) \cap \mathbb{S}(\mathbb{R}^d)$. Clearly, $\mathcal{E}(P, Q)$ is a closed cone fully determined by the compact set $\mathcal{E}_1(P, Q)$ since

$$\mathcal{E}(P, Q) = \{\alpha x\,; (\alpha, x) \in \mathbb{R}_+ \times \mathcal{E}_1(P, Q)\}.$$

We denote by $\varphi_P : \mathbb{R}^d \to \mathbb{C}$ the Fourier transform (i.e. characteristic function) of P defined for any $t \in \mathbb{R}^d$ by

$$\varphi_P(t) = \int_{\mathbb{R}^d} e^{i\langle t,x\rangle}\, dP(x).$$

Remark 1.1 (Defining the set of all rays). *Since $P_{\langle -x\rangle}$ is fully and uniquely determined by $P_{\langle x\rangle}$, the set $\mathcal{E}_1(P, Q)$ is in turn fully determined by the quotient $\mathcal{E}_1(P, Q)/\mathcal{R}$ where $\mathcal{R}$ denotes the antipodal binary equivalence relation defined by $y\,\mathcal{R}\,x$ if and only if $y = -x$. Actually, one can even replace the Euclidean centered unit radius sphere $\mathbb{S}(\mathbb{R}^d)$ by any centered sphere of positive radius for any norm on $\mathbb{R}^d$, and one can take then a suitable antipodal quotient.*

Theorem 1.2 (Cramér-Wold [CW36]). *For any couple P, Q of probability distributions on $\mathbb{R}^d$, we have $P = Q$ if and only if $\mathcal{E}_1(P, Q) = \mathbb{S}(\mathbb{R}^d)$.*

Date: First version April 2007, revised December 2014, revised November 2021, compiled November 19, 2021.

Proof. Well, the desired result is a reformulation of the fact that a probability distribution is uniquely determined by its characteristic function (i.e. its Fourier transform). This classical result is in turn a consequence of the Fourier inversion formula or of the monotone class theorem. □

More generally, for a fixed d and for any $1 \leqslant k \leqslant d$, let $\mathcal{E}_k$ be the set of all k-dimensional subspaces of $\mathbb{R}^d$. This set can be identified to the collection of projections $\pi_V : \mathbb{R}^d \to V$ where V is a k-dimensional subspace of $\mathbb{R}^d$. The set $\mathcal{E}_k$ can also be seen as a collection of k-dimensional closed subsets of $\mathbb{S}(\mathbb{R}^d)$, and for instance $\mathcal{E}_2$ is the collection of great circles of $\mathbb{S}(\mathbb{R}^2)$.

Next, we define the set $\mathcal{E}_k(P,Q) = \{\pi_V \in \mathcal{E}_k; \pi_V(P) = \pi_V(Q)\}$, where $\pi_V(P)$ denotes the projection of P onto V, i.e. the image distribution of P by the map π_V. Notice that if $\pi_V \in \mathcal{E}_k(P,Q)$, then $\pi_{V'} \in \mathcal{E}_{k'}(P,Q)$ for any k'-dimensional sub-vector space V' of V. Clearly, $\mathcal{E}_d(P,Q) = \emptyset$ if $P \neq Q$ and $\mathcal{E}_d(P,Q) = \mathbb{S}(\mathbb{R}^d)$ if $P = Q$. By the Cramér-Wold theorem, $\mathcal{E}_k(P,Q) = \mathcal{E}_k$ if and only if $P = Q$.

2. Some few moments with the problem of moments

Let P be a probability distribution on $\mathbb{R}$. The *sequence of absolute moments* (M_n) of P is given for every $n \in \mathbb{N}$ by

$$M_n = \int_{\mathbb{R}} |x|^n \, dP(x) \in [0,\infty].$$

When $M_n < \infty$, the associated moment m_n is given by

$$m_n = \int_{\mathbb{R}} x^n \, dP(x),$$

and we have $|m_n| \leqslant M_n$. The *sequence of moments* (m_n) of P is well defined if and only if P has finite absolute moments, in other words if and only if $\mathbb{R}[X] \subset \mathrm{L}^1(P)$. Notice that $(M_n) = (m_n)$ when P is supported in $\mathbb{R}_+$. We say that a probability distribution P on $\mathbb{R}$ (resp. $\mathbb{R}_+$) is *characterized by its moments* if and only if P is the unique probability distribution on $\mathbb{R}$ (resp. $\mathbb{R}_+$) with sequence of moments (m_n). Moments problems go back probably to Tchebychev, Markov, and Stieltjes, and can be subdivided into several subproblems including…

(1) **existence.** under which condition a sequence of real numbers (m_n) is the sequence of moments of a probability distribution?
(2) **uniqueness.** under which condition a probability distribution is characterized by its moments?
(3) **structure.** how to describe the convex set of all probability distributions sharing the same sequence of moments?

Notice that existence and uniqueness problems are in general sensitive to additional constraints. For the moments problem, uniqueness on $\mathbb{R}_+$ does not imply uniqueness on $\mathbb{R}$, whereas existence on $\mathbb{R}$ does not imply existence on $\mathbb{R}_+$. Stieltjes studied moments problems on $\mathbb{R}_+$. He obtained a necessary and sufficient condition for existence, and studied uniqueness. His methods involve for instance continuous fractions, see for example [Sti93]. Later, Hamburger continued the work of Stieltjes and studied moments problems on $\mathbb{R}$, see for instance [Ham21]. Actually, moments problems were studied by many people including among others Marcel Riesz, Kreĭn, Hausdorff, Hamburger, and Carleman. Nowadays, there is less activity around moments problems, but [Pak01] is a counter example. As for many problems, existence is “simpler” than uniqueness.

2.1. Existence.
The moments problem is a linear problem, with a positivity constraint.

Theorem 2.1 (Hamburger [Ham21])**.** *A sequence (m_n) of real numbers is the sequence of moments of a probability distribution on the real line if and only if the infinite Hankel matrix*

$$H = \begin{pmatrix} m_0 & m_1 & m_2 & \cdots \\ m_1 & m_2 & m_3 & \cdots \\ m_2 & m_3 & m_4 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

is positive definite: $\sum_{n,n'} m_{n+n'} u_n \overline{u}_{n'} \geqslant 0$ for any sequence (u_n) of complex numbers such that $u_n = 0$ except for a finitely many values of n.

Notice that $H_{n,n'} = m_{n+n'}$ and that the condition on H means that every finite square submatrix of H is positive definite. There exists a necessary and sufficient condition it terms of matrices for the Stieltjes moment problem. The reader will find more details on Stietljes and Hamburger moments problems in [ST43], [Akh65], and [KN77].

2.2. **Uniqueness.** One can derive sufficient conditions for uniqueness in terms of Hankel determinants. We give in the sequel some aspects of a more general approach based on quasi-analytic functions. In the sequel, and unless explicitly mentioned, we consider uniqueness on $\mathbb{R}$, not on $\mathbb{R}_+$.

Theorem 2.2 (Hausdorff, [Hau23]). *A probability distribution P on $[0,1]$ is characterized by its moments.*

Proof. The density of the polynomials for the uniform topology (Weierstass-Bernstein theorem) implies that P is characterized by its moments among compactly supported probability measures. It remains to remove this comapctness condition. See also [DF04]. □

Remark 2.3 (A counter-example by C.C. Heyde). *The log-normal distribution is not characterized by its moments. Namely, a real random variable X follows the log-normal distribution if and only if* $\log(X)$ *is a standard Gaussian distribution. The associated Lebesgue density is* $x \mapsto (2\pi)^{-1/2}x^{-1}\exp(-(\log(x))^2/2)\mathrm{I}_{\mathbb{R}_+}(x)$. *Now, for any fixed real number $a \in [-1,+1]$, let us consider the Lebesgue density f_a defined by $f_a(x) = f(x)(1 + a\sin(2\pi\log(x)))$ for every $x \in \mathbb{R}$. It turns out that f and f_a share the same sequence of moments, see for instance* [Fel71, p. 227].

Theorem 2.4 (Tchakaloff-Bayer-Teichmann [Tch57, BT06]). *Let P be a probability distribution on $\mathbb{R}^d$ with some finite absolute moments: $\max_{1\leqslant k\leqslant n} M_n < \infty$ for some integer $n \geqslant 1$. Let $\mathbb{R}_n[X]$ be the vector space of polynomial functions of total degree less than of equal to n. Then there exists a finitely supported probability distribution $P_k = p_1\delta_{x_1} + \cdots + p_k\delta_{x_k}$ with $\mathrm{supp}(P_k) = \{x_1,\ldots,x_k\} \subset \mathrm{supp}(P)$ and $k \leqslant \mathrm{Dim}\mathbb{R}_n[X]$ such that for every $f \in \mathbb{R}_n[X]$,*

$$\int_{\mathbb{R}^d} f(x)\,dP(x) = \int_{\mathbb{R}^d} f(x)\,dP_k(x) = \sum_{i=1}^{k} p_i f(x_i).$$

Proof. The proof uses basic separation properties of convex sets, the extremal points theorem of Minkowski-Carathéodory, and the transformation of the measure into a measure on the space of polynomials of bounded degree. This kind of result is also referred as *quadrature* of *cubature* formulas. See also [Put97] and [CF02]. There a link between Hamburger moments problems and the density of polynomials in Lebesgue spaces, see for instance [Sto00], [Bak01, Bak03], [Ber96], [FP05], and [PV99]. □

Theorem 2.5 (Analycity of the Fourier transform and the moments problem). *Let P be a probability distribution on $\mathbb{R}$ with well defined moments (m_n) and Fourier transform φ_P. The following propositions are equivalent.*

1. *φ_P is analytic on a neighborhood of the origin;*
2. *φ_P is analytic on $\mathbb{R}$;*
3. $\overline{\lim}_n \left(\frac{1}{n!}|m_n|\right)^{\frac{1}{n}} < \infty$.

Moreover, if they hold true, then P is characterized by its moments (m_n). It is the case in particular when P is compactly supported or when

$$\overline{\lim_n}\,\frac{1}{n}|m_n|^{\frac{1}{n}} < \infty.$$

.

Proof. For every n, we have $M_n < \infty$, and thus φ_P is n times differentiable on $\mathbb{R}$. Moreover $\varphi_P^{(n)}$ is continuous on $\mathbb{R}$ and for every $t \in \mathbb{R}$,

$$\varphi_P^{(n)}(t) = \int_{\mathbb{R}} (ix)^n e^{itx}\,dP(x).$$

In particular, $\varphi_P^{(n)}(0) = i^n m_n$, and the Taylor series of φ_P at the origin is determined by the sequence (m_n). Recall that the radius of convergence r of the power series $\sum_n a_n z^n$ associated to the sequence of complex numbers (a_n) is given by the Hadamard formula

$$r^{-1} = \overline{\lim_n}\,|a_n|^{\frac{1}{n}},$$

and consequently, 1⇔ 3 (just take $a_n = i^n m_n/n!$). In the other hand, for any $n \in \mathbb{N}$ and any $s,t \in \mathbb{R}$, we have

$$e^{isx}\left(e^{itx} - 1 - \frac{itx}{1!} - \cdots - \frac{(itx)^{n-1}}{(n-1)!}\right) \leqslant \frac{|tx|^n}{n!},$$

see for instance [Fel71, p. 512 and 514], and thus, for any $n \in \mathbb{N}$ and any $s,t \in \mathbb{R}$,

$$\left(\varphi_P(s+t)-\varphi_P(s)-\frac{t}{1!}\varphi_P'(s)-\cdots-\frac{t^{n-1}}{(n-1)!}\varphi_P^{(n-1)}(s)\right)\leqslant m_n\frac{|t|^n}{n!},$$

which implies 1 ⇔ 2. By the Stirling formula, if $\overline{\lim}_n \frac{1}{n}|m_n|^{\frac{1}{n}} < \infty$ then condition 3 holds true. If P is compactly supported, then $\sup_n |m_n| < \infty$ and thus condition 3 holds true. Suppose now that conditions 1-3 hold true. From condition 2, the analytic continuation principle states that φ_P admits a maximal simply connected analytic continuation to a neighborhood of $\mathbb{R}$ in $\mathbb{C}$, which is thus holomorphic. Next, the sequence of moments (m_n) uniquely characterizes the Taylor series at the origin, and thus uniquely characterizes the analytic continuation of φ_P by virtue of the isolated zeros theorem. In particular, the sequence of moments characterizes the function φ_P on $\mathbb{R}$, and thus P by virtue of the the Cramér-Wold theorem. □

It turns out that a probability distribution on $\mathbb{R}$ can be characterized by its moments without having an analytic Fourier transform[1]. Actually, in view of the moments problem, the main useful property here regarding analycity is that an analytic function on $\mathbb{R}$ is uniquely determined by its value and the values of all its derivatives at the origin. Quasi-analytic functions have this property. These functions where introduced by Émile Borel and Hadamard, and where later brilliantly studied by Denjoy and Carleman, see for instance the memoir of Carleman [Car26], or [Rud87, ch. 19] and [BMR97, sec. 4.2]. The Carleman condition appearing below is strictly weaker than the Hadamard condition.

For any sequence of positive real numbers (c_n) and any bounded interval $[a,b] \subset \mathbb{R}$, we denote by $\mathcal{C}([a,b],(c_n))$ the class of infinitely differentiable functions $f : [a,b] \subset \mathbb{R} \to \mathbb{C}$ such that $\sup_{[a,b]} \left|f^{(n)}\right| \leqslant r^n c_n$ for any $n \in \mathbb{N}$ and for some positive real constant r which may depend on f. The Hadamard problem consists in finding conditions on (c_n) such that any couple of functions f and g in $\mathcal{C}([a,b],(c_n))$ that are equal together with all their derivatives at some fixed point of $[a,b]$ are equal on the whole interval $[a,b]$. Such functions are called *quasi-analytic*. The analytic functions on $[a,b]$ correspond to the class $\mathcal{C}([a,b],(n!))$.

Theorem 2.6 (Denjoy-Carleman characterization of quasi-analycity)**.** *For any sequence of positive real numbers (c_n) and any bounded interval $[a,b] \subset \mathbb{R}$, the class $\mathcal{C}([a,b],(c_n))$ is quasi-analytic if and only if*

$$\sum_{n=1}^{\infty}\left(\inf_{k\geqslant n}|c_k|^{\frac{1}{k}}\right)^{-1}=\infty.$$

Analycity implies quasi-analycity but the converse if false. The Carleman condition is satisfied if the Hadamard condition is satisfied.

Corollary 2.7 (Carleman condition for the moments problem)**.** *Let P be a probability distribution on $\mathbb{R}$ with finite absolute moments (M_n) and moments (m_n). If at least one of the following conditions is satisfied*

(1) $\sum_{n=1}^{\infty} {M_{2n}}^{-\frac{1}{2n}} = \infty$;
(2) $\sum_{n=1}^{\infty} {M_n}^{-\frac{1}{n}} = \infty$;
(3) $\sum_{n=1}^{\infty} |m_n|^{-\frac{1}{n}} = \infty$

then P is characterized by its moments.

Proof. We have $|m_n| \leqslant M_n$ for every $n \in \mathbb{N}$. In the other hand, the elementary bound $2|u|^{2n+1} \leqslant |u|^{2n} + |u|^{2n+2}$ valid for every $u \in \mathbb{R}$ and $n \in \mathbb{N}$ implies that $2M_{2n}^{2n+1} \leqslant M_{2n} + M_{2n+2}$ for every $n \in \mathbb{N}$. Consequently, we obtain the following cascading Carleman like conditions:

$$\sum_{n=1}^{\infty}{M_{2n}}^{-\frac{1}{2n}}=\infty \quad\Rightarrow\quad \sum_{n=1}^{\infty}{M_n}^{-\frac{1}{n}}=\infty \quad\Rightarrow\quad \sum_{n=1}^{\infty}|m_n|^{-\frac{1}{n}}=\infty.$$

They imply that φ_P is quasi-analytic, and that it is characterized by the sequence (m_n) since $\varphi_P^{(n)}(t) = (it)^n m_n$ for every $n \in \mathbb{N}$ and $t \in \mathbb{R}$. The desired result follows then from the Denjoy-Carleman theorem with $c_n = |m_n|$ by using the bound $\inf_{k\geqslant n} c_k \leqslant c_n$. □

If φ_P is analytic on a neighborhood of the origin then the Hadamard condition $\overline{\lim}_n n^{-1}|m_n|^{\frac{1}{n}} < \infty$ holds true and implies the Carleman condition $\sum_{n=1}^{\infty} |m_n|^{-\frac{1}{n}} = \infty$.

[1] We want examples here!

Let us move now to the multivariate moment problem. We define the sequence of absolute moments (M_n) of a probability distribution P on $\mathbb{R}^d$ with $d > 1$ by

$$M_n = \int_{\mathbb{R}^d} \|x\|^n \, dP(x) \in [0, \infty]$$

for any $n \in \mathbb{N}$. If $P_{\|\cdot\|}$ denotes the image distribution of P by the map $x \mapsto \|x\|$, then $P_{\|\cdot\|}$ is a probability distribution on $\mathbb{R}_+$ with sequence of moments (M_n). By using Hölder inequality, if $M_n < \infty$ for every $n \in \mathbb{N}$, then for any multi-index $k \in \mathbb{N}^d$, one can define the moment m_k of P by

$$m_k = m_{k_1,\ldots,k_d} = \int_{\mathbb{R}^d} x_1^{k_1} \cdots x_d^{k_d} \, dP(x).$$

We say that P is *characterized by its moments* if and only if P is the unique probability distribution on $\mathbb{R}^d$ with moments (m_k). Notice that when $M_n < \infty$ for every $n \in \mathbb{N}$, the Fourier transform φ_P is infinitely Fréchet differentiable on $\mathbb{R}^d$ and for every $t \in \mathbb{R}^d$ and $k \in \mathbb{N}^d$,

$$\partial_{t_1}^{k_1} \cdots \partial_{t_d}^{k_d} \varphi_P(t_1, \ldots, t_d) = i^{k_1+\cdots+k_d} t_1^{k_1} \cdots t_d^{k_d} m_k.$$

It is delicate to make use of analycity in dimension strictly bigger than 1 due to the lack of multidimensional isolated zeros theorem (e.g. Hartog type phenomena). In some sense, the Carleman moments condition turns out to be more flexible.

Theorem 2.8 (Multidimensional case)**.** *Let P be a probability distribution on $\mathbb{R}^d$ with Fourier transform φ_P and finite absolute moments (M_n) and moments (m_k). If at least one the following propositions hold true*

(1) *$P_{\langle x \rangle}$ is characterized by its moments for every $x \in \mathbb{S}(\mathbb{R}^d)$;*
(2) *P satisfies to the Carleman condition*

$$\sum_{n=1}^{\infty} (M_n)^{-\frac{1}{n}} = \infty;$$

(3) *for every $x \in \mathbb{S}(\mathbb{R}^d)$, the function $t \in \mathbb{R} \mapsto \varphi_P(tx)$ is analytic on $\mathbb{R}$;*

then P is characterized by its moments (m_k).

Proof. For every $x \in \mathbb{S}(\mathbb{R}^d)$, the moments of the unidimensional probability distribution $P_{\langle x \rangle}$ are uniquely determined by the sequence (m_k) since for every $n \in \mathbb{N}$,

$$\int_{\mathbb{R}} u^n \, dP_{\langle x \rangle}(u) = \int_{\mathbb{R}} \langle x, y \rangle^n \, dP(y) = \sum_{k_1+\cdots+k_d=n} \binom{n}{k_1 \cdots k_d} x_1^{k_1} \cdots x_d^{k_d} m_{k_1,\ldots,k_d}.$$

By the Cramér-Wold theorem, $1 \Rightarrow P$ is characterized by (m_k).

2⇒1. For every $x \in \mathbb{S}(\mathbb{R}^d)$, the unidimensional probability distribution $P_{\langle x \rangle}$ satisfies in turn to the Carleman condition since for every $n \in \mathbb{N}$,

$$\int_{\mathbb{R}} |u|^n \, dP_{\langle x \rangle}(u) = \int_{\mathbb{R}^d} |\langle x, y \rangle|^n \, dP(y) \leqslant \|x\|^n \int_{\mathbb{R}^d} \|y\|^n \, dP(y) = M_n,$$

3⇒1. We have For every $x \in \mathbb{S}(\mathbb{R}^d)$, $\varphi_{P_{\langle x \rangle}}(t) = \varphi_P(tx)$ for every $t \in \mathbb{R}$. □

The following results shows that the Carleman condition is sharp, for the moment characterization problem and for the projection characterization problem.

Theorem 2.9 (Bélisle-Massé-Ransford [BMR97])**.** *Let K be a closed ball of $\mathbb{R}^d$ not containing the origin and let (M_n) be a sequence of positive real numbers such that*

$$M_0 = 1, \quad M_n^2 \leqslant M_{n-1} M_{n+1} \textit{ for every } n > 0, \quad \textit{and} \quad \sum_{n=1}^{\infty} M_n^{-\frac{1}{n}} < \infty.$$

Then there exists two probability measures P and Q on $\mathbb{R}^d$ such that

(1) *P and Q are mutually singular;*
(2) *P and Q share the same projection on every $(d-1)$-dimensional subspace of $\mathbb{R}^d$ not containing K;*
(3) *$\max(M_{2n}(P), M_{2n}(Q)) \leqslant M_n^2$ for every $n \geqslant 0$;*
(4) *$\int_{\mathbb{R}^d} x_1^{k_1} \cdots x_d^{k_d} \, dP(x) = \int_{\mathbb{R}^d} x_1^{k_1} \cdots x_d^{k_d} \, dQ(x)$ for every multi-index $k \in \mathbb{N}^d$.*

The proof is given in [BMR97, sec. 5.3]. A slight adaptation of this proof leads to the following result mentioned in [CAFR07, th. 2.6].

Theorem 2.10 (Cuesta-Albertos-Fraiman-Ransford [CAFR07])**.** *Let K be a proper closed subset of $\mathbb{S}(\mathbb{R}^d)$ and let (M_n) be a sequence of positive real numbers s.t.*

$$M_0 = 1, \quad M_n^2 \leqslant M_{n-1}M_{n+1} \text{ for every } n > 0, \quad \text{and} \quad \sum_{n=1}^{\infty} M_n^{-\frac{1}{n}} < \infty.$$

Then there exists two probability measures P and Q on $\mathbb{R}^d$ such that

(1) *P and Q are mutually singular;*
(2) *$K \subset \mathcal{E}_1(P,Q)$;*
(3) *$\max(M_n(P), M_n(Q)) \leqslant M_n^2$ for every $n \geqslant 0$.*

3. The Rényi-Gilbert theorem and random projections

According to the Cramér-Wold theorem in $\mathbb{R}$, for any couple P,Q of probability distributions on $\mathbb{R}$, we have $P = Q$ if and only if $\mathcal{E}_1(P,Q)$ is not empty. The aim of the sequel is to further study such kind of characterization, in higher dimensions.

Theorem 3.1 (Rényi-Gilbert [Rén52, Gil55])**.** *Let P and Q be two probability measures on $\mathbb{R}^2$. Assume that at least one of the following properties hold true.*

(1) *Rényi: the restrictions of the Fourier transforms φ_P and φ_Q of P and Q on every centered circle are analytic (as functions of the angle);*
(2) *Gilbert: P and Q have finite absolute moments and both sequences of absolute moments satisfy to the Carleman condition.*

Then $P = Q$ if and only if $\mathcal{E}_1(P,Q)$ is infinite. In particular, $P \neq Q$ if and only if $\mathcal{E}_1(P,Q)$ is finite.

Proof. If $P = Q$ then $\mathcal{E}_1(P,Q) = \mathbb{S}(\mathbb{R}^d)$ which is an infinite set. Conversely, we assume from now that $\mathcal{E}_1(P,Q)$ is infinite, and our aim is to show that $P = Q$.

Let consider the Rényi part. For any $r > 0$, we denote by $\mathcal{S}(r)$ the centered sphere of $\mathbb{R}^2$ with radius r. For each $r > 0$, the Fourier transforms φ_P and φ_Q coincide on an infinite set C_r of points of $\mathcal{S}(r)$. Since $\mathcal{S}(r)$ is compact, the infinite set C_r admits an accumulation point, and thus the two analytic functions φ_P and φ_Q coincide on the whole $\mathcal{S}(r)$. This holds for every $r > 0$, and thus $\varphi_P = \varphi_Q$. The result follows then from the Cramér-Wold theorem.

Let us consider the Gilbert part. According to the assumptions on P and Q, it is sufficient to show that the values at the origin of all the derivatives of φ_P may be computed from $\{\varphi_{P_{\langle x\rangle}}; x \in \mathcal{E}_1(P,Q)\}$. Now fix n and a finite sequence of distinct points $\{x_1,\ldots,x_n\} \subset \mathcal{E}_1(P,Q)$. We have by writing $x_k = e^{i\theta_k}$ for every $1 \leqslant k \leqslant n$:

$$\varphi^{(n)}_{P_{\langle x_k\rangle}}(0) = \sum_{j=0}^{n} \binom{n}{j} \cos^{n-j}(\theta_k)\sin^j(\theta_k)\varphi_P^{(n-j,j)}(0,0).$$

This gives $n+1$ linear equations involving the variables $\{\varphi_P^{(n-j,j)}(0,0); 0 \leqslant j \leqslant n\}$. Let us denote by A the matrix associated to this linear system. We may assume that $0 \leqslant \theta_k < \pi$ for every $1 \leqslant k \leqslant n$ without loss of generality (just replace x_k by $-x_k$ if not). Suppose for the moment that $\theta_k \neq \pi/2$ for every $1 \leqslant i \leqslant n$ and set $a_k = \tan(\theta_k)$. The determinant of the linear system of equations is a Vandermonde determinant, which is non-zero if and only if $\tan(\theta_k) \neq \tan(\theta_{k'})$ for every $1 \leqslant k \neq k' \leqslant n$. Since $0 \leqslant \theta_k < \pi$ for every $1 \leqslant k \leqslant n$, this condition writes $\theta_k \neq \theta_{k'}$ for every $1 \leqslant k \neq k' \leqslant n$, which is true since the points $x_1,\ldots,x_n$ are distinct. Now, if $\theta_k = \pi/2$, then the initial system matrix A has a whole row $A_{k,\cdot}$ of zeros except for the entry $A_{k,1} = 1$, and we may apply the argument to the cofactor of $A_{k,1}$ in A. □

It seems that the Rényi-Gilbert theorem was rediscovered independently by Ferguson, according to the abstract [Fer59] of an unpublished article.

Theorem 3.2 (Heppes [Hep56])**.** *Let P be a probability distribution on $\mathbb{R}^2$. If P has a positive Lebesgue density on a disk. Then for every finite subset F of $\mathcal{S}(\mathbb{R}^d)$, there exists a probability distribution $Q \neq P$ on $\mathbb{R}^2$ such that $F \subset \mathcal{E}_1(P,Q)$.*

3.1. **The case of finitely supported discrete probability distributions.**

It is natural to ask about the problem of determination by projections for discrete probability distributions. The following theorem gives an answer.

Theorem 3.3 (Rényi-Heppes [Rén52, Hep56]). *Let P be a discrete probability distribution on $\mathbb{R}^d$ with a support made with exactly k distinct atoms. Assume that $V_1, \ldots, V_{k+1}$ are subspaces of $\mathbb{R}^d$ of respective dimensions $d_1, \ldots, d_{k+1}$ such that no couple of them is contained in a hyperplane (i.e. no straight line is perpendicular to more than one of them). Then, for any probability distribution Q in $\mathbb{R}^d$, we have $P = Q$ if and only if $\pi_{V_i} \in \mathcal{E}_{d_i}(P, Q)$ for every $1 \leqslant i \leqslant k+1$.*

In particular, for a probability distribution made with k atoms in $\mathbb{R}^d$, we see that at most $k+1$ hyperplanes are enough to characterize the distribution. This fact can be seen as a counter example to the Rényi-Gilbert theorem where infinitely many directions are required. In some sense, since a probability distribution on $\mathbb{R}^d$ is the limit of a sequence of finitely supported discrete probability distributions, the two theorems are intuitively compatible.

The following result, taken from [BMR97, sec. 6], provides more information.

Theorem 3.4 (Bélisle-Massé-Ransford [BMR97]). *Let P and Q be two probability distributions on $\mathbb{R}^d$. Then*

$$\sup_{x \in \mathbb{R}^d} |P(\{x\}) - Q(\{x\})| \leqslant \frac{1}{\operatorname{card}(\mathcal{E}_{d-1}(P, Q))}.$$

In particular, if $\mathcal{E}_{d-1}(P, Q)$ is infinite then P and Q have the same discrete part.

Proof. Assume that $V_1, \ldots, V_k$ are distinct elements of $\mathcal{E}_{d-1}(P, Q)$. Fix $x \in \mathbb{R}^d$ and put $c = P(\{x\}) - Q(\{x\})$ and $A_i = \pi_{V_i}^{-1}(\{\pi_{V_i}(x)\}) \setminus \{x\}$ for every $1 \leqslant i \leqslant k$. Then we have for every $1 \leqslant i \leqslant k$,

$$\begin{aligned} P(A_i) &= P(\pi_{V_i}^{-1}(\{\pi_{V_i}(x)\})) - P(\{x\}) \\ &= Q(\pi_{V_i}^{-1}(\{\pi_{V_i}(x)\})) - Q(\{x\}) - c \\ &= Q(A_i) - c. \end{aligned}$$

Since $V_1, \ldots, V_k$ are distinct subspaces of codimension 1, the sets $A_1, \ldots, A_k$ are disjoint[2], and we get

$$1 \geqslant \sum_{i=1}^k Q(A_i) \geqslant \sum_{i=1}^k c = kc.$$

The desired inequality follows then by exchanging the role of P and Q in the above. □

Rényi gave in [Rén52] an example on $\mathbb{R}^2$ for which $|P(\{x\}) - Q(\{x\})| = \frac{1}{k}$. Namely, consider a polygon $\mathcal{P}$ with $2k$ sides, centered at the origin. Number the $2k$ vertices according to the trigonometric way. Let P (resp. Q) be the probability distribution with mass $1/k$ at each odd (resp. even) vertex of $\mathcal{P}$. Then P and Q share the same projections on the k straight lines orthogonal to pairs of opposite sides and going through the origin.

Remark 3.5 (Yet another generalization of Cramér-Wold in the plane). *Let E be a Borel subset of $H = \mathbb{R}_+ \times \mathbb{R}$ of positive Lebesgue measure. It was shown by Sitaram in* [Sit83] *by using elementary complex analysis that if two probability distributions P and Q on $\mathbb{R}^2$ satisfy to $P(T(E)) = Q(T(E))$ for every translation-rotation T then $P = Q$. This result constitute another generalization of the Cramér-Wold theorem in $\mathbb{R}^2$, and is actually probably valid on $\mathbb{R}^d$. According to Sitaram, it was obtained earlier by Hertle in* [Her79] *by using a Radon transform.*

3.2. **Gallery of counter examples.**

We gather here several examples and counter examples that show the complexity and some counter intuitive aspects of the characterization by projections.

Remark 3.6 (Rényi ≠ Gilbert). *The probability distribution P on $\mathbb{R}^2$ with density $re^{i\theta} \mapsto \pi^{-1}(1+r^2)^{-2}$ has a constant density and thus a constant Fourier transform on each circle, and therefore satisfies to the Rényi condition. However, P has no moments, and thus does not satisfies to the Gilbert condition. We ignore if conversely, there exists a probability distribution on $\mathbb{R}^2$ which satisfies the Gilbert condition without satisfying the Rényi condition.*

Remark 3.7 (When infinitely many directions are not enough). *Let $\varphi_1 : \mathbb{R} \to \mathbb{R}$ be defined by $\varphi_1(t) = (1-|t|)\mathrm{I}_{[-1,+1]}(t)$ for every $t \in \mathbb{R}$, and let $\varphi_2 : \mathbb{R} \to \mathbb{R}$ be the periodic extension on $\mathbb{R}$ with period 2 of the restriction of φ_1 on $[-1, +1]$. It turns out that both φ_1 and φ_2 are the Fourier transforms of univariate probability distributions. Notice that φ_1 and φ_2 coincide on the whole interval $[-1, +1]$! Furthermore, the functions $\varphi_1 \otimes \varphi_1$ and $\varphi_1 \otimes \varphi_2$ are the Fourier transforms of two probability distributions on $\mathbb{R}^2$. These Fourier transforms coincide in the directions of angle $\pi/4 \leqslant \theta \leqslant 3\pi/4$. Moreover, the associated*

[2] Indeed if $y \in A_i \cap A_j$ with $i \neq j$, then $y \neq x$ and $y \in (\pi_{V_i}(x) + V_i^{\perp}) \cap (\pi_{V_j}(x) + V_j^{\perp})$. But since V_i and V_j have codimension 1, the subspaces $\pi_{V_i} + V_i^{\perp}$ and $\pi_{V_j}(x) + V_j^{\perp}$ are two distinct straigth lines, intersecting only at x, impossible!

distributions do not have any finite moment. This example is mentioned in [Gil55]*. It is attributed to Khintchine and appears in* [Lév37] *and in* [Fel71, p. 505-506]*.*

Remark 3.8 (Finitely many directions are not enough for Gaussians)**.** *A Gaussian distribution on* $\mathbb{R}^2$ *satisfies to both Rényi and Gilbert conditions, and is thus characterized by the prescription of its projections on an infinite number of directions. It is tempting to ask if a finite number of directions is enough to characterize a Gaussian distribution of* $\mathbb{R}^2$*. Let us consider a random vector* (X,Y) *of* $\mathbb{R}^2$ *with distribution* P *with Lebesgue density* f *given for every* $(x,y)\in\mathbb{R}^2$ *by*

$$f(x,y)=\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}\Big(1+xy(x^2+y^2)e^{-\frac{1}{2}(x^2+y^2+2u)}\Big),$$

where u *is some arbitrary real number* u *such that*

$$\left|xy(x^2-y^2)e^{-\frac{1}{2}(x^2+y^2+2u)}\right|\leqslant 1.$$

This condition ensures that f *is a density. The Fourier transform of* P *is given for any* $(t_1,t_2)\in\mathbb{R}^2$ *by*

$$\varphi_P(t_1,t_2)=e^{-\frac{1}{2}(t_1^2+t_2^2)}+\frac{1}{32}t_1t_2(t_1^2-t_2^2)e^{-u-\frac{1}{4}(t_1^2+t_2^2)}.$$

It is immediate to check that $X\sim\mathcal{N}(0,1)$*,* $Y\sim\mathcal{N}(0,1)$*,* $X+Y\sim\mathcal{N}(0,2)$*,* $X-Y\sim\mathcal{N}(0,2)$*, despite the fact that* (X,Y) *is not Gaussian. One can additionally check that* X *and* Y *are uncorrelated. More generally, for an arbitrary finite sequence of directions* $(a_1,b_1),\ldots,(a_n,b_n)$ *in the plane* $\mathbb{R}^2$*, it can be shown that for suitable real numbers* u *and* v*, the function* $\varphi:\mathbb{R}^2\to\mathbb{R}$ *defined for any* (t_1,t_2) *by*

$$\varphi(t_1,t_2)=e^{-\frac{1}{2}(t_1^2+t_2^2)}+e^{-u-\frac{v}{2}(t_1^2+t_2^2)}\prod_{k=1}^n(b_k^2t_1^2-a_k^2t_2^2)$$

is the Fourier transform of a probability distribution P *on* $\mathbb{R}^2$*. One can immediately check that* $P_{\langle(a_k,b_k)\rangle}=\mathcal{N}(0,a_k^2+b_k^2)$ *for every* $1\leqslant k\leqslant n$*, despite the fact that* P *is not Gaussian. These families of counter examples where given in* [HT75]*. Later, Hamedani provided in* [Ham84] *the Lebesgue density* $f:\mathbb{R}^d\to\mathbb{R}$ *defined for any* $x\in\mathbb{R}^d$ *by*

$$f(x_1,\ldots,x_d)=(2\pi)^{-\frac{1}{2}d}\left(e^{-\frac{1}{2}\|x\|^2}+w(x_1^2-x_2^2)\prod_{k=1}^d x_k\mathrm{I}_{[-1,+1]}(x_k)\right)$$

where w *is an arbitrary real number such that*

$$\left|w(x_1^2-x_2^2)\prod_{k=1}^d x_k\mathrm{I}_{[-1,+1]}(x_k)e^{\frac{1}{2}\|x\|^2}\right|\leqslant 1.$$

Now it can be shown that if $(X_1,\ldots,X_d)\sim P$ *where* P *has density* f*, then* $X_1,\ldots,X_d$ *are Gaussians. Moreover, the variables* $X_1,\ldots,X_r$ *are independent for any* $1\leqslant r<d$*. In particular, any linear combination of* $X_1+\cdots+X_r$ *is Gaussian. Furthermore, if* $U=X_1\pm X_2$ *and* V *is any linear combination of* $X_3,\ldots,X_d$*, then* $U+V$ *is Gaussian. However,* $(X_1,\ldots,X_d)$ *is not Gaussian!*

Remark 3.9 (A counter example by Ferguson)**.** *This counter example is presented in* [BMR97, ex. 5.1]*, see also* [Fer59]*. Let* P *be a probability distribution on a random vector of the form* (U,U) *where* U *follows the Cauchy distribution of Lebesgue density* $u\mapsto\pi^{-1}(1+u^2)^{-1}$*. Let* Q *be a tensor product of two Cauchy distributions (the two components are thus independent). Then, for any* $(t_1,t_2)\in\mathbb{R}^2$*, we have* $\varphi_P(t_1,t_2)=\exp(-|t_1+t_2|)$ *and* $\varphi_Q(t_1,t_2)=\exp(-|t_1|-|t_2|)$*. It follows that* $\varphi_P(t_1,t_2)=\varphi_Q(t_1,t_2)$ *if* t_1 *and* t_2 *have the same sign, and therefore* $\mathcal{E}_1(P,Q)$ *contains the whole first and the third quadrant directions, but* $P\neq Q$*. Or course neither the Rényi condition nor the Gilbert condition are satisfied by* P *or* Q*. This counter example have strong similarities with the counter example of Gilbert given in remark 3.6 and the counter example of Lévy-Khintchine-Gilbert given in remark 3.7.*

Remark 3.10 (Periodic perturbations on the square)**.** *Let* $H:\mathbb{R}\to\mathbb{R}$ *be a* C^1 *even and periodic function of period* 1*. Notice that* H *is thus symmetric with respect to the axis* $x=1/2$*. If* $h=H'$*, we have* $\int_0^1 h(x-y)\,dx=H(1-y)-H(-y)=0$ *for every* $y\in[0,1]$ *and* $\int_0^1 h(x-y)\,dy=H(x-0)-H(x-1)=0$ *for every* $x\in[0,1]$*. Moreover,* $\int_{(x,z-x)\in[0,1]^2}h(2x-z)\,dz=H(z)-H(-z)=0$ *for every* $z\in[0,2]$*. Now, let* P *be some probability distribution on* $[0,1]^2$ *with Lebesgue density* f*. If* $\varepsilon=\int_{[0,1]^2}f>0$ *then* $(x,y)\mapsto f(x,y)+uh(x-y)$ *is the Lebesgue density of a probability distribution* Q *on* $[0,1]^2$*, provided that the real number* u *satisfies* $u\inf_{[0,1]}h\geqslant-\varepsilon$*. We have* $P\neq Q$ *if* $uh\not\equiv 0$*, but* $\{e_1,e_2,\frac{1}{2}(e_1+e_2)\}\subset\mathcal{E}_1(P,Q)$ *where* $\{e_1,e_2\}$ *is the canonical basis of* $\mathbb{R}^2$*. More generally, the method of construction of this counter* *example may be used in finite discrete settings. Recall that the set of probability distributions on the finite discrete square* $\{1,\ldots,2n+1\}^2$ *can be identified with the set of bistochastic matrices of size* $(2n+1)\times(2n+1)$*. For any real number* $\alpha>0$*, let* M_α *be the* $(2n+1)\times(2n+1)$ *antisymmetric matrix with entries in* $\{-\alpha,+\alpha\}$ *and constant diagonals with alternating sign. Then for any bistochastic matrix* P *of size* $(2n+1)\times(2n+1)$ *such that* $\varepsilon=\min_{1\leqslant i,j\leqslant 2n+1}P_{i,j}>0$*, and for any* $0<\alpha<\varepsilon$*, the* $(2n+1)\times(2n+1)$ *matrix* $P+M_\alpha$ *is bistochastic. Moreover, if* $(X,Y)\sim P$ *and* $(X',Y')\sim P+M_\alpha$ *then* $\mathcal{L}(X+Y)=\mathcal{L}(X'+Y')$*. We ignore if one can do something useful here with the Birkhoff-von Neumann theorem that says that every bistochastic matrix is a convex combination of permutation matrices. Notice that knowing the probability distributions of* X*,* Y*, and* $X+Y$ *implies the knownledge of the covariance of* X *and* Y*. There is thus a unique Gaussian on* $\mathbb{R}^2$ *with prescribed compatible Gaussian projections on the two axis and on the main diagonal (the same holds true for any proper diagonal actually).*

3.3. **Random projections.** The Rényi-Gilbert theorem is false in dimension $d>2$ since a 2-dimensional subvector space may contain an infinite number of unidimensional directions! This makes us able to add by various ways some mass in an orthogonal direction to construct several distributions which agree for an infinite number of unidimensional projections. One can get rid of this phenomenon by considering i.i.d. random directions in $\mathbb{S}(\mathbb{R}^d)$. The argument works as soon as the common distribution of the random directions in $\mathbb{S}(\mathbb{R}^d)$ has full support. We start with our own simple result, which does not need any assumption regarding moments or analycity.

Theorem 3.11 (No assumptions but random projections)**.** *Let* P *and* Q *be two probability distributions on* $\mathbb{R}^d$*. Then we have* $P=Q$ *if and only if* $\lambda(\mathcal{E}_1(P,Q))=1$*. In particular, if* (X_n) *is a sequence of i.i.d. random variables on* $\mathbb{S}(\mathbb{R}^d)$ *with common distribution* λ*, then* $P=Q$ *if and only if* $\mathbb{P}(\mathrm{Card}\{n;X_n\in\mathcal{E}_1(P,Q)\}=\infty)=1$ *and* $P\neq Q$ *if and only if* $\mathbb{P}(\mathrm{Card}\{n;X_n\in\mathcal{E}_1(P,Q)\}=\infty)=0$*.*

Proof. If $P=Q$ then $\mathcal{E}_1(P,Q)=\mathbb{S}(\mathbb{R}^d)$ and thus $\lambda(\mathcal{E}_1(P,Q))=1$. Conversely, assume that $\lambda(\mathcal{E}_1(P,Q))=1$ and suppose that $P\neq Q$. Since $P\neq Q$, by virtue of the Cramér-Wold theorem, one can then pick $x\in\mathbb{S}(\mathbb{R}^d)\setminus\mathcal{E}_1(P,Q)$. Now the set $\mathcal{E}_1(P,Q)$ is closed and thus there exists an open neighborhood $\mathcal{V}_x$ of x in $\mathbb{S}(\mathbb{R}^d)$ such that $\mathcal{V}_x\subset\mathbb{S}(\mathbb{R}^d)\setminus\mathcal{E}_1(P,Q)$. But since λ is the uniform distribution on $\mathbb{S}(\mathbb{R}^d)$, we have $\lambda(\mathcal{V}_x)>0$, which contradicts the hypothesis $\lambda(\mathcal{E}_1(P,Q))=1$. We therefore conclude that $P=Q$ if and only if $\lambda(\mathcal{E}_1(P,Q))=1$. For the last statement of the theorem, we notice that for any measurable subset A of $\mathbb{S}(\mathbb{R}^d)$, we have $\mathbb{P}(\mathrm{Card}\{n;X_n\in A\}=\infty)\in\{0,1\}$, and this probability is equal to 1 if and only if $\lambda(A)=1$. It is a trivial case of the Borel-Cantelli Lemma or of the zero-one law of Kolmogorov. Just take now $A=\mathcal{E}_1(P,Q)$ to conclude the proof. □

Under some additional hypotheses on P regarding moments or analycity, our following theorem reduces the random projections theorem above from infinite random directions to one random direction.

Theorem 3.12 (Characterization by a unique random projection)**.** *Let* P *and* Q *be two probability distribution on* $\mathbb{R}^d$ *with* $d\geqslant 2$*. Assume that at least one of the following properties holds true.*

1. *Rényi like: the restriction of the Fourier transform* φ_P *and* φ_Q *of* P *and* Q *on every centered circle*[3] *of* $\mathbb{R}^d$ *is analytic (as functions of the angle);*
2. *Gilbert like: the absolute moments of* P *and* Q *are finite and both sequences of absolute moments satisfy to the Carleman condition (implies that* P *and* Q *are characterized by their moments).*

Then $\lambda(\mathcal{E}_1(P,Q))>0$ *implies* $P=Q$*. More precisely, if* $\mathcal{E}_1(P,Q)$ *is not contained in a projective hypersurface of* $\mathbb{S}(\mathbb{R}^d)$ *then* $\mathcal{E}_1(P,Q)=1$ *and thus* $P=Q$*.*

Proof. Let $P\neq Q$ be two different probability distributions on $\mathbb{R}^d$. Since $P\neq Q$, we have $\mathcal{E}_1(P,Q)<1$ and there exists $y\in\mathbb{S}(\mathbb{R}^d)\setminus\mathcal{E}_1(P,Q)$. If $\mathcal{E}_1(P,Q)=\emptyset$, then $\lambda(\mathcal{E}_1(P,Q))=0$. If $\mathcal{E}_1(P,Q)\neq\emptyset$, then pick $x\in\mathcal{E}_1(P,Q)$. Now, let $V_{x,y}=\mathrm{Vect}\{x,y\}$ be the 2-dimensional subvector space of $\mathbb{R}^d$ containing both x and y. The projection of P (respectively Q) onto any element v of $V_{x,y}$ is equal to the projection onto v of the projection of P (respectively Q) on $V_{x,y}$. By the Rényi-Gilbert theorem applied to the projections of P and Q on $V_{x,y}$, we get that $\mathcal{E}_1(P,Q)\cap V_{x,y}$ is finite. Since this hold for any $x\in\mathcal{E}_1(P,Q)$, we get that $\mathcal{E}_1(P,Q)$ is a projective hypersurface of $\mathcal{S}(\mathbb{R}^d)$, and also that $\lambda(\mathcal{E}_1(P,Q))=0$. □

The proof of theorem 3.12 above is quite geometric since it reduces directly the proof to the Rényi-Gilbert theorem by considering 2-dimensional slices. Notice that P and Q play a symmetric role in theorem 3.12, and it particular, for the Gilbert case, the Carleman condition is required for both P and Q. Recently, Cuesta-Albertos, Fraiman, and Ransford showed in [CAFR07] that one can relax the

[3] A "centered circle" of $\mathbb{R}^d$ is the intersection of a 2-dimensional subspace of $\mathbb{R}^d$ with a centered sphere of $\mathbb{R}^d$. In other words, a centered circle of $\mathbb{R}^d$ is a great circle of a centered sphere.

Gilbert type assumption on P or Q (say Q). They call their result "Sharp Cramér-Wold theorem". Notice that one can extract the argument regarding Q from the proof of theorem 3.13 below and simply plug it inside the proof of theorem 3.12 above. However, the rest of the two proofs will remain completely different.

Theorem 3.13 (Cuesta-Albertos-Fraiman-Ransford [CAFR07])**.** *Let P be a probability distribution on $\mathbb{R}^d$ with $d \geqslant 2$ with finite absolute moments satisfying to the Carleman condition (implies that P is characterized by its moments). If Q is a probability measure on $\mathbb{R}^d$ such that $\lambda(\mathcal{E}_1(P,Q)) > 0$ then $P = Q$. More precisely, if $\mathcal{E}_1(P,Q)$ is not contained in a projective hypersurface of $\mathbb{R}^d$ then $\mathcal{E}_1(P,Q) = 1$ and thus $P = Q$.*

Proof. We begin by showing that Q has finite absolute moments. Define the set V of $\mathbb{R}^d$ by

$$V = \left\{ x \in \mathbb{R}^d; \int_{\mathbb{R}^d} |\langle x, y\rangle|^n \, dQ(y) < \infty \right\}.$$

By convexity, V is a subspace of $\mathbb{R}^d$. Moreover, if $x \in \mathcal{E}(P,Q)$, then

$$\int_{\mathbb{R}^d} |\langle x, y\rangle|^n \, dQ(y) = \int_{\mathbb{R}^d} |u|^n \, dQ_{\langle x\rangle}(y) = \int_{\mathbb{R}^d} |u|^n \, dP_{\langle x\rangle}(y) \int_{\mathbb{R}^d} |\langle x, y\rangle|^n \, dP(y) < \infty,$$

and thus $\mathcal{E}(P,Q) \subset V$. Assume from now that $\mathcal{E}(P,Q)$ is not included is a projective hypersurface of $\mathbb{R}^d$. Then necessarily $V = \mathbb{R}^d$ otherwise we can pick a non-zero $z \in V^\perp$ and for such a vector $\mathcal{E}(P,Q)$ is included in the set of zeros of the polynomial $x \mapsto \langle x, z\rangle$. Since $V = \mathbb{R}^d$, and if $e_1, \ldots, e_d$ denotes the canonical basis of $\mathbb{R}^d$, we have for any $n \in \mathbb{N}$

$$\int_{\mathbb{R}^d} \|y\|^n \, dQ(y) = \int_{\mathbb{R}^d} \left(\sum_{i=1}^d |\langle e_i, y\rangle|^2 \right)^{\frac{n}{2}} dQ(y) \leqslant c_{d,n} \sum_{i=1}^d \int_{\mathbb{R}^d} |\langle e_i, y\rangle|^n \, dQ(y) < \infty.$$

Therefore, Q has all its absolute moments finite as announced. Now, for every $n \in \mathbb{N}$, define the homogeneous polynomial $p_n : x \in \mathbb{R}^d \mapsto p(x) \in \mathbb{R}$ by

$$p_n(x) = \int_{\mathbb{R}^d} \langle x, y\rangle^n \, dP(y) - \int_{\mathbb{R}^d} \langle x, y\rangle^n \, dQ(y).$$

Notice that $p_n(x)$ is the difference between the moment of order n or $P_{\langle x\rangle}$ and the moment of order n of $Q_{\langle x\rangle}$. A simple computation shows that p_n vanishes on $\mathcal{E}(P,Q)$. Since $\mathcal{E}(P,Q)$ is not included is a projective hypersurface of $\mathbb{R}^d$ we get that $p_n \equiv 0$ for every $n \in \mathbb{N}$. This implies that $P_{\langle x\rangle}$ and $Q_{\langle x\rangle}$ share the same sequence of moments for every $x \in \mathbb{R}^d$. Now since P satisfies to the Carleman condition, then the sequence of absolute moments of $P_{\langle x\rangle}$ satisfies also to the Carleman condition for every $x \in \mathbb{R}^d$. Consequently, $P_{\langle x\rangle} = Q_{\langle x\rangle}$ for every $x \in \mathbb{R}^d$, and thus $P = Q$ by virtue of the Cramér-Wold theorem. □

Remark 3.14 (Rényi type condition)**.** *By using remark 3.6, one can easily cook up an example for which the Rényi type assumption of theorem 3.12 is satisfied whereas the Gilbert type assumption of theorem 3.12 and the assumption of theorem 3.13 are not. More generally, if R is a probability distribution on $\mathbb{R}$ which is not characterized by its moments, then the probability distribution P on $\mathbb{R}^d$ defined for every bounded measurable function $h : \mathbb{R}^d \to \mathbb{R}$ by*

$$\int_{\mathbb{R}^d} h(z) \, dP(z) = \int_{\mathbb{S}(\mathbb{R}^d)} \left(\int_{\mathbb{R}} h(rx) \, dR(r) \right) d\lambda(x)$$

is not characterized by its moments in $\mathbb{R}^d$. However, its Fourier transform is constant on each centered circle of $\mathbb{R}^d$ since P is invariant by rotation. Thus, P satisfies to the Rényi type assumptions but does not satisfies to the Gilbert type assumptions. This kind of examples shows the gap between characterization by projections and characterization by moments in multivariate settings. It is interesting to reinterpret these theorems and examples in terms of Radon transform.

As mentioned in [CAFR07], theorem 3.13 has very interesting statistical applications, since it allows the construction of distribution-free multivariate statistical tests, for instance of Kolmogorov-Smirnov type. The general idea of considering a unique random projection as a discriminative tool in high dimensional problems was put forward in the recent years by Candès and Tao for the development of their "compressive sampling theory".

Remark 3.15 (The conditions are sharp)**.** *The reader may find in* [CAFR07] *and* [BMR97] *several examples showing that the Gilbert moment condition and the projective hypersurface condition are sharp in some sense. In particular, one can find in* [CAFR07]*, for any projective hypersurface S of $\mathbb{R}^d$,* *an explicit couple of mutually singular probability distributions on $\mathbb{R}^d$, compactly supported, such that $\mathcal{E}_1(P,Q)$ is contained in S.*

Equivalently, one can express part of theorem 3.12 and theorem 3.13 as a sort of zero-one law regarding $\mathcal{E}_1(P,Q)$.

Theorem 3.16 (Zero-one law for unidimensional random projections)**.** *Under the assumptions of theorem 3.12 or 3.13 for P and Q on $\mathbb{R}^d$, we have $\lambda(\mathcal{E}_1(P,Q)) \in \{0,1\}$. In other words, $\lambda(\mathcal{E}_1(P,Q)) = 1$ if and only if $P = Q$ and $\lambda(\mathcal{E}_1(P,Q)) = 0$ if and only if $P \neq Q$.*

3.4. **A threshold phenomenon?** Let $\mathcal{Q}$ be the set of probability distributions on $\mathbb{R}^d$ ($d > 1$) with finite absolute moments which satisfy the Carleman condition, and let $\mathbb{P}$ be some probability distribution on $\mathcal{Q}$, for instance the trace of the Dirichlet distribution whose intensity measure is the Lebesgue measure Λ on $\mathbb{R}^d$. Let (X_n) be a sequence of i.i.d. random variables with uniform distribution λ on $\mathbb{S}(\mathbb{R}^d)$. Now fix $P \in \mathcal{Q}$ and define the random subset $\mathcal{Q}_n(P)$ of of $\mathcal{Q}$ by

$$\mathcal{Q}_n(P) = \{Q \in \mathcal{Q}; \{X_1, \ldots, X_n\} \subset \mathcal{E}_1(P,Q)\}.$$

Since $d > 1$, the set $\mathcal{Q}_1(P)$ is infinite for any value of X_1. Moreover, for any value of the sequence (X_n), the sequence of sets $n \mapsto \mathcal{Q}_n(P)$ is non-increasing. Furthermore, if $Q \in \cap_{n>0}\mathcal{Q}_n(P)$, then $(X_n) \subset \mathcal{E}_1(P,Q)$ and thus $\lambda(\mathcal{E}_1(P,Q)) = 1$ which gives $P = Q$. Consequently, if $\mathbb{P}$ has no atoms, then the sequence $n \mapsto \mathbb{P}(\mathcal{Q}_n(P))$ is non-increasing and $\lim_{n\to\infty} \mathbb{P}(\mathcal{Q}_n(P)) = 0$ by inferior limit. One can ask if there exists an integer $n(P)$ such that $\mathbb{P}(\mathcal{Q}_n(P)) = 0$ for every $n \geqslant n(P)$. It is true for finitely supported discrete distributions according to the Rényi-Heppes theorem. Notice that a.-s. the directions $X_1, \ldots, X_{k+1}$ fulfils the assumptions of Rényi and Heppes.

4. Other projection problems

We studied the characterization of probability distributions from one dimensional projections. Here is a small collection of related problems.

4.1. **Projections on subspaces of two or more dimensions.** Prescribing k-dimensional projections with $k > 1$ instead of 1-dimensional projections simplifies the problem, since higher dimensional projections provide in some sense more informations about the projected probability distribution. The problem was studied for instance in [BMR97] and in [CAFR07]. In particular, a sort of Rényi-Gilbert holds true for $k = d-1$, as expressed by the following theorem. Notice that the $k = d-1$ case corresponds to the X-Ray transform mentioned below.

Theorem 4.1 (Rényi [Rén52])**.** *Let P and Q be two probability distributions on $\mathbb{R}^d$ with $d > 1$. Then $P = Q$ if and only if $\mathcal{E}_1(P,Q) \cup \cdots \cup \mathcal{E}_{d-1}(P,Q) = \mathcal{S}(\mathbb{R}^d)$.*

The following result appears as a corollary of theorem 3.13, since a projective hypersurface of $\mathbb{R}^d$ can contain at most finitely many hyperplanes.

Theorem 4.2 (Bélisle-Massé-Ransford [BMR97, CAFR07])**.** *Let P and Q be two probability distributions on $\mathbb{R}^d$ with $d \geqslant 2$. If P or Q has finite absolute moments which satisfies the Carleman condition, then $P = Q$ if and only if $\mathcal{E}_{d-1}(P,Q)$ is infinite.*

However, it is shown in [CAFR07, th. 3.5] that for every projective hypersurface S of $\mathbb{R}^d$, there exists two mutually singular probability distributions P and Q on $\mathbb{R}^d$ such that both P and Q are supported by bounded subsets of $\mathbb{R}^d$ and $\mathcal{E}(P,Q) = S$. This couple of distribution is constructed by using some harmonic analysis, involving Plancherel and Paley-Wiener theorems. Recall that the Paley-Wiener theorem states that an entire function $f : \mathbb{C}^d \to \mathbb{C}$ is the Fourier transform of a compactly supported Schwartz distribution if and only if $|f(z)| \leqslant c(1+|z|)^a e^{b\mathcal{I}m(z)}$ for every $z \in \mathbb{C}^d$ and some constants a, b, c.

4.2. **Infinite dimensional extensions.** The random projection approach can be generalized without much difficulties to Hilbert spaces, by considering for example a Wiener measure (in the sense of Gross) to produce random directions, see for instance [CAFR07].

4.3. **The Radon problem.** As mentioned by Rényi in [Rén52], the characterization of two dimensional probability distributions in terms of one dimensional projections in related to a problem studied by Radon at the beginning of the twentieth century. Radon showed in [Rad17] that if a continuous function $f : K \subset \mathbb{R}^2 \to \mathbb{R}$ defined in a bounded domain K with integral equal to zero along every chord of the domain K then f is identically equal to zero. The continuity of f is required otherwise one may construct counter examples by modifying f. This result by Radon solved a conjecture by Tarski, who asked about the uniqueness when it exists of a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that its integral is constant along every chord of K. The bounded support hypothesis on f can be replaced by an integrability condition on $\mathbb{R}^2$, and it can be shown that the result is false without this integrability condition. Several authors proposed alternative proofs of this theorem by Radon, like for instance [Rén52] and [Gre58].

The Radon theorem mentioned above produced a great amount of work around the so called *Radon transform*. The Radon transform of a function $f : \mathbb{R}^d \to \mathbb{R}$ is a function $R(f) : \mathbb{S}(\mathbb{R}^d) \times \mathbb{R} \to \mathbb{R}$ defined for any $(x, y) \in \mathbb{S}(\mathbb{R}^d) \times \mathbb{R}$ by

$$R(f)(x, y) = \int_{\langle x,z\rangle = y} f(z)\, dz.$$

In other words, $R(f)(x, y)$ is the value of the integral of f over the hyperplane of equation $\langle x, z\rangle = y$ in z. This hyperplane is orthogonal to x. When f is a Lebesgue density of a probability distribution P on $\mathbb{R}^d$, then $y \mapsto R(f)(x, y)$ is the density of the probability distribution $P_{\langle x\rangle}$ on $\mathbb{R}$. Thus, $R(f)(x, \cdot)$ characterizes $P_{\langle x\rangle}$. We thus understand clearly the link between Radon transform and the characterization by projections of probability distributions problem.

Armitage and Goldstein have shown in [AG93] that there exists for each $d \geqslant 2$ a nonconstant harmonic function f on $\mathbb{R}^d$ such that f is integrable over each hyperplane, and the Radon transform of f is zero. In the case $d = 2$ this result has been obtained by Zalcman in [Zal82]. However, for any function $f \in \mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{L}^2(\mathbb{R}^d)$, we have

$$\widehat{R(f)}(x, y) = \widehat{f}(yx)$$

for evey $(x, y) \in \mathcal{S}(\mathbb{R}^d) \times \mathbb{R}$ where $\widehat{f}$ is the Fourier transform of f and where $\widehat{R(f)}$ is the Fourier transform of $R(f)$ with respect to its second argument. In particular, the Radon transform restricted to $\mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{L}^2(\mathbb{R}^d)$ is injective.

The Radon transform of a probability distribution P on $\mathbb{R}^d$ is the function $R(P)$ which associate to every $x \in \mathbb{S}(\mathbb{R}^d)$ the probability distribution $P_{\langle x\rangle}$ on $\mathbb{R}$. Notice that the Cramér-Wold theorem can be seen as an injectivity result for the Radon transform of probability distributions, based on harmonic analysis (Fourier transforms). The injectivity of the Radon transform restricted to $\mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{L}^2(\mathbb{R}^d)$ mentioned above is the function version of the Cramér-Wold theorem.

There is a huge litterature on Radon transforms of functions, dealing with regularity properties, supports, inversion formulas, etc. These aspects involve harmonic analysis, like in [Jen04]. The Radon transform of functions is used for the mathematical modeling of tomography in $\mathbb{R}^2$ and $\mathbb{R}^3$, see [Nat01]. It should not be confused with the X-Ray transform of $f : \mathbb{R}^d \to \mathbb{R}$ which associate to every $x \in \mathbb{S}(\mathbb{R}^d)$ the integral of f over the straight line $\mathbb{R}x$. In $\mathbb{R}^2$, these two transforms are identical up to parameterization. Tomography consists in reconstructing bodies from their lower dimensional projections, and is a useful tool in medicine and industry (e.g. Positon Emission Tomography).

The reconstruction of probability densities and more generally of probability distributions from partial informations on their Radon transform has been studied by the Statisticians Tsybakov, Natterer, Korostelev, and Donoho among others. See for instance [Cav00] and references therein. More precisely, they consider for example the inverse problem of the estimation of a Lebesgue density f on $\mathbb{R}^d$ from n i.i.d. observations drawn from a Lebesgue density proportional to the Radon transform $R(f)$ (each observersation is a random realisation of a random projection).

The reader may find in [HQ85] and [HHK83] very interesting results regarding Radon transforms of probability distributions. In particular, [HQ85] is entitled "*Distances between measures from* 1-*dimensional projections as implied by continuity of the inverse Radon transform*"! Some sort of speed of convergence in the Central Limit Theorem associated to Cramér-Wold theorem.

Radon transforms can be seen as the collection of integrals of a function or the collection of projections of a probability distribution over the set of proper subspaces of fixed dimension. One can thus define Radon transforms more generally on Grassmannian manifolds. The reader may find more material on Radon transforms, integral geometry, and tomography in the books [Hel99], [Dea83], [Ehr03], [Nat01], [MQ06] for instance. The reader may find for instance in [KRZ04] and [GZ98] some aspects of the geometric problem of reconstructing a convex body from lower dimensional projections.

The Radon transform considered in the above should not be confused with the spherical Radon transform, which corresponds to an integral over centered spheres:

$$\big(f : \mathbb{R}^d \to \mathbb{R}\big) \mapsto \left(r \in \mathbb{R}_+ \mapsto \int_{\mathbb{S}(\mathbb{R}^d)} f(rx)\, d\lambda(x)\right).$$

4.4. **The Strassen problem.** Strassen studied in [Str65] the problem of finding a probability distribution on a product $E \times F$ of Polish spaces, with prescribed closed support S and prescribed marginals on E and F. His method makes use of Choquet capacities and involves stochastic domination. This problem was studied by many authors. See for instance the article [Sho83] by Shortt and references therein, where the results of Strassen are generalized to countable products of more general measurable factor spaces, and beyond the context of product spaces.

Here are some connected problems: construction of Markovian coupling with prescribed marginals, construction of martingales with prescribed marginals, graphical models and G-Markov distributions, IPF algorithm, etc.

An interesting problem is graphical constraints on projections. More precisely, let $\mathcal{S}$ be a finite subset of $\mathbb{S}(\mathbb{R}^d)$, and consider the problem of finding probability distributions on $\mathbb{R}^d$ with prescribed marginals on a subset of the collection of subspaces of various dimensions generated by finite subsets of $\mathcal{S}$. The Rényi-Gilbert and Rényi-Heppes theorems and the Gaussian counter examples of Hamedani suggest that this problem is not trivial for finitely many directions, at least for probability distributions with infinite support. How about existence, uniqueness, and construction?

4.5. **More recent development.** See for instance [BEG18] and [BEGP17] and references therein.

4.6. **Final remark.** The following themes where studied by Mathematicians and are still active. They are deeply connected, but the connectivity of the citation graph does not reproduce well these connections.

- Radon transforms, integral geometry, statistical aspects of reconstruction by tomography;
- Strassen problems, couplings, and stochastic domination;
- Rény-Gilbert problems and the Cramér-Wold theorem;
- Copula in Statistics (laws on cubes with uniform marginals);
- Graphical models and marginal constraints in Statistics;
- Reconstruction of bodies from their lower dimensional projections in high dimensional convex geometry;
- Stieltjes-Hamburger moments problems and Denjoy-Carleman quasi-analycity;
- Random projections and Candès-Tao compressive sampling in Statistics.

To summarize, all these themes deal with linear problems in infinite dimensions, with often some kind of positivity constraints.

5. Acknowledgments

These notes benefited from the feedback of Jordan Stoyanov (Newcastle, UK) and Oskar Laverny (Lyon, France).

References

[AG93] D. H. Armitage and M. Goldstein, *Nonuniqueness for the Radon transform*, Proc. Amer. Math. Soc. **117** (1993), no. 1, 175–178. MR MR1106177 (93c:44004) 4.3

[Akh65] N. I. Akhiezer, *The classical moment problem and some related questions in analysis*, Translated by N. Kemmer, Hafner Publishing Co., New York, 1965. MR MR0184042 (32 #1518) 2.1

[Bak01] A. G. Bakan, *Polynomial density in $L_p(\mathbb{R}, d\mu)$ and representation of all measures which generate a determinate Hamburger moment problem*, Approximation, optimization and mathematical economics (Pointe-à-Pitre, 1999), Physica, Heidelberg, 2001, pp. 37–46. MR MR1842874 (2002i:41004) 2.2

[Bak03] ———, *A criterion for the density of algebraic polynomials in the spaces $L_p(\mathbb{R}, d\mu)$, $1 \leqslant p < \infty$*, Ukraïn. Mat. Zh. **55** (2003), no. 5, 701–705. MR MR2071709 (2005e:46042) 2.2

[BEG18] Bernhard G Bodmann, Martin Ehler, and Manuel Gräf, *From low-to high-dimensional moments without magic*, Journal of Theoretical Probability **31** (2018), no. 4, 2167–2193. 4.5

[BEGP17] Anna Breger, Martin Ehler, Manuel Gräf, and Thomas Peter, *Cubatures on grassmannians: moments, dimension reduction, and related topics*, Compressed Sensing and its Applications, Springer, 2017, pp. 235–259. 4.5

[Ber96] Ch. Berg, *Moment problems and polynomial approximation*, Ann. Fac. Sci. Toulouse Math. (6) (1996), no. Special issue, 9–32, 100 ans après Th.-J. Stieltjes. MR MR1462705 (98h:44002) 2.2

[BMR97] C. Bélisle, J.-C. Massé, and Th. Ransford, *When is a probability measure determined by infinitely many projections?*, Ann. Probab. **25** (1997), no. 2, 767–786. MR MR1434125 (98c:60011) 2.2, 2.9, 2.2, 3.1, 3.4, 3.9, 3.15, 4.1, 4.2

[BT06] Ch. Bayer and J. Teichmann, *The proof of Tchakaloff's theorem*, Proc. Amer. Math. Soc. **134** (2006), no. 10, 3035–3040 (electronic). MR MR2231629 (2007d:44004) 2.4

[CAFR07] J. A. Cuesta-Albertos, R. Fraiman, and Th. Ransford, *A sharp form of the Cramér-Wold theorem*, to appear in Journal of Theoretical Probability, 2007. 2.2, 2.10, 3.3, 3.13, 3.3, 3.15, 4.1, 4.2, 4.1, 4.2

[Car26] T. Carleman, *Fonctions quasi-analytiques*, Gauthier-Villars, Paris, 1926. 2.2

[Cav00] L. Cavalier, *Efficient estimation of a density in a problem of tomography*, Ann. Statist. **28** (2000), no. 2, 630–647. MR MR1790012 (2001h:62060) 4.3

[CF02] R. E. Curto and L. A. Fialkow, *A duality proof of Tchakaloff's theorem*, J. Math. Anal. Appl. **269** (2002), no. 2, 519–532. MR MR1907129 (2003e:41047) 2.2

[CW36] H. Cramér and H. Wold, *Some theorems on distribution functions*, J. London Math. Soc. (1936), no. 11, 290–295. 1.2

[Dea83] S. R. Deans, *The Radon transform and some of its applications*, A Wiley-Interscience Publication, John Wiley & Sons Inc., New York, 1983. MR MR709591 (86a:44003) 4.3

[DF04] P. Diaconis and D. Freedman, *The Markov moment problem and de Finetti's theorem. I*, Math. Z. **247** (2004), no. 1, 183–199. MR MR2054525 (2005m:60004) 2.2

[Ehr03] L. Ehrenpreis, *The universality of the Radon transform*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2003, With an appendix by Peter Kuchment and Eric Todd Quinto. MR MR2019604 (2007d:58047) 4.3

[Fel71] W. Feller, *An introduction to probability theory and its applications. Vol. II.*, Second edition, John Wiley & Sons Inc., New York, 1971. MR MR0270403 (42 #5292) 2.3, 2.2, 3.7

[Fer59] T. Ferguson, *On the determination of joint distributions from the marginal distributions of linear combinations*, Ann. Math. Stat. **30** (1959), no. 1, 255, abstract only. 3, 3.9

[FP05] L. Fialkow and S. Petrovic, *A moment matrix approach to multivariable cubature*, Integral Equations Operator Theory **52** (2005), no. 1, 85–124. MR MR2138699 (2005k:47032) 2.2

[Gil55] W. M. Gilbert, *Projections of probability distributions*, Acta Math. Acad. Sci. Hungar. (1955), no. 6, 195198. 3.1, 3.7

[Gre58] J. W. Green, *On the determination of a function in the plane by its integrals over straight lines*, Proc. Amer. Math. Soc. **9** (1958), 758–762. MR MR0096770 (20 #3253) 4.3

[GZ98] P. Goodey and G. Zhang, *Inequalities between projection functions of convex bodies*, Amer. J. Math. **120** (1998), no. 2, 345–367. MR MR1613642 (99h:52010) 4.3

[Ham21] H. Hamburger, *Über eine Erweiterung des Stieltjesschen Momentenproblems I-III*, Math. Ann. **81, 82, 83** (1920, 1921, 1921), 235–319, 120–164, 168–187. 2, 2.1

[Ham84] G. G. Hamedani, *Nonnormality of linear combinations of normal random variables*, Amer. Statist. **38** (1984), no. 4, 295–296. MR MR770261 (86b:62081) 3.8

[Hau23] F. Hausdorff, *Momentenprobleme für ein endliches Intervall*, Math. Z. **16** (1923), 220–248. 2.2

[Hel99] S. Helgason, *The Radon transform*, second ed., Progress in Mathematics, vol. 5, Birkhäuser Boston Inc., Boston, MA, 1999. MR MR1723736 (2000m:44003) 4.3

[Hep56] A. Heppes, *On the determination of probability distributions of more dimensions by their projections*, Acta Math. Acad. Sci. Hungar. (1956), no. 7, 403410. 3.2, 3.3

[Her79] A. Hertle, *Zur Radon-Tranformation von Functionen und Massen*, Ph.D. thesis, Erlangen, 1979, Thesis. 3.5

[HHK83] M. G. Hahn, P. Hahn, and M. J. Klass, *Pointwise translation of the Radon transform and the general central limit problem*, Ann. Probab. **11** (1983), no. 2, 277–301. MR MR690129 (84h:60047) 4.3

[HQ85] M. G. Hahn and E. T. Quinto, *Distances between measures from* 1*-dimensional projections as implied by continuity of the inverse Radon transform*, Z. Wahrsch. Verw. Gebiete **70** (1985), no. 3, 361–380. MR MR803678 (87b:60026) 4.3

[HT75] G. G. Hamedani and M. N. Tata, *On the determination of the bivariate normal distribution from distributions of linear combinations of the variables*, Amer. Math. Monthly **82** (1975), no. 9, 913–915. MR MR0383643 (52 #4523) 3.8

[Jen04] S. R. Jensen, *Sufficient conditions for the inversion formula for the k-plane Radon transform in* $\mathbf{R}^n$, Math. Scand. **94** (2004), no. 2, 207–226. MR MR2053741 (2005e:44002) 4.3

[KN77] M. G. Kreĭn and A. A. Nudel′man, *The Markov moment problem and extremal problems*, American Mathematical Society, Providence, R.I., 1977, Ideas and problems of P. L. Čebyšev and A. A. Markov and their further development, Translated from the Russian by D. Louvish, Translations of Mathematical Monographs, Vol. 50. MR MR0458081 (56 #16284) 2.1

[KRZ04] A. Koldobsky, D. Ryabogin, and A. Zvavitch, *Projections of convex bodies and the Fourier transform*, Israel J. Math. **139** (2004), 361–380. MR MR2041799 (2005f:52015) 4.3

[Lév37] P. Lévy, *Théorie de l'addition des variables aléatoires*, first edition, Gauthier-Villars, Paris, 1937, reprint of second edition published in 2003 by Jacques Gabay with ISBN 2876472074. 3.7

[MQ06] A. Markoe and E. T. Quinto (eds.), *Integral geometry and tomography*, Contemporary Mathematics, vol. 405, American Mathematical Society, Providence, RI, 2006, Papers from the AMS Special Session on Tomography and Integral Geometry held at Rider University, Lawrenceville, NJ, April 17–18, 2004. MR MR2233611 (2007c:00008) 4.3

[Nat01] F. Natterer, *The mathematics of computerized tomography*, Classics in Applied Mathematics, vol. 32, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001, Reprint of the 1986 original. MR MR1847845 (2002e:00008) 4.3

[Pak01] A. G. Pakes, *Remarks on converse Carleman and Kreĭn criteria for the classical moment problem*, J. Aust. Math. Soc. **71** (2001), no. 1, 81–104. MR MR1840495 (2002e:44007) 2

[Put97] M. Putinar, *A note on Tchakaloff's theorem*, Proc. Amer. Math. Soc. **125** (1997), no. 8, 2409–2414. MR MR1389533 (97j:65045) 2.2

[PV99] M. Putinar and F.-H. Vasilescu, *Solving moment problems by dimensional extension*, Ann. of Math. (2) **149** (1999), no. 3, 1087–1107. MR MR1709313 (2001c:47023b) 2.2

[Rad17] J. Radon, *Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten*, Ber. Math.-Phys. Kl. Sdchs. Ges. Wiss. Leipzig **59** (1917), 262–277. 4.3

[Rén52] A. Rényi, *On projections of probability distributions*, Acta Math. Acad. Sci. Hungar. (1952), no. 3, 131142. 3.1, 3.3, 3.1, 4.1, 4.3

[Rud87] Walter Rudin, *Real and complex analysis*, third ed., McGraw-Hill Book Co., New York, 1987. MR MR924157 (88k:00002) 2.2

[Sho83] R. M. Shortt, *Strassen's marginal problem in two or more dimensions*, Z. Wahrsch. Verw. Gebiete **64** (1983), no. 3, 313–325. MR MR716489 (84i:60023) 4.4

[Sit83] A. Sitaram, *A theorem of Cramér and Wold revisited*, Proc. Amer. Math. Soc. **87** (1983), no. 4, 714–716. MR MR687648 (84d:60013) 3.5

[ST43] J. A. Shohat and J. D. Tamarkin, *The Problem of Moments*, American Mathematical Society Mathematical surveys, American Mathematical Society, New York, 1943. MR MR0008438 (5,5c) 2.1

[Sti93] Th. J. Stieltjes, *Œuvres complètes/Collected papers. Vol. I, II*, Springer-Verlag, Berlin, 1993, Reprint of the 1914–1918 edition, Edited and with a preface and a biographical note by Gerrit van Dijk, With additional biographical and historical material by Walter Van Assche, Frits Beukers, Wilhelmus A. J. Luxemburg and Herman J. J. te Riele. MR MR1272017 (95g:01033) 2

[Sto00] J. Stoyanov, *Krein condition in probabilistic moment problems*, Bernoulli **6** (2000), no. 5, 939–949. MR MR1791909 (2001i:44014) 2.2

[Str65] V. Strassen, *The existence of probability measures with given marginals*, Ann. Math. Statist **36** (1965), 423–439. MR MR0177430 (31 #1693) 4.4

[Tch57] V. Tchakaloff, *Formules de cubatures mécaniques à coefficients non négatifs*, Bull. Sci. Math. (2) **81** (1957), 123–134. MR MR0094632 (20 #1145) 2.4

[Zal82] L. Zalcman, *Uniqueness and nonuniqueness for the Radon transform*, Bull. London Math. Soc. **14** (1982), no. 3, 241–245. MR MR656606 (83h:42020) 4.3

URL: `https://djalil.chafai.net/`