View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Available online at www.sciencedirect.com
ScienceDirect
Indagationes Mathematicae 23 (2012) 848–883
indagationes mathematicae
www.elsevier.com/locate/indag

Conditioning moments of singular measures for entropy optimization. I☆

Marko Budišić[a,*], Mihai Putinar[b]

[a] *Department of Mechanical Engineering, University of California, Santa Barbara, CA 93106, United States*
[b] *Department of Mathematics, University of California, Santa Barbara, CA 93106, United States*

Abstract

In order to process a potential moment sequence by the entropy optimization method one has to be assured that the original measure is absolutely continuous with respect to Lebesgue measure. We propose a non-linear exponential transform of the moment sequence of any measure, including singular ones, so that the entropy optimization method can still be used in the reconstruction or approximation of the original. The Cauchy transform in one variable, used for this very purpose in a classical context by A.A. Markov and followers, is replaced in higher dimensions by the Fantappiè transform. Several algorithms for reconstruction from moments are sketched, while we intend to provide the numerical experiments and computational aspects in a subsequent article. The essentials of complex analysis, harmonic analysis, and entropy optimization are recalled in some detail, with the goal of making the main results more accessible to non-expert readers.
© 2012 Royal Dutch Mathematical Society (KWG). Published by Elsevier B.V. All rights reserved.

keywords: Fantappiè transform; Entropy optimization; Moment problem; Tube domain; Exponential transform

1. Introduction

In sciences and engineering, a particular inverse problem arises often, requiring approximation of a measure by a density function from knowledge of linear data, e.g., integrals of a function

☆ Israel Gohberg, in memoriam.

* Corresponding author. Tel.: +1 805 893 5095.
E-mail addresses: mbudisic@engr.ucsb.edu (M. Budišić), mputinar@math.ucsb.edu (M. Putinar).

0019-3577/$ - see front matter © 2012 Royal Dutch Mathematical Society (KWG). Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.indag.2012.05.008

basis against a measure. Classical moment problem considers integrals of monomials, the power moments, as the set of known measurements, while the generalized moment problems extend the admissible inputs to integrals of orthogonal polynomials, Fourier basis, wavelets, or other functional bases.

The list of applications of the moment problem is long, ranging from engineering, through physics, statistics, well into applied mathematics. While pure mathematical settings allow for infinite moment sequences, leading to classical moment problems of Hausdorff, Hamburger, and Stieltjes, the applied settings almost exclusively assume knowledge of only a finite number of moments, which is known as the truncated moment problem.

In early 1980s, statistical physics and signal processing communities recognized that a practical solution to the truncated moment problem, which is mathematically under-determined, can be found through optimization of the Shannon entropy, a nonlinear functional acting on the density of the measure [39,29]. Initial success, in the numerically unfavorable setting of power moments, generated sufficient interest to improve on the original method [30,8,1,11] and arrive at a routinely-used method not only in physics, but also in statistics and control theory [22,14,23]. Furthermore, optimization of entropy has been shown to be of theoretical importance: it can be used to fully characterize the moment sequences representable by densities based on truncated moment data [12], and arbitrarily incomplete moment data [5].

However, not every moment sequence is a suitable input for the entropy optimization. In particular, it is easy to demonstrate that the moment sequence of the Dirac-δ distribution is not a feasible input, as the optimization does not converge in that case. Such singular measures captured our focus, as we were motivated by potential applications to inverse problems in dynamical systems.

Measures invariant under evolution of dynamical systems are of particular interest, with increasing activity driven by applied problems. On chaotic attractors, trajectories of dynamical systems are known to be non-robust to any errors and behavior is more reliably represented using statistical methods [18]. Surprisingly, even in chaotic regimes, the moment data of invariant measures can be reliably computed from simulated and experimental trajectories by averaging moment functions along them, despite the errors inherent to those procedures [47]. Singular invariant measures abound in dynamical systems, e.g., a system with an attracting fixed point preserves a Dirac-δ distribution, whose moments are easily computed by averaging along any trajectory in the basin of attraction. As mentioned before, entropy optimization would not converge for such a common invariant measure.

To overcome the obstacle of singular measures, we propose a three step process: (i) regularization, (ii) entropy optimization, and (iii) inversion. Regularization conditions the moment sequence into a feasible input to the entropy optimization, converting the original moment sequence into moments of a bounded, integrable *phase function*. The entropy optimization step can then be used to recover a closed expression for the phase function approximation. In the inversion step, point-wise evaluations of the phase function are used to recover an approximant of the original measure.

The proposed regularization of the moment sequence of a singular, positive measure derives from an original idea of Markov to study the moment sequence through its complex generating function. We start by a simple observation that an analytic function mapping a domain into the open upper-half plane admits an analytic logarithm whose imaginary part (the phase) is bounded from below by 0 and from above by π. The passage from a positive measure to the phase function through a canonical integral transform, obeying the above principle, has circulated in the Russian literature in connection with the century old works devoted to the one dimensional

L-problem of moments. The early articles by Krein, Akhiezer and Nudelman on the subject offer a comprehensive account of this method [34,35].

In the present article we go beyond one dimension, considering Fantappiè transforms of positive measures supported by a wedge in $\mathbb{R}^d$ [6,38]. The existing methods of harmonic analysis on tube domains enter naturally into the picture offering to the maximum entropy reconstruction method a solid background. The much nicer sequence of moments of the phase function are obtained from the moment sequence of the original measure via a non-linear recurrent operation. A thorough investigation of the multivariate moment via asymptotic expansions of the Fantappiè transform of the underlying measure was undertaken by Henkin and Shananin [24,25], whose work we take as a basis for ours.

While the entropy optimization provides a standard reconstruction procedure for the phase function, the approaches to inversion for one- and multi-variate problems are different. In one-dimensional case, we can make use of the well known Plemelj–Sokhotski formulas [28,31] to complete the inversion step. The formulas, however, are difficult to generalize to multivariate settings [21]; instead, we propose a ray beam disintegration, based on a refined and partially forgotten one-dimensional analysis of the phase regularization due to Aronszajn and Donoghue [7]. The ray beam approach reduces the problem to a setting similar to medical tomography, based on inverse Radon or Laplace transform methods [41,42]. We believe this will be a fruitful approach that we plan on exploring in follow-up papers, so we only draft it in this paper.

The paper is organized using the following outline. Section 2 briefly introduces the multivariate moment problem and the entropy optimization, including an example illustrating lack of convergence for a Dirac-δ measure. In Section 3 we expose the elementary aspects of the entropy optimization method, in the case of one real variable for unbounded and bounded supports, using, respectively, power moments and trigonometric moments, i.e., Fourier coefficients. Section 5 is devoted to generalization to multivariate problems, through the phase regularization of the Fantappiè transform of a measure supported by a wedge in Euclidean space (Section 5.1) and by special compact domains in Euclidean space (Section 5.2). A Riesz–Herglotz formula is derived, in the spirit of [32,2], with a couple of examples on product domains.

The present article remains at a theoretical level, leaving for a continuation of it to deal with further practical aspects: numerical experiments, the error analysis and examples from dynamical systems. We do, however, present practical algorithms that are essential for moment conditioning, the Miller–Nakos algorithm in Appendix A, described in [40], and a recent algorithm for entropy optimization, described in [8], in Appendix B.

We dedicate this work to the late Israel I. Gohberg, legendary figure of modern operator theory and function theory. His original and highly influential ideas have permanently shaped moment problems and the entropy method referred to in the following pages.

2. Preliminaries

Let $d \geq 1$ be a fixed dimension and let K be a closed subset of the Euclidean space $\mathbb{R}^d$. Fix a finite set $A \subset \mathbb{N}^d$ of multi-indices. The *truncated moment problem* with supports on K and monomials labeled by A consists in finding (as effectively as possible) a positive measure μ supported by K, with prescribed moments

$$\gamma_\alpha = \int_K x^\alpha d\mu(x), \quad \alpha \in A. \tag{1}$$

In case the set K is unbounded, it is implicit that the above integrals converge in Lebesgue sense. Throughout this article we adopt the multi-index notation

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}, \quad x \in \mathbb{R}^d.$$

A few basic questions are in order:

1. *Characterize all sequences of moments $(a_\alpha)_{\alpha\in A}$ associated to positive measures carried by the set K.*

This question can be rephrased in terms of the formal integration functional

$$L(f) = \sum_{\alpha\in A} c_\alpha \gamma_\alpha, \quad f = \sum_{\alpha\in A} c_\alpha x^\alpha.$$

Let us denote by $\mathbb{R}[x]_A$ the linear span, in the ring of polynomials $\mathbb{R}[x]$, of all monomials x^α, $\alpha \in A$.

A necessary and sufficient condition that a linear functional $L : \mathbb{R}[x] \longrightarrow \mathbb{R}$ is representable by a positive measure supported by the set K is that L in non-negative on all elements $f \in \mathbb{R}[x]$ which are non-negative on K. Then L can be extended via a Hahn–Banach construction to a positive linear functional on the space of continuous functions on K, with polynomial growth at infinity. This observation remains however of a limited theoretical importance, and it becomes effective only when simple characterizations of non-negative polynomials on K is available. Fortunately, in the case when K is a basic semi-algebraic set, such "Positivstellensätze" were recently resurrected and a good collection of examples is available, see [44].

The single variable case is the simplest and best understood. The following result goes back to Marcel Riesz [45].

Theorem 1. *Let n be a fixed degree and (a, b) an interval on the real line, bounded or not. A positive measure μ carried by the closure of (a, b) exists, with moments*

$$\gamma_k = \int x^k d\mu(x), \quad 0 \le k < n,$$

and

$$\gamma_n \ge \int x^n d\mu$$

if and only if the associated functional L satisfies $L(f) \ge 0$ for all polynomials $f(x) = c_0 + c_1 x + \cdots + c_n x^n$ which are non-negative on (a, b).

Three cases are distinguished, and they correspond to classical moment problem studies: $(a, b) = (0, 1)$, known as the Hausdorff moment problem, $(a, b) = (0, \infty)$ known as Stieltjes moment problem, and $(a, b) = (-\infty, \infty)$ known as the Hamburger moment problem. In each separate situation a full characterization of all non-negative polynomials on (a, b) is available, with the result of making the above Riesz result effective. We refer the reader to [4,46] for full details.

2. *Knowing that problem (1) is solvable, find constructively one particular solution.*

As a general rule, any attempt to solve the truncated problem (1) starts with the observation that the set of all solutions

$$\Sigma = \left\{ \mu \ge 0; \ \int_K x^\alpha d\mu = \gamma_\alpha, \ \alpha \in A \right\}$$

is convex and closed in the weak-* topology. If we include $\alpha = 0$ among the elements of the index set A, then all elements of Σ have fixed total variation. Thus, in this case, on a compact support K, the set of solutions Σ is compact in the weak-* topology of all measures.

Among all elements of the solution set Σ the extremal ones are the first to be detected by linear optimization methods. For example, in the case of the three classical truncated moment problems on the line, they correspond to convex combinations of point masses. Their support is identified with the zero set of orthogonal polynomials, and the multipliers of the Dirac measures are also computable in terms of the diagonal Padé approximation of the series:

$$-\frac{\gamma_0}{z} - \frac{\gamma_1}{z^2} - \cdots - \frac{\gamma_n}{z^{n+1}}.$$

Stieltjes original memoir remains unsurpassed for a careful analysis of this approximation scheme, see for instance [4]. A basic observation in this direction, providing an extremal solution to Stieltjes moment problem with the data $(\gamma_0, \ldots, \gamma_{2n-1})$ is the following: assuming that the Hankel matrices

$$\begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_2 & \cdots & \gamma_n \\ \vdots & & & \vdots \\ \gamma_{n-1} & \gamma_n & \cdots & \gamma_{2n-1} \end{pmatrix}, \qquad \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_n \\ \gamma_2 & \gamma_3 & \cdots & \gamma_{n+1} \\ \vdots & & & \vdots \\ \gamma_n & \gamma_{n+1} & \cdots & \gamma_{2n-1} \end{pmatrix} \tag{2}$$

are positive definite, the one step completion $(\gamma_0, \ldots, \gamma_{2n-1}, \tilde{\gamma_{2n}})$ so that the determinant

$$\begin{vmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_n \\ \gamma_1 & \gamma_2 & \cdots & \gamma_{n+1} \\ \vdots & & & \vdots \\ \gamma_n & \gamma_{n+1} & \cdots & \tilde{\gamma_{2n}} \end{vmatrix} = 0$$

vanishes, has a unique, necessarily finite, atomic solution.

Since the computation of the roots of an orthogonal polynomial is not friendly from the numerical point of view, the search for other special solutions of the truncated moment problem led to adopt a statistical point of view, and consider "the most probable" solutions, with respect to a non-linear, concave functional. Recent applications (in particular to continuum mechanics) use to this aim the Boltzmann–Shannon entropy, see [9,13,22,30,37,39]. The entropy maximization method for the trigonometric moment problem stands aside for clarity and depth in this framework, see [36].

A great deal of recent work, cf. [30,23], has clarified the existence of maximum-entropy solutions, especially in some degenerate cases. We start from there, and add a computational/numerical analysis component to the study.

3. Maximal entropy solutions in 1D

For the sake of clarity we digress and specialize the above discussion to the simplest and best-understood framework. Namely, we discuss below the existence and uniqueness of maximum entropy solutions to the truncated moment problem in the case of a single variable.

3.1. Basic properties

Although an abstract, fairly general treatment of the maximum entropy method is nowadays available, see or instance [30,13], we specialize below on an interval of the real line. To this aim, we go back to Riesz' existence theorem stated in the previous section. Namely, n is a fixed degree and (a, b) is an interval on the real line, bounded or not. We start with the moment data $\gamma_0, \ldots, \gamma_n$, and seek a positive measure μ carried by the closure of (a, b) satisfying

$$\gamma_k = \int x^k d\mu(x), \quad 0 \le k < n,$$

and

$$\gamma_n \ge \int x^n d\mu.$$

We search μ of the from $d\mu(x) = \exp(\lambda_0 + \lambda_1 x + \cdots + \lambda_n x^n)dx$, assuming that the integrability condition

$$\int_a^b \exp(\lambda_0 + \lambda_1 x + \cdots + \lambda_n x^n)dx < \infty$$

is assured by the choice of the parity and sign of the leading term. For instance, in case $a = 0, b = \infty$ we must have $\lambda_p < 0$ and $\lambda_{p+1} = \lambda_{p+2} = \lambda_n = 0$; or in the case $a = -\infty, b = \infty$ we must have $\lambda_{2p} < 0$ and $\lambda_{2p+1} = \lambda_{2p+2} = \lambda_n = 0$. We denote by Λ (by omitting the subscript n) the set of all such multipliers which produce integrable exponentials.

The proper choice of the parameters λ_k is made by imposing the optimality (maximum entropy) condition:

$$\sup\left\{\lambda_0\gamma_0 + \cdots + \lambda_n\gamma_n - \int_a^b \exp(\lambda_0 + \lambda_1 x + \cdots + \lambda_n x^n)dx\right\} \tag{3}$$

where the supremum is taken over all admissible (i.e. integrable exponential) tuples $\lambda = (\lambda_0, \ldots, \lambda_n)$. Let us similarly denote $\gamma = (\gamma_0, \ldots, \gamma_n)$ and $\mathbf{x} = (1, x, x^2, \ldots, x^n)$, where the latter is considered as a variable point on the Veronese curve described by the list of the first monomials.

The starting point of our discussion is the observation that the functional

$$L : \Lambda \longrightarrow \mathbb{R}, \qquad L(\lambda) = \lambda \cdot \gamma - \int_a^b \exp[\lambda \cdot \mathbf{x}]dx,$$

is concave. Indeed, whenever the partial derivatives are defined (for instance in the Euclidean interior of Λ), we have

$$\frac{\partial^2 L}{\partial\lambda_i \partial\lambda_j} = -\int_a^b x^{i+j} \exp[\lambda \cdot \mathbf{x}]dx.$$

In the above Hessian, we recognize the negative of the Hankel matrix of a non-atomic positive measure, whence the strict negative definiteness. Moreover, the inner critical points of the functional are given by the vanishing gradient conditions:

$$\frac{\partial L}{\partial\lambda_j} = \gamma_j - \int_a^b x^j \exp[\lambda \cdot \mathbf{x}]dx = 0.$$

The difficulty related to the described method lies in the complicated structure of the set Λ of admissible multipliers. While for a bounded interval (a, b) this set is the whole Euclidean space $\Lambda = \mathbb{R}^{n+1}$, the case $(a, b) = (0, \infty)$ requires:

$$\Lambda = [\mathbb{R}^n \times (-\infty, 0)] \cup [\mathbb{R}^{n-1} \times (-\infty, 0) \times \{0\}] \cup \cdots \cup [\mathbb{R} \times \{0\} \times \cdots \times \{0\}].$$

And similarly when $(a, b) = (-\infty, \infty)$. On the positive side, we remark following Junk [30] that in all cases the assumption that γ is a moment sequence implies

$$\lim_{|\lambda| \to \infty} L(\lambda) = -\infty.$$

Thus, in the bounded interval case, the optimization problem (3) always has a solution, and by strict convexity, this is unique. Note that in this situation, the positivity conditions in Riesz Theorem (or equivalently Hausdorff finite difference conditions) are necessary and sufficient for the existence of an exponential type solution to the truncated moment problem, see also [39] for a detailed discussion.

A much more delicate analysis is required in the case of Stieltjes moment problem $(a, b) = (0, \infty)$. For this case it is very possible that the extremal value in problem (3) is attained on the boundary of the set Λ. Assume for instance that

$$\begin{aligned} &\sup \left\{ \lambda_0 \gamma_0 + \cdots + \lambda_n \gamma_n - \int_a^b \exp(\lambda_0 + \lambda_1 x + \cdots + \lambda_n x^n) dx \right\} \\ &\quad = \sigma_0 \gamma_0 + \cdots + \sigma_n \gamma_n - \int_a^b \exp(\sigma_0 + \lambda_1 x + \cdots + \sigma_n x^n) dx \end{aligned}$$

where $\sigma = (\sigma_0, \ldots, \sigma_n) \in \Lambda \setminus \operatorname{int} \Lambda$. That is, there exists an index $0 < p < n$ with the property

$$\sigma_{p-1} < 0 = \sigma_p = \cdots = \sigma_n$$

if $p > 1$, or simply

$$0 = \sigma_1 = \cdots = \sigma_n$$

in case $p = 1$. Anyway, then only lateral partial derivatives $\frac{\partial L}{\partial \lambda_j}(\sigma)$ exist for all $p \le j \le n$. Since σ is a global maximum, we infer

$$\begin{cases} \gamma_j - \displaystyle\int_0^\infty x^j \exp[\sigma \cdot \mathbf{x}] dx = \frac{\partial L}{\partial \lambda_j}(\sigma) \ge 0, & p \le j \le n, \\ \gamma_j - \displaystyle\int_0^\infty x^j \exp[\sigma \cdot \mathbf{x}] dx = \frac{\partial L}{\partial \lambda_j}(\sigma) = 0, & j < p. \end{cases}$$

Note that above, the exponential density depends only on p parameters $(\sigma_0, \ldots, \sigma_{p-1})$, whence it is normal to expect that only the first p moments are matched.

A detailed analysis of the decision tree resulting from the above observations goes as back as 1977 to Einbu [19] and it was much clarified in the recent works by Junk [30] and Hauck et al. [23]. We reproduce below, following Einbu and Junk, the main phenomenon, in the form of an analysis of a one step extension.

Suppose that, for the truncated version of Stieltjes moment problem, the initial segment of moments

$$(\gamma_0, \gamma_1, \ldots, \gamma_{n-1})$$

is realized by the maximum entropy method, that is there is an admissible tuple $\sigma = (\sigma_0, \ldots, \sigma_n)$, such that

$$\gamma_j = \int_0^\infty x^j \exp[\sigma \cdot \mathbf{x}]dx, \quad 0 \le j \le n-1.$$

This implies that Hankel's positivity conditions (2) hold true, and that the (lateral) partial derivatives of the function $L(\lambda)$ vanish at $\lambda = \sigma$.

We assume next that the extended moment sequence $(\gamma_0, \gamma_1, \ldots, \gamma_{n-1}, \delta)$ is also realizable by the maximal entropy method. Hankel's positivity conditions (2) imply

$$\delta \ge \gamma_n(\min),$$

where the bound $\gamma_n(\min)$ is a rational function of the data $(\gamma_0, \gamma_1, \ldots, \gamma_{n-1})$, expressed as a quotient of Hankel type determinants. Define

$$\gamma_n(\max) = \int_0^\infty x^n \exp[\sigma \cdot \mathbf{x}]dx.$$

This corresponds to the boundary point $(\sigma_0, \ldots, \sigma_n, 0) \in \Lambda_n$, and in addition we know that the function $L : \Lambda_n \longrightarrow \mathbb{R}$, when restricted to $\Lambda_{n-1} \times \{0\}$, has null partial (lateral) derivatives at $(\sigma_0, \ldots, \sigma_n, 0)$. Assume that

$$\gamma_j = \int_0^\infty x^j \exp[\tau \cdot \mathbf{x}]dx, \quad 0 \le j \le n,$$

where $\tau \in \Lambda_n$. In particular $\tau_n < 0$, or $\tau_n = 0$, in which case, by the uniqueness of the maximum entropy solution $\tau = (\sigma_0, \ldots, \sigma_n, 0)$ and $\delta = \gamma_n(\max)$.

Assume that $\tau_n < 0$, so that

$$\gamma_j - \int_0^\infty x^j \exp[\tau \cdot \mathbf{x}]dx = \frac{\partial L}{\partial \lambda_j}(\tau) = 0,$$

where $\gamma_n = \delta$. Thus τ is a global maximum for the function L defined on Λ_n, and in particular $L(\tau) \ge L(\sigma_0, \ldots, \sigma_n, 0)$. By analyzing the restriction of the concave function L to the linear segment joining inside the set Λ_n the points τ and $(\sigma_0, \ldots, \sigma_n, 0)$ we infer $\left.\frac{\partial L(t\tau+(1-t)(\sigma_0,\ldots,\sigma_n,0))}{\partial t}\right|_{t=0} \le 0$, or in other terms

$$\delta \le \lambda_n(\max).$$

In conclusion, assuming that the finite moment sequence $(\gamma_0, \gamma_1, \ldots, \gamma_{n-1})$ is representable by a maximum entropy solution *of the same degree*, the extension $(\gamma_0, \gamma_1, \ldots, \gamma_{n-1}, \gamma_n)$ has the same property only if

$$\gamma_n(\min) \le \gamma_n \le \gamma_n(\max).$$

One step further, when investigating only the solvability of Stieltjes problem with data $(\gamma_0, \ldots, \gamma_n)$ by the maximum entropy solution without assumptions on the projected string $(\gamma_0, \ldots, \gamma_{n-1})$, the upper bound $\gamma_n(\max)$ may become infinite, see for details [30].

3.2. Recurrence relation for the moments of an exponential weight

The maximum entropy method for solving the truncated moment problem invites us to have a closer look at the full string of moments of an exponential of a polynomial weight. We enter below into the details of these computations, in the case of Stieltjes moment problem.

Fix an integer $n > 0$ and consider the polynomial

$$P(x) = \sigma_0 + \sigma_1 x + \cdots + \sigma_n x^n,$$

with real coefficients and $\sigma_n < 0$. Denote by

$$\gamma_k = \int_0^\infty x^k \exp[P(x)]dx, \quad k \geq 0,$$

the moments of the density $e^{P(x)}dx$. An integration by parts yields, for all $k \geq 0$:

$$\gamma_k = \int_0^\infty x^k e^P dx = \frac{x^{k+1}}{k+1} e^P \Big|_0^\infty - \int_0^\infty \frac{x^{k+1}}{k+1} P' e^P dx = -\int_0^\infty \frac{x^{k+1}}{k+1} P' e^P dx.$$

Hence, a finite difference equation relates every string of $n+1$ consecutive moments:

$$(k+1)\gamma_k + \sigma_1 \gamma_{k+1} + 2\sigma_2 \gamma_{k+2} + \cdots + n\sigma_n \gamma_{k+n} = 0, \quad k \geq 0. \tag{4}$$

Since $\sigma_n \neq 0$, we obtain the following simple observation.

Lemma 2. *Let $P(x)$ be a polynomial of degree n, with negative leading term. The moments of the density $e^{P(x)}dx$ are recurrently determined by (4) from the first n moments.*

Specifically, the linear dependence

$$\gamma_{k+n} = -\frac{k+1}{n\sigma_n}\gamma_k - \frac{\sigma_1}{n\sigma_n}\gamma_{k+1} - \cdots - \frac{(n-1)\sigma_{n-1}}{n\sigma_n}\gamma_{k+n-1},$$

holds. By changing the running index, we find for all $m > n$:

$$\gamma_m = -\frac{m-n+1}{n\sigma_n}\gamma_{m-n} - \frac{\sigma_1}{n\sigma_n}\gamma_{m-n+1} - \cdots - \frac{(n-1)\sigma_{n-1}}{n\sigma_n}\gamma_{m-1}.$$

Let $M' = \max_{i=1}^{n-1} \left|\frac{i\sigma_i}{n\sigma_n}\right|$ and $M = \max\left(M', \left|\frac{n-1}{n\sigma_n}\right|\right)$, so that

$$\max_{j \leq m} |\gamma_j| \leq \left(\frac{m}{|n\sigma_n|} + nM\right) \max_{j \leq m-1} |\gamma_j|.$$

Therefore there is a positive constant C and a positive integer N, such that

$$\max_{j \leq m} |\gamma_j| \leq C^m (m+N)!, \quad m \geq 0.$$

Consequently, Stirling's formula implies

$$\frac{\ln \max_{j \leq m} |\gamma_j|}{m} \leq C + \frac{(m+N)(\ln(m+N)-1)}{m} + \frac{\ln(2\pi(m+N))}{2m},$$

and in particular

$$\frac{\ln \max_{j \leq m} |\gamma_j|}{m} \leq C' + \frac{\ln(m+N)}{m},$$

where C' is a positive constant.

In conclusion, there is a positive constant γ, such that

$$\sum_{m=0}^{\infty} \frac{1}{|\gamma_m|^{1/m}} \geq \sum_{m=0}^{\infty} \frac{1}{[\max_{j\leq m} |\gamma_j|]^{1/m}} \geq \gamma \sum_{m=0}^{\infty} \frac{1}{m+N} = \infty.$$

According to Carleman's uniqueness criterion (see for instance [4]) we obtain the following result.

Theorem 3. *Let $P(x)$ be a non-constant polynomial with negative leading term. Then the moment problem with density $e^P dx$ is determined.*

We translate this statement for the reader who is not familiar with the terminology: if a positive measure μ on $[0, \infty)$ has the same moments as $e^P dx$, then $\mu = e^P dx$.

3.3. Existence

We have seen in the previous sections that not every truncated sequence of moments $(\gamma_0, \gamma_1, \ldots, \gamma_n)$ on the semi-axis can be achieved by the maximum entropy method, within the same degree. That is, it is not true that there always exists an admissible polynomial $P(x)$ of degree n or less, such that

$$\gamma_k = \int_0^{\infty} x^k e^{P(x)} dx, \quad k \leq n. \tag{5}$$

To give the simplest example, consider the sequence

$$\gamma_0 = 1, \qquad \gamma_1 = \gamma_2 = \cdots = \gamma_n = 0.$$

Obviously, the Dirac mass δ_0 has these very moments. However, there is no polynomial P, of any degree, such that

$$0 = \gamma_1 = \int_0^{\infty} x e^{P(x)} dx.$$

Simply because the integrand is non-negative and non-null on the interval of integration.

Our study is motivated by the need to solve this pathology. In the following sections we indicate a method to overcame the limitation of the maximum entropy method to absolutely continuous measures. Along the same lines, some recent works proposed different regularizations, see for instance [15].

4. Single variable: conditioning using the Cauchy transform

The recent works of Junk [30] and Hauck et al. [23] clarified which positive densities ρ are appropriate for the maximum entropy reconstruction method. A thorough analysis of the convex structure of the truncated moment set of these distributions, e.g., extreme points, facets, was carried out in the cited works, with significant applications for the kinetic theory of gases. In particular, singular measures are especially poor candidates for maximum entropy reconstruction, as seen from the example in Section 3.3. In an attempt to enlarge the class of measures for which such well established reconstruction methods work, we propose a regularization procedure which will produce an admissible input for the entropy optimization procedure.

The goal of our procedure is to reconstruct a possibly singular measure μ by transforming it to a continuous measure $\phi(t)dt$, whose density ϕ we term *the phase function*. The entire measure reconstruction procedure can broken down into three steps:

1. regularization based on moment data of μ,
2. density reconstruction (using entropy optimization) of ϕ,
3. inversion, i.e., recovering a measure $\mu^* \approx \mu$, from point-wise knowledge of ϕ.

We stress here that it is not our aim to improve on the density reconstruction procedure, i.e., the entropy optimization, itself. Rather we focus on moving the density reconstruction where it can be performed with assured convergence, by inserting the regularization and inversion steps. It could be very well possible that other density reconstruction methods, e.g., basis pursuit, wavelet-based reconstruction, could be used instead of the maximum entropy for the general reconstruction problem, however, we do not explore these options here.

In this section, we first focus on measures whose support lies in a one-dimensional space. In this case, the entire procedure is based on a simple idea of Markov [3], widely used in function theory, employing Cauchy transforms. Cauchy transforms serve as an analytic tool to study complex generating functions of the moment sequences. The regularization step is based on representation theorems for the generating function of the moment sequence, while the inversion step is grounded in Plemelj–Sokhotski formulas, which can be used to reconstruct the densities on the original domain. When the domain is one-dimensional, Plemelj–Sokhotski formulas can be formulated through a Hilbert transform, which is easily evaluated numerically. Therefore, such a reconstruction results in an algorithm that can easily be implemented in a computer code.

In Section 4.1, we first give the procedure for measures with arbitrary supports in $\mathbb{R}$, based on power, i.e., monomial, moments of the measure μ as input data. If the support of measure is contained in a compact interval, we can employ trigonometric moments instead, which are preferred numerically to power moments. The regularization procedure for compact supports is developed in Section 4.2, and is somewhat more technical than for the unbounded case, yet the spirit is the same. Based on insights for one-dimensional domains, in Section 5 we discuss how the procedure might be extended to measures supported in $\mathbb{R}^d$.

4.1. Unbounded support

Define the Cauchy transform of a measure μ, with support in $\mathbb{R}$, as

$$\mathcal{C}\mu(z) = \int_{\mathbb{R}} \frac{d\mu(x)}{x - z}. \tag{6}$$

Markov's observation is the following: assuming all integrals exist, the Cauchy transform of a *positive* measure on the line is of Nevanlinna class, i.e., it has a positive imaginary part in the upper-half plane:

$$\frac{\mathcal{C}\mu(z) - \mathcal{C}\mu(\overline{z})}{2i} = \int_{\mathbb{R}} \frac{\Im z d\mu(x)}{|x - z|^2} > 0, \quad \Im z > 0.$$

Hence the *phase* $\Im \ln[\mathcal{C}\mu(z)]$ is a harmonic function in the upper half-plane, uniformly bounded from below by zero and from above by π. The boundary values along the real line of $\Im \ln[\mathcal{C}\mu(z)]$ produce an integrable, positive and bounded density ϕ, satisfying:

$$1 + \mathcal{C}\mu(z) = \exp \int_{\mathbb{R}} \frac{\phi(x)dx}{x - z}, \quad \Im z > 0 \tag{7}$$

i.e.

$$1 + \mathcal{C}\mu(z) = \exp \mathcal{C}\phi(z), \quad \Im z > 0, \tag{8}$$

where we slightly abuse the notation when we use $\mathcal{C}\phi$. The dictionary between properties of μ and density ϕ was established by Aronszajn and Donoghue [7]. The most important, of course, is the existence and boundedness of ϕ. As ϕ is bounded even if μ is singular, we consider $\phi(t)dt$ to be a *regularization* of $d\mu$.

Practical benefit of this expression comes from the ability to use it without knowing the closed-form expressions for measures involved. The Cauchy transform is the (complex) generating function for moments of μ, i.e., its expansion at $z = \infty$ is given by

$$(\mathcal{C}\mu)(z) = -\sum_{n=0}^{\infty} \frac{a_\mu(n)}{z^{n+1}},$$

where $a_\mu(n) \triangleq \int_{\mathbb{R}} t^n d\mu(t)$, and $a_\phi(n)$ defined analogously.[1] Solving for $\mathcal{C}\phi$ and using the series expansion $\ln(1+z) = -\sum_{n=1}^{\infty}(-1)^n z^n/n$ yields the following equality between power series:

$$\mathcal{C}\phi(z) = \ln[1 + \mathcal{C}\mu(z)] = -\sum_{k=1}^{\infty} \frac{1}{k}[-\mathcal{C}\mu(z)]^k$$

$$\sum_{n=0}^{\infty} \frac{a_\phi(n)}{z^{n+1}} = \sum_{k=1}^{\infty} \frac{1}{k}\left[\sum_{n=0}^{\infty} \frac{a_\mu(n)}{z^{n+1}}\right]^k.$$

The Miller–Nakos Theorem [40], whose complete proof we bring in the Appendix A, gives a recursion for evaluation of moments $a_\phi(n)$ from moments $a_\mu(k)$ for $k = 0, \ldots, n$,

$$a_\phi(N) = \sum_{k=1}^{N} \frac{1}{k}[S_N(z)]_N^k, \tag{9}$$

where $[S_N(z)]_N^k$ indicates the coefficient next to $z^{-(N+1)}$, in the k-th power of the truncation $S_N(z) = \sum_{n=0}^{N} a_\mu(n) z^{-(n+1)}$ of the generating power series. Such a triangular property is essential for practical problems: we will typically have access only to truncated moment data and we do not wish to establish any a priori ansatz, especially not $a_\mu(n) = 0$ for $n > N$.

At this point, we have set up moment data such that most density reconstruction procedures apply: density ϕ is bounded and compactly supported, making it possible to reconstruct it using entropy optimization described in Section 3. Such a procedure produces an approximant

$$\phi^*(x) = \exp \sum_{k=0}^{N} \alpha_k x^k \tag{10}$$

that converges to density ϕ as the number of available moments N increases.

The inversion step describes how the point-wise knowledge of approximant $\phi^* \approx \phi$ is used to compute an absolutely continuous measure μ^* that approximates the original measure μ. In this paper we do not claim to obtain quantitative convergence results on $\mu^* \to \mu$, especially when μ

[1] The non-linear transform of the moment sequence was exploited in the theory of the phase shift of perturbed spectra in quantum mechanics, see [10,43].

is a singular measure, however, we stress that, for singular measures, a classical reconstruction procedure like entropy optimization might not produce any results. Therefore, we view our results in this paper as a starting point for further investigations of approximation of singular measures.

To a smooth entropy optimizer ϕ^* corresponds a measure μ^* with a density $\rho = d\mu^*/dx$, which we use to approximate the original measure μ. The lynchpin of the inversion procedure, i.e., evaluation of ρ from knowledge of ϕ^*, is the existence of boundary limits $\lim_{\epsilon\to 0} \mathcal{C}\mu(x \pm i\epsilon)$, for $\epsilon > 0$. The limits exist independently pointwise, and, assuming that $\rho \in L^1(\mathbb{R})$ is of Hölder-class, the Plemelj–Sokhotski formulas, e.g., [28, Section 14.11] or [31, Section 3.7], establish that it is possible to evaluate ρ pointwise from Cauchy transforms of μ as

$$\rho(x) = \frac{1}{2\pi i} \lim_{\epsilon\downarrow 0} [\mathcal{C}\mu^*(x + i\epsilon) - \mathcal{C}\mu^*(x - i\epsilon)].$$

As limits exist independently, and exp is analytic, we can formulate them in terms of analogous limits for Cauchy transforms of the phase function $\mathcal{C}\phi^*$, i.e. by (8),

$$\begin{aligned}\rho(x) &= \frac{1}{2\pi i} \lim_{\epsilon\downarrow 0} [\exp \mathcal{C}\phi^*(x + i\epsilon) - \exp \mathcal{C}\phi^*(x - i\epsilon)] \\ &= \frac{1}{2\pi i} \left[\exp \lim_{\epsilon\downarrow 0} \mathcal{C}\phi^*(x + i\epsilon) - \exp \lim_{\epsilon\downarrow 0} \mathcal{C}\phi^*(x - i\epsilon) \right].\end{aligned}$$

Moreover, the Plemelj–Sokhotski formulas provide explicit expressions for each limit:

$$\lim_{\epsilon\downarrow 0} \frac{1}{2\pi i} \mathcal{C}\phi^*(x \pm i\epsilon) = \pm\frac{1}{2}\phi^*(x) + \frac{i}{2}\mathcal{H}\phi^*(x), \tag{11}$$

where the Hilbert transform is

$$\mathcal{H}\phi^*(x) = \frac{1}{\pi} \fint \frac{\phi^*(t)dt}{t - x}. \tag{12}$$

It follows then that the expression for ρ is given by:

$$\rho(x) = \frac{1}{\pi} \exp\left[-\pi \mathcal{H}\phi^*(x)\right] \sin \pi\phi^*(x). \tag{13}$$

This formula connects density $\rho = d\mu/dx$ with the phase density function ϕ, or, in the case of moment closure by entropy optimization, a smooth approximant ϕ^* to the phase density function ϕ.

The formula (13) is numerically practical: the entropy optimization provides us with a closed formula for ϕ^*, while its Hilbert transform is easily numerically evaluated via the Fast Fourier Transform algorithm. Therefore, here we have obtained a practical inversion formula for approximating a singular measure μ via an absolutely continuous measure μ^* with density $\rho = d\mu^*/dx$.

4.2. Compact support

When the measure μ is supported on a known compact interval, we can use trigonometric moments, instead of power moments, in the process given above. The resulting process is more numerically robust, as trigonometric functions are orthonormal and bounded as a family, unlike the family of monomials on an arbitrary interval.

Let μ be a measure on the interval $\Delta = [-\pi, \pi)$ that induces the measure $\breve{\mu}$ on the boundary $\partial\mathbb{D}$ of the unit disk $\mathbb{D} \subset \mathbb{C}$. A known relation is then $d\mu(\theta) = -i\bar{\zeta}d\breve{\mu}(\zeta)$, for $\zeta = e^{i\theta} \in \partial\mathbb{D}$. Define the circular Cauchy transformation

$$\mathcal{K}\mu(z) \triangleq \frac{1}{2\pi}\int_{\partial\mathbb{D}} \frac{d\breve{\mu}(\zeta)}{\zeta - z},$$

for $z \notin \partial\mathbb{D}$. Using the equivalent arc-length formulation clarifies the difference between

$$\mathcal{K}\mu(z) = \frac{i}{2\pi}\int_{-\pi}^{\pi} \frac{d\mu(\theta)}{1 - e^{-i\theta}z}$$

and the Cauchy transform on the line $\mathcal{C}\mu(z) = \int_{-\pi}^{\pi} d\mu(x)/(x - z)$, cf. (6).

The function $\mathcal{K}\mu$ is holomorphic inside $\operatorname{int}\mathbb{D}$, $\mathcal{K}\mu \in \mathcal{O}(\mathbb{D})$, where it has the Taylor expansion

$$\mathcal{K}\mu(z) = i\sum_{k=0}^{\infty} \tau_\mu(k) z^k,$$

with complex trigonometric moments

$$\tau_\mu(k) \triangleq \frac{1}{2\pi}\int_{\partial\mathbb{D}} \bar{\zeta}^k \frac{d\breve{\mu}(\zeta)}{i\zeta} \tag{14}$$

serving as coefficients.

The imaginary part $\Im\mathcal{K}\mu(z)$, is positive for positive measures, as the imaginary part of the kernel is

$$\frac{1}{4\pi i}\left(\frac{i}{1 - e^{-i\theta}z} - \frac{-i}{1 - e^{i\theta}\bar{z}}\right) = \frac{1 - \Re(e^{-i\theta}z)}{2\pi\left|1 - e^{-i\theta}z\right|^2},$$

and $\left|ze^{-i\theta}\right| < 1$ when $z \in \operatorname{int}\mathbb{D}$. Consequently, the argument of $\mathcal{K}\mu(z)$, with the appropriately chosen branch of the logarithm,

$$F(z) \triangleq -i\ln\mathcal{K}\mu(z) \in \mathcal{O}(\mathbb{D}), \tag{15}$$

is of Caratheodory class: it is a positive function, with a bounded real part $\Re F(z) \in [0, \pi]$, which corresponds to the bounded angle of $\mathcal{K}\mu(z)$.

The following classical theorem allows us to obtain a representation of Caratheodory class functions in terms of bounded densities on a circle (e.g. [27], Section 12.10):

Theorem 4 (*Riesz–Herglotz*)**.** *Let $F \in \mathcal{O}(\mathbb{D})$ be such that $\Re F(z) \in [0, c]$ for some fixed $c > 0$. Then there exists a function $\breve{\phi} \in L^1(\partial\mathbb{D})$ for which*

$$F(z) = i\Im F(0) + \mathcal{P}\phi(z),$$

where $\breve{\phi}(\zeta) \in [0, c]$ pointwise and

$$\mathcal{P}\phi(z) \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z}\phi(\theta)d\theta,$$

is the Poisson integral of $\phi(\theta) \equiv \breve{\phi}(e^{i\theta})$.

The Poisson integral $\mathcal{P}\phi$ can be rewritten in terms of the circular Cauchy transform $\mathcal{K}\phi$:

$$\mathcal{P}\phi(z) = \frac{1}{2\pi}\int_{\partial\mathbb{D}} \frac{\zeta+z}{\zeta-z}\frac{\phi(\zeta)d\zeta}{i\zeta} = -\tau_\phi(0) - i2\mathcal{K}\phi(z),$$

where, again, $\tau_\phi(k)$ are trigonometric moments of ϕ, defined analogously to (14). The Riesz–Herglotz formula then reads

$$F(z) = -i2\mathcal{K}\phi(z) - \tau_\phi(0) + i\Im F(0).$$

We can compute the constants in the formula by evaluating it at $z = 0$ and comparing it to evaluation of the definition (15) at the same point:

$$F(0) = i\Im F(0) - \tau_\phi(0) - i2[i\tau_\phi(0)] = \tau_\phi(0) + i\Im F(0)$$
$$F(0) = -i\ln i\tau_\mu(0) = \frac{\pi}{2} - i\ln\tau_\mu(0),$$

concluding that

$$\tau_\phi(0) = \frac{\pi}{2}, \qquad \Im F(0) = -\ln\tau_\mu(0).$$

Substituting these constants into the Riesz–Herglotz formula, and using the definition of $F(z)$, we obtain the exponential representation of $\mathcal{K}\mu(z)$:

$$\mathcal{K}\mu(z) = -i\tau_\mu(0)\exp[2\mathcal{K}\phi(z)]. \tag{16}$$

To compute the moments of ϕ, we relate the Taylor expansions of the functions above, and use $\tau_\phi(0) = \pi/2$ to obtain

$$1 + \sum_{n=1}^{\infty}\hat{\tau}_\mu(n)z^n = \exp\left[2i\sum_{n=1}^{\infty}\tau_\phi(n)z^n\right],$$

where $\hat{\tau}_\mu(n) \triangleq \tau_\mu(n)/\tau_\mu(0)$. As before, we use the expansion $\ln(1+z) = -\sum_{n=1}^{\infty}(-1)^n z^n/n$ to relate the series through expression

$$\sum_{k=1}^{\infty}\tau_\phi(k)z^k = \frac{i}{2}\sum_{k=1}^{\infty}\frac{(-1)^k}{k}\left[\sum_{n=1}^{\infty}\hat{\tau}_\mu(n)z^n\right]^k.$$

A finite number M of trigonometric moments $\tau_\phi(k)$ can then be computed using the Miller–Nakos algorithm (see Appendix A) when M moments $\tau_\mu(k)$ are known.

To invert the procedure, we assume that to approximate ϕ, we are given a smooth density $\phi^* : [-\pi,\pi] \to \mathbb{R}$, which corresponds to a continuous μ^* with density $\rho : [-\pi,\pi] \to \mathbb{R}$, i.e., $d\mu^*(\theta) = \rho(\theta)d\theta$. Density ρ can be evaluated point-wise using Plemelj–Sokhotski formulas (see Dynkin's chapter, Section 6 in [16]), which evaluate non-tangential limits $\text{i-lim}_{\xi\to z}\mathcal{K}\mu^*(\xi)$ and $\text{e-lim}_{\xi\to z}\mathcal{K}\mu^*(\xi)$ at $z \in \partial\mathbb{D}$, with the argument in domains $\xi \in \operatorname{int}\mathbb{D}$ and $\xi \in \mathbb{C}/\mathbb{D}$, respectively. The A. Calderón's theorem asserts existence of such limits for $\breve{\phi}^* \in L^1(\partial\mathbb{D})$.

Due to analyticity of exp in (16), we can evaluate the non-tangential limits of $\mathcal{K}\mu^*$, in terms of non-tangential limits $\mathcal{K}\phi^*$ of $\breve{\phi}^*(\zeta)d\zeta$[2]

$$\text{i-}\lim_{\xi\to z}\mathcal{K}\mu^*(\xi) = -i\tau_\mu(0)\exp[2\,\text{i-}\lim_{\xi\to z}\mathcal{K}\phi^*(\xi)]$$

[2] We slightly abuse the notation when we use $\mathcal{K}\phi^*$.

$$\operatorname*{e-lim}_{\xi\to z} \mathcal{K}\mu^*(\xi) = -i\tau_\mu(0)\exp[2\operatorname*{e-lim}_{\xi\to z}\mathcal{K}\phi^*(\xi)].$$

Privalov's Lemma establishes that the non-tangential limits satisfy Plemelj–Sokhotski formulas for $z \in \partial\mathbb{D}$:

$$\operatorname*{i-lim}_{\xi\to z} \mathcal{K}\phi^*(\xi) = \mathcal{Q}\phi^*(z) + \frac{i}{2}\breve{\phi}^*(z),$$
$$\operatorname*{e-lim}_{\xi\to z} \mathcal{K}\phi^*(\xi) = \mathcal{Q}\phi^*(z) - \frac{i}{2}\breve{\phi}^*(z),$$

where $\mathcal{Q}$ indicates the singular integral

$$\mathcal{Q}\phi^*(z) \triangleq \frac{1}{2\pi}\fint_{\partial\mathbb{D}} \frac{\breve{\phi}^*(\zeta)d\zeta}{\zeta - z},$$

with analogous expressions holding for $\mathcal{K}\mu^*(z)$, with density $\rho(\theta) = d\mu^*/d\theta$ instead of ϕ^*. Therefore, to evaluate $\breve{\rho}(z)$ on $\partial\mathbb{D}$ we seek the difference between the non-tangential limits:

$$\begin{aligned}\breve{\rho}(z) &= -\tau_\mu(0)\left\{\exp[2\operatorname*{i-lim}_{\xi\to z}\mathcal{K}\phi^*(\xi)] - \exp[2\operatorname*{e-lim}_{\xi\to z}\mathcal{K}\phi^*(\xi)]\right\}\\ &= -i2\tau_\mu(0)\exp[2\mathcal{Q}\phi^*(z)]\sin\phi^*(z).\end{aligned}$$

The singular integral $\mathcal{Q}\phi^*(z)$ can be evaluated on $z \equiv e^{i\theta}$ using the circular Hilbert transform of ϕ^*:

$$\begin{aligned}\mathcal{Q}\phi^*(z) &= \frac{1}{2\pi}\fint_{\partial\mathbb{D}} \frac{\breve{\phi}^*(\zeta)d\zeta}{\zeta - z} = \frac{i}{4\pi}\fint_{\partial\mathbb{D}} \frac{\zeta + z}{\zeta - z}\frac{\breve{\phi}^*(\zeta)d\zeta}{i\zeta} + \frac{i}{2}\tau_\phi(0)\\ &= \frac{i}{4\pi}\fint_{-\pi}^{\pi} \frac{e^{i\sigma} + e^{i\theta}}{e^{i\sigma} - e^{i\theta}}\phi^*(\sigma)d\sigma + \frac{i\pi}{4}\\ &= \frac{1}{4\pi}\fint_{-\pi}^{\pi} \cot\frac{\sigma - \theta}{2}\phi^*(\sigma)d\sigma + \frac{i\pi}{4}\\ &= \frac{1}{2}\mathcal{H}\phi^*(\theta) + \frac{i\pi}{4},\end{aligned}$$

where $z \equiv e^{i\theta}$. The circular Hilbert transform is, by one convention,

$$\mathcal{H}\phi^*(\theta) \triangleq \frac{1}{2\pi}\fint_{-\pi}^{\pi} \cot\frac{\sigma - \theta}{2}\phi^*(\sigma)d\sigma.$$

Finally, substituting this expression into $\breve{\rho}(\zeta) \equiv \rho(\theta)$, we get the evaluation of the density $\rho(\theta)$ as

$$\rho(\theta) = 2\tau_\mu(0)\exp[\mathcal{H}\phi^*(\theta)]\sin\phi^*(\theta).$$

Practically, Hilbert transform is easily evaluated on a fixed grid using numerical Fourier Transform, e.g., FFT, so this formula can be employed when we have access to the evaluation of ϕ^* on a fixed grid in $[-\pi, \pi]$.

5. Several variables: conditioning using complex Fantappiè transforms

The reminder of this paper deals with generalization of the regularization procedure to measures of several variables. To do so, we will replace Cauchy transform with a Fantappiè

transform, which is usually defined as a real integral transform, relying on two sources: the harmonic analysis on tube domains over convex cones [20] and the complete monotonicity results, à la Bernstein, characterizing the Laplace and Fantappiè transforms of positive measures on convex cones [24]. For expository material on Fantappiè transform, see [6, Section 3].

We propose three different ways to complexify it, in order to use the general Riesz–Herglotz representation theory and obtain analogs to the phase function ϕ. The choice of the complexification procedure is based on a trade-off: presently we are able to obtain either theoretically general results with little practical value, or practically useful results which do not allow for as much theoretical breadth. We expect that the future research will bridge this gap between theory and computation.

Take a solid, acute, closed convex cone $\Gamma \subset \mathbb{R}^d$, and its associated *polar cone*

$$\Gamma^* \triangleq \{x \in \mathbb{R}^d;\ \omega \cdot x \geq 0,\ \omega \in \Gamma\}.$$

The cone Γ^* will carry the support of the measure μ, while Γ will play the role of the "frequency parameter", to use the language of signal processing and applied Fourier/Laplace analysis. The following characterization of the real-valued Fantappiè transform is due to [24]:

Theorem 5 (*Henkin–Shananin*)**.** *A function $\Phi : (0, \infty) \times \Gamma \to \mathbb{R}$ is the Fantappiè transform*

$$\Phi(\omega_0, \omega) = \int_{\Gamma^*} \frac{d\mu(x)}{\omega_0 + \omega \cdot x}, \tag{17}$$

of a positive measure μ supported by Γ^* if and only if Φ is (i) *continuous,* (ii) *completely monotonic,*[3] *and* (iii) *homogeneous of degree -1, i.e.,*

$$\Phi(\lambda\omega_0, \lambda\omega) = \lambda^{-1}\Phi(\omega_0, \omega), \quad \omega_0 > 0,\ \omega \in \Gamma,\ \lambda > 0.$$

To extend the Fantappiè transform to complex domains, one has a choice of complexifying the offset parameter ω_0, the normal parameter ω, or both. We start with the full generality in Section 5.1, complexifying both ω_0 and ω to tube domains, and develop the full regularization procedure, at a cost of providing no inversion formulas. Next, we constrain $\omega_0 = 1$ in Section 5.2, obtaining a restricted tube domain, to provide some practical regularization formulas for measures on familiar compact domains, which stand in direct analogy to one-dimensional problem. A future research direction could explore Clifford algebras as a setting for generalization of the Plemelj–Sokhotski formulas, which were used to complete the inversion process in the one-dimensional case in Section 4. A shorter, perhaps more immediately practical, procedure is given in Section 5.3, where we treat ω as a parameter, complexifying only ω_0. As a consequence, we recover a single variable procedure at each value of ω, which we term *the partial Fantappiè transform*. The family of solutions, parametrized by ω, could be used in a tomographic procedure to recover an approximant to the original measure μ.

5.1. Unbounded supports and tube domains

The Fantappiè integral transform extends analytically to the complex domain:

$$\Phi(u_0, u) \triangleq \int_{\Gamma^*} \frac{d\mu(x)}{u_0 + u \cdot x}, \quad \Re u \in \operatorname{int}\Gamma,\ \Re u_0 > 0,$$

[3] A function Φ is *completely monotonic* if it satisfies inequalities $(-1)^k D_{\xi_1} \cdots D_{\xi_k} \Phi(p) \geq 0,\ \forall k \geq 0,\ p \in \operatorname{int}\Gamma$, where D_{ξ_k} are partial derivatives along coordinates $\xi_1, \ldots, \xi_k \in \Gamma$.

retaining, by definition, homogeneity of degree -1 in the complex argument $(u_0, u) \in \mathbb{C} \times \mathbb{C}^d$. Let $\Omega = (0, \infty) \times \operatorname{int} \Gamma \subset \mathbb{R}^{d+1}$ be the interior of the domain of continuity for the real Fantappiè transform.

The associated *tube domain* is the set $T_\Omega = \Sigma + i\Omega$, where $\Sigma = \mathbb{R}^{d+1}$. Due to the different role played by the first axis, we denote elements by $(z_0, z) = (\sigma_0 + i\omega_0, \sigma + i\omega) \in T_\Omega$. With this notation[4] the domain of analyticity can be written as $-iT_\Omega$, i.e., $\Phi \in \mathcal{O}(-iT_\Omega)$. For clarity, we will use u to denote elements of $-iT_\Omega$, and z for elements of T_Ω, with the obvious change of coordinates $u = -iz = \omega - i\sigma$, when $z = \sigma + i\omega$.

Notice that when $(u_0, u) \in -iT_\Omega$,

$$\Re \Phi(u_0, u) = \int_{\Gamma^*} \frac{\omega_0 + \omega \cdot x}{|u_0 + u \cdot x|^2} d\mu(x) > 0,$$

following from the definition of the polar cone Γ^*. Therefore, function $i\,\Phi(u_0, u)$ is analytic, and has a positive imaginary part. It follows that its complex phase $\ln i\,\Phi(u_0, u)$ is well defined on $-iT_\Omega$, with the property

$$\Im \ln i\,\Phi(u_0, u) \in (0, \pi), \quad (u_0, u) \in -iT_\Omega.$$

Converting this expression to the tube domain, define $F(z_0, z)$ on the tube domain T_Ω by setting

$$F(z_0, z) \triangleq -i \ln i\,\Phi(-iz_0, -iz),$$

with a further simplification

$$F(z_0, z) = -i \ln[-\Phi(z_0, z)], \tag{18}$$

due to homogeneity of Φ, or

$$\Phi(z_0, z) = -\exp i F(z_0, z). \tag{19}$$

Defined this way, the function $F(z_0, z)$ is analytic on T_Ω, and satisfies $\Re F(z_0, z) \in [0, \pi]$.

These properties make it possible to reveal the structure of functions F via a straightforward generalization of Riesz–Herglotz formula for analytic functions of positive real part (see Theorem 4), which is our next goal.

Let $\Omega \subset \mathbb{R}^d$ be an open, acute and solid convex cone, with associated tube domain $T_\Omega = \mathbb{R}^d + i\Omega$. The Hardy space $H^2(T_\Omega)$ is defined as the space of analytic functions $F : T_\Omega \longrightarrow \mathbb{C}$, such that

$$|F|^2 = \sup_{\omega \in \Omega} \int_{\mathbb{R}^d} |F(\sigma + i\omega)|^2 \, d\sigma < \infty.$$

By a celebrated theorem of Paley and Wiener, $H^2(T_\Omega)$ is the space of Fourier–Laplace transforms of square integrable functions defined on the polar cone.

The following result characterizes real Fourier–Laplace transforms [24]:

Theorem 6 (*Bernstein, Bochner, Gilbert*)**.** *A function $F : \Omega \longrightarrow \mathbb{R}$ is the Laplace transform*

$$F(\omega) = \int_{\Omega^*} e^{-\omega \cdot x} d\mu(x),$$

[4] Such a choice conforms with the existing conventions of harmonic analysis, and we apologize in advance for all the resulting multiplicative imaginary unities. Fantappiè transforms in the later sections simplify this convention.

of a positive measure μ supported by Ω^* if and only if F is continuous on Ω and of class C^∞ and completely monotonic in the interior int Ω.

The extension from the cone Ω to tube domain T_Ω, for $f \in L^2(\Omega^*, dx)$, is given by

$$F(z) = \frac{1}{\sqrt{(2\pi)^d}} \int_{\Omega^*} e^{iz\cdot x} f(x)dx,$$

for $z \in T_\Omega$. Function F then belongs to $H^2(T_\Omega)$ and the map $f \mapsto F$ is an isometric isomorphism between the two Hilbert spaces. For a proof and an overview of the theory of Hardy spaces on tube domains see [20].

The reproducing kernel of the Hardy space, also known as *Szegö's kernel* is

$$S(z, w) = \frac{1}{(2\pi)^d} \int_{\Omega^*} e^{i(z-\overline{w})\cdot x} dx,$$

for $z, w \in T_\Omega$. Remark the homogeneity property:

$$S(\lambda z, \lambda w) = \lambda^{-d} S(z, w), \quad \lambda > 0.$$

The reproducing property has the following form: if $F \in H^2(T_\Omega)$, then the boundary limits, still denoted by F, satisfy

$$F(\sigma) = \lim_{\omega \to 0} F(\sigma + i\omega),$$

where the limit exists in $L^2(\mathbb{R}^d)$ and

$$F(z) = \int_{\mathbb{R}^d} S(z, \sigma) F(\sigma) d\sigma.$$

We focus next on functions $F \in A(T_\Omega)$ which are analytic in T_Ω and uniformly bounded and continuous on its closure. Then the function $z \mapsto S(z, w)F(z)$ belongs to $H^2(T_\Omega)$, for every $w \in T_\Omega$, and

$$S(z, w)F(z) = \int_{\mathbb{R}^d} S(z, \sigma) S(\sigma, w) F(\sigma) d\sigma,$$

and by complex conjugation

$$S(z, w)\overline{F(w)} = \int_{\mathbb{R}^d} S(z, \sigma) S(\sigma, w) \overline{F(\sigma)} d\sigma.$$

By adding the two identities we obtain

$$S(z, w)\frac{F(z) + \overline{F(w)}}{2} = \int_{\mathbb{R}^d} S(z, \sigma) S(\sigma, w) \Re F(\sigma) d\sigma.$$

The restriction to the diagonal of the above formula yields

$$\Re F(z) = \int_{\mathbb{R}^d} P(z, \sigma) \Re F(\sigma) d\sigma,$$

where

$$P(z, \sigma) = \frac{|S(z, \sigma)|^2}{S(z, z)}, \quad z \in T_\Omega,\ \sigma \in \mathbb{R}^d,$$

is Poisson's kernel. Again, see [20] for full details.

Fix a point $\alpha \in i\Omega$, and subtract the identities

$$S(z,\alpha)[F(z)+\overline{F(\alpha)}] = \int_{\mathbb{R}^d} 2S(z,\sigma)S(\sigma,\alpha)\Re F(\sigma)d\sigma,$$
$$S(z,\alpha)\frac{F(\alpha)+\overline{F(\alpha)}}{2} = \int_{\mathbb{R}^d} \frac{|S(\sigma,\alpha)|^2\, S(z,\alpha)}{S(\alpha,\alpha)}\Re F(\sigma)d\sigma.$$

One finds

$$S(z,\alpha)[F(z)-i\Im F(\alpha)] = \int_{\mathbb{R}^d}\left[2S(z,\sigma)S(\sigma,\alpha) - \frac{|S(\sigma,\alpha)|^2\, S(z,\alpha)}{S(\alpha,\alpha)}\right]\Re F(\sigma)d\sigma,$$

a tube domain analogue of the Schwarz formula, which relates the values of an analytic function (in the disk) to the boundary values of its real part. We call, by way of natural analogy to the similar integral kernel for the disk,

$$H(z,w;\alpha) = 2\frac{S(z,w)S(w,\alpha)}{S(z,\alpha)} - \frac{|S(w,\alpha)|^2}{S(\alpha,\alpha)},$$

the *Herglotz kernel* associated to the tube domain T_Ω, with $z, w \in T_\Omega$, $\alpha \in i\Omega$.

In particular cases one can obtain from here an integral representation of all analytic functions in T_Ω possessing positive real part, what it is customarily called the Riesz–Herglotz formula, see for details [2]. Fortunately, our aim is more modest, having to deal only with the Fantappiè transforms of measures appearing in the previous section.

From now on, we return to the convex cone $\Omega = (0,\infty)\times \operatorname{int}\Gamma \subset \mathbb{R}^{d+1}$ appearing in definition of the real Fantappiè transform (17). Specifically, the analytic function

$$F(z_0,z) = -i\ln[-\Phi(z_0,z)], \quad (z_0,z)\in T_\Omega = \mathbb{R}^{d+1}+i\Omega$$

satisfies

$$0 < \Re F(z_0,z) < \pi, \quad (z_0,z)\in T_\Omega$$

and, due to the homogeneity of Φ it growths logarithmically along rays contained in T_Ω:

$$|F(\lambda z_0,\lambda z)| \le |F(z_0,z)| + |\ln\lambda|, \quad (z_0,z)\in T_\Omega,\ \lambda>0.$$

Fix a point $\alpha \in i\Omega$ and consider the translated functions

$$F_\epsilon(z_0,z) = F[(z_0,z)+\epsilon\alpha],$$

so that they are analytic on the closure of T_Ω and of logarithmic growth along rays. Due to the homogeneity of degree $-(d+1)$ of the reproducing kernel S of T_Ω, we deduce that for every $\epsilon>0$, the function $F_\epsilon S \in H^2(T_\Omega)$, and in view of the computations above:

$$F_\epsilon(\zeta) = i\Im F_\epsilon(\alpha) + \int_{\mathbb{R}^{d+1}} H(\zeta,\sigma;\alpha)\Re F_\epsilon(\sigma)d\sigma, \quad \zeta=(z_0,z)\in T_\Omega,\ S(\zeta,\alpha)\neq 0.$$

Note that $\Re F_\epsilon \in (0,\pi)$ on T_Ω, as restriction of the original function to a subset of the tube domain. By passing with ϵ to zero and to a weak-* limit in $L^\infty(\mathbb{R}^{d+1})$ we obtain a function $\phi \in L^\infty(\mathbb{R}^{d+1})$, $0\le\phi\le\pi$, a.e., representing F as follows:

$$F(\zeta) = i\Im F(\alpha) + \int_{\mathbb{R}^{d+1}} H(\zeta,\sigma;\alpha)\phi(\sigma)d\sigma, \quad \zeta=(z_0,z)\in T_\Omega,\ S(\zeta,\alpha)\neq 0.$$

In all instances of interest Szegö's kernel $S(\zeta,\alpha)$ does not vanish at all points $\zeta,\alpha \in T_\Omega$, producing a genuine integral representation of F.

We collect the above remarks into a formal statement.

Proposition 7. *Let $\Gamma \subset \mathbb{R}^d$ be a closed, solid and acute convex cone and let μ be a finite mass positive measure supported by the polar cone Γ^*. The Fantappiè transform of the measure μ admits the exponential representation:*

$$\int_{\Gamma^*} \frac{d\mu(x)}{z_0 + z\cdot x} = -\exp[iF(z_0,z)], \quad (z_0,z) \in T_\Omega,$$

where $\Omega = (0,\infty)\times \operatorname{int}\Gamma$. In its turn, the analytic function F admits the integral representation

$$F(\zeta) = iC + \int_{\mathbb{R}^{d+1}} H(\zeta,\sigma;\alpha)\phi(\sigma)d\sigma, \quad \zeta = (z_0,z) \in T_\Omega,$$

where $\phi : \mathbb{R}^{d+1} \longrightarrow [0,\pi]$ is a measurable function and $C \in \mathbb{R}$ is a real constant.

We assume in the above statement that $\alpha \in i\Omega$ is fixed and $S(\zeta,\alpha) \neq 0$. In this sense the, possibly singular, measure μ is regularized by the absolutely continuous measure $\phi(\sigma)d\sigma$. The integral kernel $H(\zeta,\sigma;\alpha)$ is of little practical use in its full generality. However, in particular cases, to be discussed in the rest of the article, the preceding Markov exponential transform regularization becomes more accessible.

One question which naturally arises in the above statement is: is it possible to characterize the bounded densities ϕ appearing in the integral representation of an analytic function $F \in \mathcal{O}(T_\Omega)$ that satisfies $0 \le \Re F \le 1$? The answer is yes, but the conditions imposed on ϕ are not friendly. They were discovered a long time ago, in the case of the polydisk [33] and more general symmetric domains [2]. We merely indicate these conditions in the case of non-vanishing Szegö kernel and sketch the proof.

Proposition 8. *Let $\Omega \subset \mathbb{R}^d$ be an open, acute, solid cone, let $\alpha \in i\Omega$ and assume $S(\zeta,\xi) \neq 0$, $\zeta,\xi \in T_\Omega$. An element $\phi \in L^\infty(\mathbb{R}^d)$ is the phase of an analytic function $F \in \mathcal{O}(T_\Omega)$, $0 \le \Re F \le 1$:*

$$F(\zeta) = iC + \int_{\mathbb{R}^d} H(\zeta,u;\alpha)\phi(u)du, \quad \zeta \in T_\Omega, \tag{20}$$

where $C \in \mathbb{R}$, if and only if $0 \le \phi \le 1$, a.e., and the "moment conditions"

$$\int_{\mathbb{R}^d} [H(\zeta,\sigma;\alpha) + H(\alpha,\sigma;\xi) - H(\zeta,\sigma;\xi)]\phi(\sigma)d\sigma = 0, \quad \zeta,\xi \in T_\Omega$$

hold.

Proof. In order to prove the non-trivial implication, let $\phi \in L^\infty(\mathbb{R}^d)$, $0 \le \phi \le 1$, a.e., and define the function $F(\zeta)$ by formula (20). In view of the definition of Herglotz' kernel, by taking $\zeta = \alpha$ we find

$$F(\alpha) - iC = \int \frac{|S(u,\alpha)|^2}{S(\alpha,\alpha)}\phi(u)du,$$

whence by addition

$$F(\zeta) + \overline{F(\alpha)} = 2\int \frac{S(\zeta,u)S(u,\alpha)}{S(\zeta,\alpha)}\phi(u)du.$$

Write this formula for $F(\xi) + \overline{F(\alpha)}$, too. According to the moment conditions in the statement, we find again by addition

$$F(\zeta) + \overline{F(\xi)} = 2\int \frac{S(\zeta, u)S(u, \xi)}{S(\zeta, \xi)}\phi(u)du, \quad \zeta, \xi \in T_\Omega.$$

In particular, $\Re F$ is the Poisson's transform of ϕ, and therefore $0 \le \Re F \le 1$. $\square$

5.2. *Supports in special sets and restricted tube domains*

The reader should be puzzled by now by the way too abstract and useless level of this article. It is time perhaps for some examples of phase regularity applied to measures supported by three basic convex shapes in euclidean space: the orthant, the euclidean l_2-ball and the l_1-ball. The first one will especially be dear to control theorists, because it contains in the particular case of one dimension familiar computations of Laplace transforms. We include the euclidean ball as the commonly occurring domain for measures, and the l_1 ball as it results in trigonometric moment data for phase functions, which is an appealing set up for entropy optimization.

To remove some normalizations that encumber the computations we break with the generic convention $T_\Omega = \mathbb{R}^d + i\Omega$. Instead, at the beginning of each example, we specify the domain, redefine the Fantappiè transforms and then summarize the derivation which was detailed in the Section 5 to obtain the final result.

5.2.1. *The orthant*

Let $\Omega = (0, \infty)^d$ be the open positive orthant in $\mathbb{R}^d$, self-dual in the sense $\Omega^* = \overline{\Omega}$, the closure of itself. Szegö's kernel of the tube domain over Ω is

$$S(z, w) = \frac{1}{(2\pi)^d}\int_{\Omega^*} e^{i(z-\overline{w})\cdot u}du = \frac{1}{(2\pi i)^d}\prod_{k=1}^{d}\frac{1}{\overline{w_k} - z_k}, \quad z, w \in T_\Omega = \mathbb{R}^d + i\Omega.$$

With the selection of the reference point $\alpha = (i, i, \ldots, i) \in i\Omega$, Herglotz kernel becomes

$$\begin{aligned} H(z, u; \alpha) &= 2\frac{S(z, u)S(u, \alpha)}{S(z, \alpha)} - \frac{|S(u, \alpha)|^2}{S(\alpha, \alpha)} \\ &= \frac{1}{(2\pi)^d}\left[2\prod_{k=1}^{d}\frac{1 - iz_k}{(u_k - z_k)(u_k + i)} - \prod_{k=1}^{d}\frac{2}{1 + u_k^2}\right] \\ &= 2\prod_{k=1}^{d}\frac{1}{2\pi i}\left(\frac{1}{u_k - z_k} - \frac{1}{u_k + i}\right) - \prod_{k=1}^{d}\frac{1}{2\pi i}\left(\frac{1}{u_k - i} - \frac{1}{u_k + i}\right). \end{aligned}$$

In particular, for $d = 1$ we recover the familiar Szegö kernel $S(z, w) = \frac{1}{2\pi i}\frac{1}{\overline{w} - z}$ of the upper half plane, and Herglotz kernel becomes

$$H(z, u; i) = \frac{1}{\pi i}\left(\frac{1}{u - z} - \frac{u}{u^2 + 1}\right).$$

Let $\Phi(z)$ be an analytic function, mapping the open upper half-plane into itself and satisfying $\lim_{s\to\infty}\Phi(is) = 0$. Then $\ln\Phi(z)$ is well defined, with $\Im\ln\Phi(z) \in (0, \pi)$. The argument preceding Proposition 7 remains valid, with the result

$$\Phi(z) = \exp[iF(z)], \quad \Im z > 0,$$

where $F(z)$ is analytic, $\Re F(z) \in (0, \pi)$ and

$$F(z) = i\Im F(i) + \int_{\mathbb{R}} H(z, u : i)\phi(u)du,$$

and $\phi \in L^\infty(\mathbb{R}),\ 0 \le \phi \le \pi$. In conclusion, we obtain the representation

$$\Phi(z) = e^{\Im F(i)} \exp \int_{\mathbb{R}} \left(\frac{1}{u - z} - \frac{u}{u^2 + 1} \right) \frac{\phi(u)}{\pi} du,$$

a formula already invoked in (21).

5.2.2. The l_2 ball

Let μ be a positive Borel measure supported by the closed unit ball $\mathbf{b}$ of $\mathbb{R}^d$, and denote by

$$a_\alpha(\mu) = \int_{\mathbf{b}} x^\alpha d\mu(x), \quad \alpha \in \mathbb{N}^d.$$

We consider a version of the Fantappiè transform of μ:

$$\Phi(\mu)(z) = \int_{\mathbf{b}} \frac{d\mu(x)}{1 - x \cdot z},$$

where $u \cdot v = u_1 v_1 + \cdots + u_d v_d$.

Note that $\Phi(\mu)$ is an analytic functions defined in the open unit ball $\mathbf{B}$ of $\mathbf{C}^d$. Its Taylor series expansion at $z = 0$ is reducible, modulo universal constants, to the moments of μ:

$$\Phi(\mu)(z) = \sum_\alpha \frac{|\alpha|!}{\alpha!} a_\alpha(\mu) z^\alpha.$$

Remark also that, for all $z \in \mathbf{B}$:

$$\Re \Phi(\mu)(z) = \int_{\mathbf{b}} \frac{1 - \Re x \cdot z}{|1 - x \cdot z|^2} d\mu(x) \ge 0.$$

By the maximum principle for pluri-harmonic functions, equality sign can happen only if $\Phi(\mu)(z)$ is identically equal to a purely imaginary constant, which is impossible, since $\Phi(\mu)(0) = \mu(\mathbf{b})$.

Define the function

$$F(z) \triangleq -i \ln i\, \Phi(\mu)(z)$$

on the unit ball, such that $i\,\Phi(\mu)(z) = \exp i F(z)$, with $\Re F(z) \in [0, \pi]$ for $z \in B$. The function $F(z)$ is analytic in the ball, and has a positive, bounded real part there. By the generalized Riesz–Herglotz formula, see [33,38], we infer:

$$F(z) = i\Im F(0) + \int_{\partial \mathbf{B}} [2S(z, w) - 1]\phi(w) d\sigma(w),$$

where $\partial \mathbf{B}$ is the unit sphere, $\sigma(w)$ is the surface element on $\partial \mathbf{B}$, normalized to have mass equal to one,

$$S(z, w) = \frac{1}{(1 - z \cdot \overline{w})^d}$$

is Szegö's kernel of the ball, and $\phi(w)$ is a measurable function on the sphere, satisfying

$$0 \le \phi(w) \le \pi, \quad w \in \partial\mathbf{B}.$$

Let us deal first with the free term, similarly as we did in the one-dimensional case:

$$\Phi(\mu)(0) = \mu(\mathbf{b}),$$

and therefore

$$F(0) = -i \ln i\mu(\mathbf{b}) = \underbrace{-i \ln \mu(\mathbf{b})}_{i\Im F(0)} + \frac{\pi}{2}.$$

The total mass of $\phi(w)$ is now easily computed by setting $z = 0$ in the Riesz–Herglotz formula, to obtain

$$-i \ln i\mu(\mathbf{b}) + \frac{\pi}{2} = -i \ln i\mu(\mathbf{b}) + \int_{\partial\mathbf{B}} \phi(w) d\sigma(w),$$

or

$$\int_{\partial\mathbf{B}} \phi(w) d\sigma(w) = \frac{\pi}{2}.$$

The Fantappié transform of the measure μ simplified by substituting the exact values for the free term and the total mass of the phase function ϕ:

$$\begin{aligned} i\,\Phi(\mu)(z) &= \exp i F(z) \\ &= \exp\left[\ln \mu(\mathbf{b}) + i \int_{\partial\mathbf{B}} [2S(z,w) - 1]\phi(w) d\sigma(w)\right] \\ \Phi(\mu)(z) &= \mu(\mathbf{b}) \exp\left\{i \int_{\partial\mathbf{B}} [2S(z,w) - 1]\phi(w) d\sigma(w) - i\frac{\pi}{2}\right\} \\ &= \mu(\mathbf{b}) \exp\left\{2i \int_{\partial\mathbf{B}} [S(z,w) - 1]\phi(w) d\sigma(w)\right\}. \end{aligned}$$

This formula relates, via a triangular, non-linear transformation, the moments $(a_\mu(\alpha))$ of μ to the moments $(a_\phi(\alpha))$ of the density ϕ. Additionally, note that the Koranyi–Pukánszky theorem asserts that all the multivariate moments of ϕ at mixed-sign indices are zero, i.e., for a multi-index α, $a_\phi(\alpha) = 0$, except possibly when $\forall i, \alpha_i \ge 0$, or $\forall i, \alpha_i \le 0$.

5.2.3. The l_1 ball

Let $\Delta = \{x \in \mathbb{R}^d;\ |x_1| + \cdots + |x_d| \le 1\}$ and consider a positive measure μ supported by Δ. Its Fantappiè transform is analytic in the unit polydisk $\mathbf{D}^d$, and has there a power series expansion

$$\Phi(\mu)(z) = \sum_\alpha \frac{|\alpha|!}{\alpha!} a_\mu(\alpha) z^\alpha.$$

Arguing as in the case of the ball, the function $\ln \Phi(\mu)$ is analytic in the open polydisk, and has non-negative imaginary part there bounded above by π. The analog of Riesz–Herglotz formula (as derived for the first time by Koranyi and Pukánszky [32]) yields a measurable function ϕ defined on the unit torus $\mathbb{T}^d$, with values in the interval $[0, \pi]$, so that:

$$\Phi(\mu)(z) = \mu(\Delta) \exp\left\{2i \int_{\mathbb{T}^d} [\Pi(z,w) - 1]\phi(w) d\theta(w)\right\},$$

where this time

$$\Pi(z, w) = \prod_{j=1}^{d} \frac{1}{1 - z_j \overline{w}_j},$$

and

$$d\theta(w) = \prod_{j=1}^{d} \frac{dw_j}{2\pi i w_j}.$$

Let $a_\phi(\alpha) = \int_{\mathbb{T}^d} \phi(w)\overline{w}^\alpha d\theta(w)$ denote the Fourier coefficients of the function ϕ (i.e. its trigonometric moments on the torus).

At the level of generating series we obtain the following transform

$$\sum_\alpha \frac{|\alpha|!}{\alpha!} a_\mu(\alpha) z^\alpha = \mu(\Delta) \exp\left[2i \sum_{\alpha \neq 0} a_\phi(\alpha) z^\alpha\right].$$

Consider a simple example.

Example. To verify that the above formulas are correct, we consider the 1D case, with the Dirac measure $d\mu = c\delta_a$, where $c > 0$, $a \in [-1, 1]$. The Fantappiè transform is the analytic function in the unit disk:

$$f(z) = \frac{c}{1 - az}.$$

Since $f(z)$ has positive real part, $if(z)$ has positive imaginary part in the disk. Whence $\ln if(z)$ is well defined, analytic, and has imaginary part in the interval $[0, \pi]$, thus the classical Riesz–Herglotz formula is applicable to the function $-i \ln if(z)$:

$$-i \ln if(z) = i\Im[-i \ln if(0)] + \int_{\mathbb{T}} \frac{1 + z\overline{w}}{1 - z\overline{w}} \phi(w) \frac{dw}{2\pi i w}.$$

Above, $\phi(w) = \Re(-i \ln if(w))$ is a measurable function on the torus with values in $[0, \pi]$. Note that $f(0) = c$, so that $\ln if(0) = \ln c + i\pi/2$, and $i\Im[-i \ln if(0)] = -i \ln c$. By evaluating $z = 0$ in the formula we obtain

$$-i \ln c + \pi/2 = -i \ln c + \int_{\mathbb{T}} \phi(w) \frac{dw}{2\pi i w},$$

that is

$$\pi/2 = \int_{\mathbb{T}} \phi(w) \frac{dw}{2\pi i w}.$$

Finally we get

$$\begin{aligned}
\ln i + \ln f(z) &= \ln c + i \int_{\mathbb{T}} \frac{1 + z\overline{w}}{1 - z\overline{w}} \phi(w) \frac{dw}{2\pi i w} \\
\ln f(z) &= \ln c + i \int_{\mathbb{T}} \left[\frac{1 + z\overline{w}}{1 - z\overline{w}} - 1\right] \phi(w) \frac{dw}{2\pi i w} \\
&= \ln c + 2i \int_{\mathbb{T}} \frac{z\overline{w}}{1 - z\overline{w}} \phi(w) \frac{dw}{2\pi i w} \\
&= \ln c + 2i \int_{\mathbb{T}} \left[\frac{1}{1 - z\overline{w}} - 1\right] \phi(w) \frac{dw}{2\pi i w},
\end{aligned}$$

which is consistent with our general formulas.

One step further, we can easily compute the positive ($n > 0$) Fourier coefficients of $\phi(w) = \Re(-i \ln i f(w))$:

$$\int_{\mathbb{T}} \phi(w)\overline{w}^n \frac{dw}{2\pi i w} = \int_{\mathbb{T}} \frac{-i}{2} \ln i f(w)\overline{w}^n \frac{dw}{2\pi i w}$$
$$= \frac{1}{2i}\int_{\mathbb{T}} \ln \frac{c}{1-aw}\overline{w}^n \frac{dw}{2\pi i w} = \frac{1}{2i}\frac{a^n}{n}.$$

The final verification:

$$\frac{c}{1-az} = c \exp \sum_{n=1}^{\infty} \frac{a^n}{n} z^n.$$

5.3. *The partial Fantappiè transform*

In this section we continue analytically the Fantappiè transform of a measure supported by a convex cone in a single direction, treating the rest of the variables as parameters, with a double benefit: a simple and well known formulas in 1D, and a tight control of the growth of the phase function. We closely follow below the article [7], although similar computations have appeared much earlier in the work of Nevanlinna and Verblunsky.

5.3.1. *Regularization by parametrized single-variable transforms*

The setting is the same: $\Gamma \subset \mathbb{R}^d$ is a closed, solid, acute convex cone and μ is a finite positive measure supported by its polar cone Γ^*. We consider the analytic extension of the Fantappiè transform:

$$\Phi(-z, y) = \int_{\Gamma^*} \frac{d\mu(x)}{-z + x \cdot y}, \quad y \in \Gamma,\ \Im z > 0.$$

Since

$$\frac{1}{-\overline{z} + x \cdot y} - \frac{1}{-z + x \cdot y} = \frac{\overline{z} - z}{|z - x \cdot y|^2}$$
$$(y \in \Gamma,\ \Im z > 0) \Rightarrow \Im \Phi(-z, y) > 0.$$

Thus, for any fixed $y \in \Gamma$, the function $z \mapsto i\,\Phi(-z, y)$ preserves the upper half-plane and hence it can be represented for $\Im w > 0$ as

$$\int_{\Gamma^*} \frac{d\mu(x)}{-z + x \cdot y} = \Phi(-z, y) = C(y) \exp \int_{\mathbb{R}} \left(\frac{1}{t - z} - \frac{t}{1 + t^2} \right) \phi_y(t) dt, \tag{21}$$

where $C(y) > 0$ and $0 \le \phi_y(t) \le 1$ both depending measurably on y, respectively y and t, see [7]. Both functions $C(y), \phi_y(t)$ are uniquely determined by $\Phi(z, y)$, hence by the measure μ. For illustration, we provide a simple example of a point mass:

Example. Take $\mu = c\delta_0$, where $c > 0$. First we obtain by direct integration

$$\frac{-1}{z} = \exp \int_0^{\infty} \left(\frac{1}{t - z} - \frac{t}{1 + t^2} \right) dt,$$

hence

$$\int_{\Gamma^*} \frac{c\delta_0(x)}{-z + ix \cdot y} = c \exp \int_0^{\infty} \left(\frac{1}{t - z} - \frac{t}{1 + t^2} \right) dt,$$

obtaining

$$C(y) = c, \qquad \phi_y(t) = \chi_{[0,\infty)}(t), \quad y \in \Gamma.$$

It is rather annoying that for such a simple measure as a point mass, the phase function has an unbounded support. Fortunately, there is a simple remedy, derived from an observation of Verblunsky [48,17].

Theorem 9. *Let μ be a finite positive measure supported on the cone Γ^*. For every $y \in \Gamma$ there exists a phase function $\xi_y \in L^1([0,\infty), dt)$, $0 \le \xi_y \le 1$, measurably depending on y, such that*

$$1 + \int_\Gamma \frac{d\mu(x)}{x \cdot y - z} = \exp \int_0^\infty \frac{\xi_y(t)dt}{t - z}, \quad \Im z > 0. \tag{22}$$

Moreover, if $\int_{\Gamma^*} |x|^n \, d\mu(x) < \infty$ for some $n \in \mathbb{N}$, then $\int_0^\infty t^n \xi_y(t)dt < \infty$ for all $y \in \Gamma$.

Proof. Fix a point $y \in \Gamma$ and denote by μ_y the push forward of the measure μ via the map $x \mapsto x \cdot y$. Specifically, for a test function $f \in C_0([0,\infty))$:

$$\int_{\Gamma^*} f(x \cdot y) d\mu(x) = \int_0^\infty f(t) d\mu_y(t). \tag{23}$$

In these terms, Fantappiè's transform of the measure μ becomes

$$\Phi(z, y) = \int_0^\infty \frac{d\mu_y(t)}{z + t}.$$

According to Verblunsky's theorem [7], there exists a function $\xi_y \in L^1([0,\infty), dt)$, $0 \le \xi_y \le 1$ with the property

$$1 + \Phi(-z, y) = \exp \int_0^\infty \frac{\xi_y(t)dt}{t - z}, \quad \Im z > 0.$$

The boundary limit operations producing ξ_y from μ_y imply that the dependence $y \mapsto \xi_y$ is (weakly) measurable, as a map from Γ to $L^1([0,\infty), dt)$. Finally, Theorem A.b of [7] implies the finiteness of the first $n + 1$ moments of every ξ_y, provided that the moments of μ_y of the same order are finite. □

Returning to our simple example, $\mu = c\delta_0$, we find this time

$$1 + \Phi(-z, y) = 1 + \int_{\Gamma^*} \frac{c\delta_0(x)}{x \cdot y - z} = 1 - \frac{c}{z} = \exp \int_0^c \frac{dt}{t - z},$$

whence

$$\xi_y = \chi_{[0,c]}, \quad y \in \Gamma.$$

Assume that $\int_{\Gamma^*} |x|^n \, d\mu(x) < \infty$ for a positive value of n and denote the initial moments of μ as:

$$\gamma_\alpha = \int_{\Gamma^*} x^\alpha d\mu(x), \quad |\alpha| \le n.$$

Similarly, denote

$$c(y)_j = \int_0^\infty t^j \xi_y(t)dt, \quad 0 \le j \le n.$$

Then one can identify asymptotically the series expansion (in a wedge with vertex at $z = 0$) of the two terms in (22), obtaining the algebraic relation

$$1 - \sum_{k=0}^{n} \frac{k!}{z^{k+1}} \left(\sum_{|\alpha|=k} \frac{y^\alpha}{\alpha!} \gamma_\alpha \right) + O(z^{-n-2}) = \exp\left(- \sum_{j=0}^{n} \frac{c_j(y)}{z^{j+1}} \right). \tag{24}$$

For details see again [7]. In particular, by equating the coefficients of z^{-1} one finds Verblunsky's identity

$$c_0(y) = \gamma_0, \quad y \in \Gamma.$$

Note that, after expanding the exponential series, the moment $c_j(y)$ is given by a universal polynomial function in the variables γ_α, $|\alpha| \le j$. It is exactly this system of polynomial equation which was discovered and exploited by Markov (in dimension one).

5.3.2. Inversion through the Radon transform

Using entropy optimization, for each $p \in \Gamma$, we obtain functions ξ_p^* that approximate Aronszajn–Donoghue phase functions ξ_p of push forward measures μ_p defined by (23). The inversion procedure given in Section 4 would be able to recover μ_p^* measures, however, piecing together approximation μ^* from "slices" μ_p^* would be a challenging task. Instead, we seek to recover the Radon transform $\mathcal{R}\mu^*(p, t)$ of approximation μ, which is then inverted by one of the standard algorithms for inverse Radon transform.

We start by relating the partial Fantappiè transform to Cauchy transform

$$\Phi\mu(p, -z) = \int_{\Gamma^*} \frac{d\mu(x)}{-z + p \cdot x} = \mathcal{C}\mu_p(z) = \exp \mathcal{C}\xi_p(z) - 1,$$

with a slight abuse of notation for the Cauchy transform, see (22). To avoid the singular integral in passing from complex to the real Fantappiè directly, we sum the Plemelj–Sokhotski formulas (11) to obtain

$$\begin{aligned} \Phi\mu(p, p_0) &= \mathcal{C}\mu_p(-p_0) = \lim_{\epsilon \downarrow 0} \frac{1}{2} \left[\mathcal{C}\mu_p(-p_0 - i\epsilon) + \mathcal{C}\mu_p(-p_0 + i\epsilon) \right] \\ &= \exp\left[-\pi \mathcal{H}\xi_p(-p_0) \right] \cos\left[\pi \xi_p(-p_0) \right] - 1, \end{aligned}$$

using the derivation analogous to derivation of (13), where we defined function f_p just to relieve notation for the rest of the procedure.

To connect the Fantappiè transform to the Radon transform of a measure, we again follow Henkin and Shananin [24]. For a positive measure μ with density ξ supported in the positive orthant $\mathbb{R}^n_+$, define the Radon transform as

$$\mathcal{R}\mu(\theta, s) \triangleq \int_{H(\theta,s)} \xi(x) d\sigma(x), \tag{25}$$

$$= \int_{\mathbb{R}^n} \delta(s - x \cdot \theta) \xi(x) d\sigma(x), \tag{26}$$

where $H(\theta, s) \triangleq \{x \in \mathbb{R}^n : x \cdot \theta = s\}$ is the integration hyperplane and $d\sigma(x)$ its surface area measure. Henkin and Shananin give the following relation between Fantappiè and Radon transforms of rapidly decaying measure μ:

$$\Phi\mu(p, p_0) = \int_0^\infty \frac{\mathcal{R}\mu(p, \tau)}{\tau + p_0} d\tau, \tag{27}$$

valid for $p_0 > 0$, $p \in \mathbb{R}^n_+$. This expression can be interpreted as a composition of Hilbert and Radon transforms

$$\Phi\mu(p, p_0) = [\mathcal{H}\mathcal{R}\mu(p, \cdot)](-p_0),$$

where Hilbert transform is taken along the offset parameter. As $-\mathcal{H}^2$ is the identity operator, by applying another Hilbert transform, we can evaluate the Radon transform as

$$\mathcal{R}\mu(p, p_0) = -\frac{1}{\pi}[\mathcal{H}f_p](p_0),$$

where we define

$$f_p(p_0) \triangleq \exp\left[-\pi\mathcal{H}\xi_p(p_0)\right]\cos\left[\pi\xi_p(p_0)\right] - 1.$$

As mentioned before, Hilbert transform can be efficiently evaluated using FFT. Therefore, for each selected p, we can evaluate the Radon transform along p_0 axis. One might validly ask why we decided to go through this labyrinthine process only to evaluate a Radon transform of μ. The answer lies in the method used to obtain moment data. If the original measure μ has a density fully accessible for arbitrary Radon-type measurements, as it is in medical tomography, then the entire procedure is superfluous as we can access $\mathcal{R}_\mu(p)$ directly at any $p \in \Gamma$. However, for singular measures and in some settings, e.g., invariant measures on attractors of dynamical systems, Radon-measurements are not directly possible and the described procedure becomes an acceptable path to reconstruction.

An unfortunate obstacle prevents us for completing the process by invoking a readily-available inversion algorithm for the Radon transform. The relation between the Fantappiè and Radon transform (27) holds only in the positive orthant $(p_0, p) \in \mathbb{R}^{n+1}_+$, which is, in general, not enough for a numerically-stable reconstruction [41]. It is possible that this obstacle can be removed by considering certain symmetries of the problem. However, these considerations would lead us too far from the central theme of this paper and we plan to explore them in a subsequent paper.

6. Conclusions

In this paper, we presented an approach aimed at representing singular measures using bounded densities. Our goal was to use existing entropy optimization methods to solve truncated moment problems for singular measures. Previously, the entropy optimization could not be applied to singular measures due to lack of convergence in the optimization procedure.

The paper adds two steps that bookend the entropy optimization. The regularization step uses a triangular, recursive transformation on the input moment set to produce the conditioned moment set which is a feasible input for entropy optimization. The inversion step uses the density that solves the entropy optimization constrained by conditioned moments, and recovers an approximation to the original measure.

We presented both steps in detail for the support in a one-dimensional space. Two different settings, unbounded and bounded supports, were analyzed, resulting in, respectively, a more general formulation using power moments, and a numerically favorable formulation using trigonometric moments.

We generalize the regularization step to multivariate domains, in particular supports of measures in wedges and particular compact domains in $\mathbb{R}^d$. The inversion, however, proved to

be a technically more demanding task, with the direct generalization of the Plemelj–Sokhotski formulas requiring a detailed application of theory of Clifford algebras. We did not tackle such generalization in this paper, nevertheless, we presented an outline of a simpler tomographic process. It would reduce a multivariate inversion problem to a family of inversion along rays in the domain, which can be solved using the one-dimensional method. The information based on ray transforms could then be integrated into the approximation of the original measure using methods of tomography.

This paper is the first part of this research effort. We plan to explore the full solution to the multivariate inversion step using tomography in one of the follow-up papers. Furthermore, a future paper should formulate error analysis and present numerical confirmation of the usefulness of our method, applied to concrete examples.

Appendix A. The Miller–Nakos algorithm for exponentiation of a power series

Computing coefficients of an exponentiated (formal) power series can be performed recursively, i.e., from knowledge of the coefficients of the base and coefficients of the lower coefficients. The original algorithm in a single variable is given by Peter Henrici [26] who attributes it to J.C.P. Miller. The extension to the multivariate case is due to George Nakos [40], which is unfortunately published in a journal that is not easily accessible. Therefore, we give the algorithm in its entirety here, stressing that all the original ideas were present in the above papers. Ours are the choice of notation and the (trivial) extension to the case of series with zero free terms, mentioned at the end of the section.

Let $\alpha, \beta, \mu, \gamma \in \mathbb{N}_0^d$ denote multi-indices and let the basis multi-index be $\epsilon_i = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$ where 1 is at the i-th position. Unless noted otherwise, sums over multi-indices range over all multi-indices $\mathbb{N}_0^d$. Furthermore, denote $\partial_i = \frac{\partial}{\partial x_i}$.

Theorem 10 (*J.C.P. Miller, G. Nakos*)**.** *Let $A(x)$ and $B(x)$ be multivariate power series with non-zero free terms, given by expressions*

$$A(x) \triangleq \sum_{\alpha} a_\alpha x^\alpha \qquad B(x) \triangleq \sum_{\beta} b_\beta x^\beta,$$

with $\alpha_0 \neq 0$, $\beta_0 \neq 0$.

If the power series are related by equation

$$B(x) = A(x)^k,$$

then coefficients b_β can be computed recursively by expressions

$$\begin{aligned} b_0 &= a_0^k \\ b_\mu &= \sum_{0<\gamma\leq\mu} \frac{1}{a_0}\left[(k+1)\frac{|\gamma/\mu|}{|\mu/\mu|} - 1\right] a_\gamma b_{\mu-\gamma}, \end{aligned} \tag{A.1}$$

where

$$|\alpha/\beta| = \sum_{i:\beta_i\neq 0} \alpha_i/\beta_i.$$ [5]

[5] *Particularly,$|\alpha/\alpha|$counts the nonzero elements in α.*

Proof. The proof of the algorithm rests on the identity

$$A(x)\partial_i B(x) = kB(x)\partial_i A(x), \tag{A.2}$$

derived by expanding $\partial_i B(x)$ by chain rule and multiplying both sides by $A(x)$. Series expansions for derivatives $\partial_i A(x)$, and analogously for $\partial_i B(x)$, is

$$\partial_i A(x) = \sum_\alpha a_\alpha \alpha_i x^{\alpha-\epsilon_i}.$$

Expanding (A.2) into series we obtain

$$\left(\sum_\alpha a_\alpha x^\alpha\right)\left(\sum_\beta b_\beta \beta_i x^{\beta-\epsilon_i}\right) = k\left(\sum_\beta b_\beta x^\beta\right)\left(\sum_\alpha a_\alpha \alpha_i x^{\alpha-\epsilon_i}\right).$$

Introduce following change of indices: on the left hand side over $\beta, \beta-\epsilon_i \mapsto \beta, \beta \mapsto \beta+\epsilon_i, \beta_i \mapsto \beta_i+1$, and the analogous substitutions on the right hand side in sum over α, to obtain the identity

$$\sum_\alpha a_\alpha x^\alpha \sum_\beta (\beta_i+1) b_{\beta+\epsilon_i} x^\beta = k\sum_\beta b_\beta x^\beta \sum_\alpha (\alpha_i+1) a_{\alpha+\epsilon_i} x^\alpha.$$

At this point, the goal is to compute coefficient b_μ from knowledge of a_α and b_γ for $\gamma<\mu$. The products of power series $\sum_\alpha a_\alpha x^\alpha \sum_\beta b_\beta x^\beta = \sum_\gamma c_\gamma x^\gamma$ are computed using convolution over coefficients $c_\gamma \triangleq \sum_{\alpha+\beta=\gamma} a_\alpha b_\beta$. The coefficients at index $\mu-\epsilon_i$ on both sides of the above identity have to match, yielding coefficient identity

$$\sum_{\alpha+\beta=\mu-\epsilon_i} a_\alpha b_{\beta+\epsilon_i}(\beta_i+1) = k\sum_{\alpha+\beta=\mu-\epsilon_i} b_\beta a_{\alpha+\epsilon_i}(\alpha_i+1).$$

The expressions can be simplified by another re-indexing, taking into account constraints on indices in summation, and symmetry in summation: on LHS, $\alpha\mapsto\gamma, \epsilon_i+\beta\mapsto\mu-\gamma$, on RHS $\beta\mapsto\mu-\gamma, \epsilon_i+\alpha\mapsto\gamma$, to obtain

$$\sum_{0\le\gamma\le mu} a_\gamma b_{\mu-\gamma}(\mu_i-\gamma_i) = k\sum_{0\le\gamma\le\mu} b_{\mu-\gamma} a_\gamma \gamma_i,$$

or

$$\sum_{0\le\gamma\le\mu} a_\gamma b_{\mu-\gamma}\left[\mu_i-(k+1)\gamma_i\right] = 0.$$

To solve for b_μ, extract $\gamma=0$ case from the sum to obtain

$$a_0 b_\mu \mu_i = \sum_{0<\gamma\le\mu} a_\gamma b_{\mu-\gamma}\left[(k+1)\gamma_i-\mu_i\right]$$

$$b_\mu = \frac{1}{a_0}\sum_{0<\gamma\le\mu} a_\gamma b_{\mu-\gamma}\left[(k+1)\frac{\gamma_i}{\mu_i}-1\right].$$

Since this expression is valid if and only if $\mu_i\neq 0$, we can sum over all such cases to obtain

$$|\mu/\mu|\, b_\mu = \frac{1}{a_0}\sum_{0<\gamma\le\mu} a_\gamma b_{\mu-\gamma}\left[(k+1)\,|\gamma/\mu| - |\mu/\mu|\right],$$

which, through dividing by $|\mu/\mu|$, yields the expression (A.1). □

The moment expansion series that we use have a zero free term, which is essential for solving for b_μ in the last step of the proof. This can be resolved by adding an extra step which uses the binomial expansion

$$[(A(x)+1)-1]^n = \sum_{0\le k\le n} \binom{n}{k} (-1)^{n-k} [A(x)+1]^k ,$$

where $A(x)$ is the series without a free term. In general, this step does involve n extra computations, however, the moment conversions such as (9) require all the powers between 0 and n anyway, so there is no additional cost involved for our purposes.

Appendix B. The fast iterative algorithm for entropy optimization

This algorithm has been described in [8], and it is based on explicit discretization of the entropy functional, which reduces computation of moments to a matrix multiplication. For completeness, we present it here in a distilled form.

First, assume the domain is [0, 1] interval, and choose points x_k with quadrature weights w_k for $k = 1, \dots, K$, i.e.,

$$\int_0^1 f(x)dx \approx \sum_{j=1}^{J} f(x_j) w_j .$$

For an arbitrary distribution $p(x)$, we will write $\mathbf{p} = (p_j)$, $p_j := p(x_j)$, and $\tilde{\mathbf{p}} = (\tilde{p}_j)$, $\tilde{p}_j = w_j p_j$. To evaluate moments of $p(x)$ we employ the matrix $\mathbf{A} = (a_{ij})$, whose rows are evaluations of monomials on the array x_k. For first N power moments, $\mathbf{A}$ will be a row-truncated Vandermonde matrix $a_{ij} = x_j^i$, where $i = 1, \dots, N$, and $j = 1, \dots, K$. The vector of moments $\mu = (\mu_i)$ for a discretized distribution p_j is easily evaluated by taking the product $\mu = \mathbf{A} \cdot \tilde{\mathbf{p}}$.

Let $\alpha_i,\ i = 1, \dots, N$ be the set of the Lagrange multipliers (dual variables), in which the entropy optimization is unconstrained. The primal is then evaluated by function

$$p(\alpha)_j = \exp\left[(\mathbf{A}^T \cdot \alpha)_j - 1\right], \tag{B.1}$$

with the goal of finding α^* such that $p_j(\alpha^*) \approx p_j^*$. Constraint deviation vector is

$$h_i(\alpha) = [\mathbf{A} \cdot \mathbf{p}(\alpha)]_i - \mu_i . \tag{B.2}$$

The optimization program strives to achieve

$$\min \left(\mathbf{1}^T \mathbf{p}(\alpha) - \mu^T \alpha\right),$$

by cyclically updating components of α. Denote kth iteration of α by $\alpha^{(k)}$.

Let k be a step counter, and set $i = (k \bmod N) + 1$ to be the index of the Lagrange multiplier updated in the kth step. Fix the convergence tolerance $\epsilon > 0$, and denote initial step by $k = 0$. The initial vector $\alpha^{(0)}$ can be chosen as random numbers in some interval, a constant vector, or some other vector.

Compute the correction factor

$$\lambda^{(k)} = \ln \frac{\mu_i}{\left[\mathbf{A} \cdot \mathbf{p}(\alpha^{(k)})\right]_i}, \tag{B.3}$$

and update ith Lagrange multiplier

$$\alpha_i^{(k+1)} = \alpha_i^{(k)} + \lambda^{(k)}$$
$$\alpha_j^{(k+1)} = \alpha_j^{(k)}, \quad \text{for } j \neq i.$$

If $\left\| h(\alpha^{(k+1)}) \right\| < \epsilon$, then $\alpha^* = \alpha^{(k+1)}$, otherwise, increase k by one, and restart from computation of the correction factor.

The paper [8] asserts that the algorithm converges when $\mu_i > 0$ and $a_{ij} \in [0, 1], \forall i$, or when $\mu_i < 0$ and $a_{ij} \in [-1, 0], \ \forall i$. When this is not the case, the authors provide a pre-conditioning step that modifies $\mathbf{A}$ and μ to ensure convergence.

The presented algorithm can be extended to cases where generalized moments are taken, instead of power moments. In those cases, the moments are not necessarily positive, nor is the matrix $\mathbf{A}$, so they have to be rescaled before the correction factor formula (B.3) can be used.

To generalized moments on another interval, let $T_i(x)$ be linearly independent functions which are used to generate generalized moments. The matrix A is then given by elements $a_{ij} = T_i(x_j)$ for $i = 1, \ldots, N$, $j = 1, \ldots, K$. Choose a positive constant $\delta > 0$ and let

$$\begin{aligned}
&\text{(needed offset)} && u_i = -\min_j(a_{ij}) + \delta \\
&\text{(scale)} && M_i = \max_j(u_j + a_{ij}) \\
&\text{(scaling factor)} && t_i = \frac{1}{(M_i + \delta)}.
\end{aligned}$$

The original paper used $\delta = 1$ and $t_i = [N(M_i + \delta)]^{-1}$ but we found that such settings result in somewhat slower convergence rates.

The conditioned matrix $\mathbf{A}'$ and moment vector μ' are then computed as

$$a'_{ij} = t_i(u_i + a_{ij})$$
$$\mu'_i = t_i(u_i + \mu_i),$$

which ensures convergence conditions.

Now the original program is modified by replacing the primal formula (B.1), the correction factor formula (B.3), and constraint Eq. (B.2) by, respectively,

$$\mathbf{p}'(\alpha) = \exp\left[(\mathbf{A}'^T \cdot \alpha)_j - 1\right],$$
$$\lambda'^{(k)} = \ln \frac{\mu'_i}{\left[\mathbf{A}' \cdot \mathbf{p}'(\alpha^{(k)})\right]_i},$$

and

$$h'_i(\alpha) = \left[\mathbf{A} \cdot \mathbf{p}'(\alpha)\right]_i - \mu_i.$$

Note that in the constraint deviation, we evaluate the unconditioned moments of the conditioned primal. The justification is the fact that $\mathbf{p}^* = \mathbf{p}'(\alpha)$ simultaneously solves the conditioned problem $\mathbf{A}'\mathbf{p} = \mu'$ and unconditioned problem $\mathbf{A}\mathbf{p} = \mu$, due to linearity of the pre-conditioning step. We, therefore, find the unconditioned constraint deviation condition more intuitive to use, which makes it easier to choose a desired convergence tolerance ϵ.

References

[1] R.V. Abramov, An improved algorithm for the multidimensional moment-constrained maximum entropy problem, Journal of Computational Physics 226 (1) (2007) 621–644. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR\&s1=MR2356853.

[2] L.A. Aizenberg, Š.A. Dautov, Holomorphic functions of several complex variables with nonnegative real part. Traces of holomorphic and pluriharmonic functions on the Šilov boundary, Mat. Sb. (NS) 99(141) (3) (1976) 342–355. 479.

[3] N.I. Akhiezer, M.G. Krein, Some Questions in the Theory of Moments, in: Translations of Mathematical Monographs, vol. 2, American Mathematical Society, Providence, RI, 1962, Translated by W. Fleming and D. Prill.

[4] N.I. Akhiezer, The Classical Moment Problem, Hafner, New York, 1965.

[5] C.-G. Ambrozie, Multivariate truncated moments problems and maximum entropy. arXiv.orgmath.FA. URL: http://arxiv.org/abs/1111.7123v4.

[6] M. Andersson, M. Passare, R. Sigurdsson, Complex Convexity and Analytic Functionals, in: Progress in Mathematics, vol. 225, Birkhäuser, Basel, 2004, http://dx.doi.org/10.1007/978-3-0348-7871-5. URL: http://www.ams.org/mathscinet-getitem?mr=2060426.

[7] N. Aronszajn, W. Donoghue, On exponential representations of analytic functions in the upper half-plane with positive imaginary part, Journal d'Analyse Mathématique 5 (1) (1956) 321–388.

[8] K. Bandyopadhyay, A. Bhattacharya, P. Biswas, D. Drabold, Maximum entropy and the problem of moments: a stable algorithm, Physical Review E. Statistical, Nonlinear, and Soft Matter Physics 71 (5) (2005) http://dx.doi.org/10.1103/PhysRevE.71.057701.

[9] C.M. Bender, L.R. Mead, N. Papanicolaou, Maximum entropy summation of divergent perturbation series, Journal of Mathematical Physics 28 (5) (1987) 1016–1018. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR\&s1=MR887016.

[10] M.S. Birman, D.R. Yafaev, The spectral shift function. The papers of M.G. Krein and their further development, Algebra i Analiz 4 (5) (1992) 1–44. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR\&s1=MR1202723.

[11] P. Biswas, H. Shimoyama, L.R. Mead, Lyapunov exponents and the natural invariant density determination of chaotic maps: an iterative maximum entropy ansatz, Journal of Physics A: Mathematical and General 43 (12) (2010) 125103. 12. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR&s1=MR2599797.

[12] G. Blekherman, J.B. Lasserre, The truncated K-moment problem for closure of open sets. arXiv.org math.OC. URL: http://arxiv.org/abs/1108.0627v1.

[13] J.M. Borwein, A.S. Lewis, On the convergence of moment problems, Transactions of the American Mathematical Society 325 (1) (1991) 249–271. http://dx.doi.org/10.2307/2001670.

[14] J.M. Borwein, Maximum entropy and feasibility methods for convex and nonconvex inverse problems, Optimization. A Journal of Mathematical Programming and Operations Research 61 (1) (2012) 1–33. http://dx.doi.org/10.1080/02331934.2011.632502. URL: http://dx.doi.org/10.1080/02331934.2011.632502.

[15] Z. Cai, R. Li, Numerical regularized moment method of arbitrary order for Boltzmann-BGK equation, SIAM Journal on Scientific Computing 32 (5) (2010) 2875–2907. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR\&s1=MR2729444.

[16] Commutative Harmonic Analysis. I, in: Encyclopaedia of Mathematical Sciences, vol. 15, Springer-Verlag, Berlin, 1991.

[17] H. Delange, On two theorems of S. Verblunsky, Proceedings of the Cambridge Philosophical Society 46 (1950) 57–66.

[18] J. Eckmann, D. Ruelle, Ergodic theory of chaos and strange attractors, Reviews of Modern Physics 57 (3) (1985) 617–656. http://dx.doi.org/10.1103/RevModPhys.57.617. URL: http://link.aps.org/doi/10.1103/RevModPhys.57.617.

[19] J.M. Einbu, On the existence of a class of maximum-entropy probability density functions, Institute of Electrical and Electronics Engineers. Transactions on Information Theory IT-23 (6) (1977) 772–775. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR&s1=MR0529946.

[20] J. Faraut, A. Korányi, Analysis on Symmetric Cones, in: Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1994.

[21] B.A. Fuks, Theory of Analytic Functions of Several Complex Variables, in: Translations of Mathematical Monographs, vol. 8, American Mathematical Society, Providence, RI, 1963, URL: http://www.ams.org/mathscinet-getitem?mr=MR0168793.

[22] T.T. Georgiou, Relative entropy and the multivariable multidimensional moment problem, Institute of Electrical and Electronics Engineers. Transactions on Information Theory 52 (3) (2006) 1052–1066. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR\&s1=MR2238070.

[23] C.D. Hauck, C.D. Levermore, A.L. Tits, Convex duality and entropy-based moment closures: characterizing degenerate densities, SIAM Journal on Control and Optimization 47 (4) (2008) 1977–2015. http://dx.doi.org/10.1137/070691139. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR\&s1=MR2421338.

[24] G.M. Henkin, A.A. Shananin, Bernstein Theorems and Radon Transform. Application to the Theory of Production Functions, in: Transl. Math. Monogr., vol. 81, Amer. Math. Soc., Providence, RI, 1990, URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR&s1=MR1104018.

[25] G.M. Henkin, A.A. Shananin, $\mathbb{C}^n$-capacity and multidimensional moment problem, in: Proceedings of the Symposium on Value Distribution Theory in Several Complex Variables: On the occasion of the Inauguration of Wilhelm Stoll as the Vincent F. Duncan and Annamarie Micus Duncan Professor of Mathematics, in: University of Not., vol. 12, 1992, pp. 69–85. April 28–29, 1990. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR\&s1=MR1243019.

[26] P. Henrici, Automatic computations with power series, Journal of the Association for Computing Machinery 3 (1956) 10–15.

[27] P. Henrici, Applied and Computational Complex Analysis. Vol. 2, in: Wiley Classics Library, vol. 2, Wiley-Interscience, New York, 1977, URL: http://www.ams.org/mathscinet-getitem?mr=MR1164865.

[28] P. Henrici, Applied and Computational Complex Analysis. Vol. 3, in: Pure and Applied Mathematics (New York), vol. 3, Wiley-Interscience, New York, 1977, URL: http://www.ams.org/mathscinet-getitem?mr=MR822470.

[29] E.T. Jaynes, On the rationale of maximum-entropy methods, Proceedings of the IEEE 70 (9) (1982) 939–952.

[30] M. Junk, Maximum entropy for reduced moment problems, Mathematical Models & Methods in Applied Sciences 10 (7) (2000) 1001–1025. URL: http://dx.doi.org/10.1142/S0218202500000513.

[31] F.W. King, Hilbert Transforms. Vol. 1, in: Encyclopedia of Mathematics and its Applications, vol. 124, Cambridge University Press, Cambridge, 2009, URL: http://www.ams.org/mathscinet-getitem?mr=MR2542214.

[32] A. Korányi, L. Pukánszky, Holomorphic functions with positive real part on polycylinders, Transactions of the American Mathematical Society 108 (1963) 449–456. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR\&s1=MR0159029.

[33] A. Korányi, The Poisson integral for generalized half-planes and bounded symmetric domains, Annals of Mathematics 82 (1965) 332–350. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR\&s1=MR0200478.

[34] M.G. Krein, The ideas of P.L. Čebyšev and A.A. Markov in the theory of limiting values of integrals and their further development, American Mathematical Society Translations 12 (1959) 1–121. URL: http://www.ams.org/mathscinet-getitem?mr=MR0113106.

[35] M.G. Krein, A.A. Nudelman, The Markov Moment Problem and Extremal Problems, American Mathematical Society, Providence, RI, 1977.

[36] H.J. Landau, Maximum entropy and the moment problem, American Mathematical Society. Bulletin. New Series 16 (1) (1987) 47–77. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR&s1=MR866018.

[37] C.D. Levermore, Moment closure hierarchies for kinetic theories, Journal of Statistical Physics 83 (5–6) (1996) 1021–1065. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR&s1=MR1392419.

[38] J.E. McCarthy, M. Putinar, Positivity aspects of the Fantappié transform, Journal d'Analyse Mathématique 97 (2005) 57–82. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR&s1=MR2274973.

[39] L.R. Mead, N. Papanicolaou, Maximum entropy in the problem of moments, Journal of Mathematical Physics 25 (8) (1984) 2404–2417. URL: http://www.ams.org/mathscinet/search/publications.html?pg1=MR&s1=MR751523.

[40] G. Nakos, Expansions of powers of multivariate formal power series, The Mathematica Journal 3 (1) (1993) 45–47.

[41] F. Natterer, The Mathematics of Computerized Tomography, in: Classics in Applied Mathematics, vol. 32, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.

[42] V. Palamodov, Reconstructive Integral Geometry, in: Monographs in Mathematics, vol. 98, Birkhäuser Verlag, Basel, 2004.

[43] A.G. Poltoratski, The Krein spectral shift and rank one perturbations of spectra, Algebra i Analiz 10 (5) (1998) 143–183. URL: http://arxiv.org/abs/math/9601206v1.

[44] A. Prestel, C.N. Delzell, Positive Polynomials, in: Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2001, URL: http://www.ams.org/mathscinet-getitem?mr=MR1829790.

[45] M. Riesz, Sur le problème des moments, troisième note, Arkiv för Matematik, Astronomi och Fysik 17 (16) (1923) 1–52.

[46] J.A. Shohat, J.D. Tamarkin, The Problem of Moments, in: American Mathematical Society Mathematical Surveys, vol. II, American Mathematical Society, New York, 1943.

[47] H. Sigurgeirsson, A.M. Stuart, Statistics from computations, in: Foundations of Computational Mathematics, Oxford, 1999, Cambridge Univ. Press, Cambridge, 2001, pp. 323–344. URL: http://www.ams.org/mathscinet-getitem?mr=MR1839148.

[48] S. Verblunsky, On the initial moments of a bounded function, Proceedings of the Cambridge Philosophical Society 43 (1947) 275–279.