# On Positive and Conditionally Negative Definite Functions with a Singularity at Zero, and their Applications in Potential Theory

T. R. L. Phillips

August 2018

A thesis submitted for the degree of Doctor of Philosophy
Supervised by Prof. Anatoly Zhigljavsky and Prof. Karl Michael Schmidt

## Acknowledgements

To a great friend and an even greater mind, Anatoly; thank you. I wholeheartedly appreciate everything you have done for me. Our discussions on life and betting will be sorely missed.

Professor Schmidt; thank you for answering every one of my questions, no matter how obscure. Your level of knowledge never ceases to amaze me.

To my biggest fans, Mum and Dad; because I owe a lot of my success to you. Thank you.

## Abstract

It is widely known that positive and conditionally negative definite functions take finite values at the origin. Nevertheless, there exist functions with a singularity at zero, arising naturally e.g. in potential theory or the study of (continuous) extremal measures, which still exhibit the general characteristics of positive or conditional negative definiteness.

Taking a framework set up by Lionel Cooper as a motivation, we study the general properties of functions which are positive definite in an extended sense. We prove a Bochner-type theorem and, as a consequence, show how unbounded positive definite functions arise as limits of classical positive definite functions, as well as that their space is closed under convolution. Moreover, we provide criteria for a function to be positive definite in the extended sense, showing in particular that complete monotonicity in conjunction with local absolute integrability is sufficient.

The celebrated Schoenberg theorem establishes a relation between positive definite and conditionally negative definite functions. By introducing a notion of conditional negative definiteness which accounts for the classical, non-singular conditionally negative definite functions, as well as functions which are unbounded at the origin, we extend this result to real-valued functions with a singularity at zero. Moreover, we demonstrate how singular conditionally negative definite functions arise as limits of classical conditionally negative definite functions and provide several examples of functions which are unbounded at the origin and conditionally negative definite in an extended sense.

Finally, we study the convexity and minimisation of the energy associated with various singular, completely monotone functions, which have not previously been considered in the field of potential theory or experimental design and solve the corresponding energy problems by means of numerically computing approximations to the (optimal) minimising measures.

**Keywords:** unbounded positive definite function; Bochner's theorem; completely monotone function; singular conditionally negative definite function; Schoenberg's theorem; convexity; energy problem; optimal measure.

## Declaration

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ......................... (candidate) Date ...............

**Statement 1**
This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD (mathematics).

Signed ......................... (candidate) Date ...............

**Statement 2**
This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ......................... (candidate) Date ...............

**Statement 3**
I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ......................... (candidate) Date ...............

# Contents

# 1 Introduction

The concept of positive definite sequences, arising in the context of a problem in complex function theory posed by Carathéodory [13], was introduced in 1911 by Toeplitz [44]. Herglotz [18] established a connection between positive definite sequences and the trigonometric moment problem. Motivated by the work of Carathéodory and Toeplitz, Mathias [23] and later Bochner [9] defined and studied the properties of positive definite functions, specifically their harmonic analysis. Before these developments, however, Mercer [24] had studied the more general concept of positive definite kernels in research on integral equations.

According to the classical standard definition, a function $f : \mathbb{R}^d \to \mathbb{C}$ is positive definite if

$$\sum_{i,j=1}^{n} f(\mathbf{x}_i - \mathbf{x}_j)\, v_i\, \overline{v_j} \geq 0 \tag{1}$$

for all $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ and $v_1, v_2, \ldots, v_n \in \mathbb{C}$, with any $n \in \mathbb{N}$; in other words, if the matrix $\left[f(\mathbf{x}_i - \mathbf{x}_j)\right]_{i,j=1}^{n}$ is non-negative definite for all $n \in \mathbb{N}$ and $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$. Using (1) with $n = 2$, $\mathbf{x}_1 = \mathbf{0}$, $\mathbf{x}_2 = \mathbf{x}$, $v_1 = 1$ and $v_2$ such that $v_2 f(\mathbf{x}) = -|f(\mathbf{x})|$, it can be shown that $|f(\mathbf{x})| \leq f(\mathbf{0})$ for all $\mathbf{x} \in \mathbb{R}^d$. Hence positive definite functions by the standard definition are always bounded.

One of the principle results on this subject is Bochner's theorem [9, Chapter IV.20], which states that a function $f : \mathbb{R}^d \to \mathbb{C}$ is continuous and positive definite if and only if it is the (inverse) Fourier transform of a finite, non-negative measure $\mu$ on $\mathbb{R}^d$, i.e.

$$f(\mathbf{x}) = \check{\mu}(\mathbf{x}) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{i\mathbf{x}\cdot\mathbf{z}}\, \mu(d\mathbf{z}) \qquad (\mathbf{x} \in \mathbb{R}^d).$$

Thus, Bochner's theorem provides an equivalent characterisation of whether or not a given continuous function $f$ is positive definite. The concept of positive definite functions was extended to positive definite distributions by L. Schwartz [41, Chapter VII, §9], and his analogue of Bochner's theorem states that a distribution is positive definite (and tempered) if and only if it is the Fourier transform of a non-negative measure of slow increase, i.e. such that the measure of balls is polynomially bounded in terms of the radius.

As shown above, positive definite functions in the sense of the standard definition (1) are always bounded by their value at zero. However, there exist functions such as $f = |\cdot|^{-\alpha}$ $(0 < \alpha < 1)$, which have a singularity at the origin, yet still exhibit properties similar to those of positive definite functions. Such functions arise in potential theory (see, e.g. [8], [21] and [33]), and recently appeared in the context of extremal measures ([36], [37]). Functions defined on $\mathbb{R}$ which are unbounded at the origin and positive definite in the following extended sense were studied by Cooper [14].

**Definition.** A function $f : \mathbb{R} \to \mathbb{C}$ is called positive definite w.r.t. a set $J$ of functions if for every $\phi \in J$, the integral

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f(x-y)\phi(x)\overline{\phi(y)}\, dxdy \tag{2}$$

exists (in the Lebesgue sense) and is non-negative [14, p. 54].

Let $\mathrm{P}(J)$ denote the class of all functions which are positive definite w.r.t. $J$. Cooper's definition coincides with the standard definition (1) when $f$ is continuous and $J = \mathrm{C}_0(\mathbb{R})$, see e.g. [14, p.53]. However, for certain spaces of functions $J$, the above definition enables us to extend the concept of positive definiteness to functions which have a singularity at the origin. In particular, we shall consider the spaces $J = \mathrm{L}^p(\mathbb{R}^d)$ (and their local versions) for various values of $p$.

Building on the foundations set by Cooper, we study unbounded positive definite functions in more detail. One of our main results is Theorem 2.5.1, which, in analogy to Bochner's theorem for the classical case, characterises a larger class of (generally unbounded) positive definite functions. Various results follow from this theorem. For example, functions which are positive definite w.r.t. $\mathrm{L}^2(\mathbb{R}^d)$ can be approximated, in the $\mathrm{L}^1(\mathbb{R}^d)$ sense, by a sequence of continuous, classically positive definite functions (see Corollary 2.6.1). Functions which arise as 'convolution squares' are positive definite in the new sense (see Corollary 2.6.4), and conversely, a function which is positive definite w.r.t. $\mathrm{L}^2(\mathbb{R}^d)$ can be written, in some sense, as a convolution square (see Corollary 2.6.5). We also show that the even reflections of locally integrable, completely monotone functions are positive definite w.r.t. $\mathrm{L}^2_0(\mathbb{R})$, the set of functions in $\mathrm{L}^2(\mathbb{R})$ with compact essential support (see Corollary 2.7.6), and, subsequently, provide many examples of unbounded functions which are positive definite in the extended sense.

A comprehensive study of conditionally negative definite functions, which appear naturally in both probability and potential theory, can be found in [3], [4] and [35]. A function $f : \mathbb{R}^d \to \mathbb{C}$ is conditionally negative definite in the standard sense if $f$ is conjugate symmetric, that is $f(\mathbf{x}) = \overline{f(-\mathbf{x})}$ for all

$\mathbf{x} \in \mathbb{R}^d$, and

$$\sum_{i,j=1}^{n} f(\mathbf{x}_i - \mathbf{x}_j)\, v_i\, \overline{v_j} \le 0 \tag{3}$$

for all $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ and $v_1, v_2, \ldots, v_n \in \mathbb{C}$ satisfying $\sum_{i=1}^n v_i = 0$, with any $n \in \mathbb{N}$.

It is clear that if $f$ is a classically positive definite function, then $-f$ is conditionally negative definite. The converse is not necessarily true (see Section 3.1 for details). Similarly to positive definite functions, conditionally negative definite functions take finite values at zero. However, unlike their positive counterparts, they need not be bounded away from the origin.

Motivated by Cooper's definition of positive definiteness, we define an extended notion of conditionally negative definite functions on $\mathbb{R}^d$ as follows.

**Definition.** A function $f : \mathbb{R}^d \to \mathbb{C}$ is said to be conditionally negative definite w.r.t. a set $J$ of functions if $f$ is conjugate symmetric a.e., that is $f(\mathbf{x}) = \overline{f(-\mathbf{x})}$ f.a.a. $\mathbf{x} \in \mathbb{R}^d$, and for every $\phi \in J$ satisfying $\int_{\mathbb{R}^d} \phi(\mathbf{x})\, d\mathbf{x} = 0$,

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{x} - \mathbf{y}) \phi(\mathbf{x}) \overline{\phi(\mathbf{y})}\, d\mathbf{x} d\mathbf{y} \le 0.$$

Let $\mathrm{CN}(J)$ denote the class of all functions which are conditionally negative definite w.r.t. the set $J$. For suitably chosen $J$, $\mathrm{CN}(J)$ contains the classical conditionally negative definite functions, which take finite values at zero, as well as functions which are singular at the origin (see Section 3.5).

The renowned Schoenberg theorem [39, Th. 2] establishes a relation between the classical positive definite and conditionally negative definite functions. In particular, it states that a function $f : \mathbb{R}^d \to \mathbb{C}$ is conditionally negative definite if and only if for all $t > 0$, $g : \mathbf{x} \mapsto e^{-tf(\mathbf{x})}$ is positive definite.

The central result of Section 3 is Theorem 3.4.1, which is a generalisation of Schoenberg's theorem to real-valued (generally unbounded) functions in $\mathrm{P}(J)$ and $\mathrm{CN}(J)$, for $J = \mathrm{L}_0^2(\mathbb{R}^d)$. Several subsequent results concerning the class $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ are also established. For instance, we demonstrate that functions in $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ are locally integrable (see Lemma 3.3.1), that $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ is a closed subset of $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ (see Lemma 3.3.2) and that real-valued functions which are conditionally negative definite w.r.t. $\mathrm{L}_0^2(\mathbb{R}^d)$ can be approximated, in the $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ sense, by a sequence of infinitely differentiable, classically conditionally negative definite functions (see Lemma 3.4.3). Moreover, using Theorem 3.4.1, we indicate how to construct numerous examples of singular functions $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$, such that $-f \notin \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$ (see Section 3.5).

The classical energy problem in the field of potential theory is based on finding the measure(s) $\mu \in \mathcal{M}(\mathcal{X})$ which minimise(s)/maximise(s) the energy integral

$$I_f(\mu) := \int_{\mathcal{X}} \int_{\mathcal{X}} f(\mathbf{x}-\mathbf{y})\mu(d\mathbf{x})\mu(d\mathbf{y}) \tag{4}$$

for a given a function $f : \mathbb{R}^d \to \mathbb{R}$ and set $\mathcal{M}(\mathcal{X})$ of non-negative unit Borel measures with support in the compact set $\mathcal{X} \subset \mathbb{R}^d$. We define a minimising measure $\mu^* \in \mathcal{M}(\mathcal{X})$ to be optimal if it is the unique minimiser for $I_f$.

The case when $f = -\log|\cdot|$ has been widely studied, see e.g. [33], [34]. In fact, the logarithmic energy problem has been solved for various sets $\mathcal{X}$, see e.g. [33, Chapter I.1], where circles, discs and line segments are considered. The more general case (for $d = 1$) when

$$f(x) = \begin{cases} (1-|x|^{\alpha-1})/(\alpha-1) & \text{if} \quad \alpha \neq 1, \\ -\log|x| & \text{if} \quad \alpha = 1 \end{cases} \qquad (x \in \mathbb{R}\backslash\{0\})$$

has been studied in [37], where it is shown that for any $\alpha \in (0,2)$, $I_f(\mu)$ is strictly convex on the set of all probability measures on the set of Borel subsets of $[0,1]$, and that the measure with generalised arcsine density,

$$p_{1-\alpha/2}(t) = \frac{\Gamma\big(2-\alpha\big)\, t^{-\alpha/2}\,(1-t)^{-\alpha/2}}{\Gamma^2\big(1-\alpha/2\big)},$$

is optimal for $I_f$, see [37, Th. 2]. The energy problem has also been greatly considered in the literature for the Riesz kernel $\kappa_\alpha(\mathbf{x},\mathbf{y}) = |\mathbf{x}-\mathbf{y}|^{\alpha-d}$ ($0 < \alpha < d$, $\mathbf{x},\mathbf{y} \in \mathbb{R}^d$) and the classical Newtonian kernel, i.e. when $f(\mathbf{x}) = |\mathbf{x}|^{2-d}$ ($d > 2$, $\mathbf{x} \in \mathbb{R}^d$), see e.g. [16], [21], [28], [32]. In the non-singular case, when $f = |\cdot|^\alpha$ ($\alpha > 0$), properties of the maximising measures and their potentials

$$\mathcal{P}_\mu(\mathbf{y}) := \int_{\mathcal{X}} f(\mathbf{x}-\mathbf{y})\,\mu(d\mathbf{x}) \quad (y \in \mathcal{X})$$

have been explored in [8].

The problem of finding optimal designs in experimental design is very closely related to the energy problem in potential theory. In particular, the functional $I_f(\mu)$ arises as an optimality criterion in the optimal design problem with correlated observations for the location model $y_j = \theta + \epsilon_j$, see e.g. [47, Eq. 5]. The measure $\mu^*$ that minimises $I_f$ on the set of probability measures defined on a compact subset of $\mathbb{R}$, say $[0, 1]$, defines an optimal design for a suitable correlation function $f$. Standard correlation functions are positive definite in the classical sense, however, as in [37, Corollary 1], we extend the optimal design problem to the case when $f$ is singular at the origin and positive definite in an extended sense (see Section 4 for details).

We study the energy integral $I_{f(|\cdot|)}$ for several singular, completely monotone functions $f$, which have not previously been considered in the field of potential theory, and solve the corresponding energy problems by means of numerically computing densities of measures which minimise $I_{f(|\cdot|)}$. The main result of Section 4 is an algorithm for constructing continuous probability measures which approximate the minimising measures for such energies $I_{f(|\cdot|)}$. Firstly, we consider the case when $f = (\cdot)^{-\alpha}$ ($\alpha \in (0, 1)$), which corresponds to the case of the Riesz kernel (see Section 4.5), and, secondly, when $f$ is replaced with a variety of alternative completely monotone functions with a singularity at zero (see Section 4.6).

# 2 On Unbounded Positive Definite Functions

Positive definite functions are bounded, taking their maximum absolute value at 0. Nevertheless, there are unbounded functions, arising e.g. in potential theory or the study of (continuous) extremal measures, which still exhibit the general characteristics of positive definiteness. Taking a framework set up by Lionel Cooper [14] as a motivation, we study the general properties of such functions which are positive definite in an extended sense.

Our central result is Theorem 2.5.1, which, in analogy to Bochner's theorem for the classical case, characterises a larger class of (generally unbounded) positive definite functions. Numerous results follow from this theorem. For instance, functions which are positive definite w.r.t. $\mathrm{L}^2(\mathbb{R}^d)$ can be approximated, in the $\mathrm{L}^1(\mathbb{R}^d)$ sense, by a sequence of continuous, classically positive definite functions (see Corollary 2.6.1). Functions which arise as 'convolution squares' are positive definite in the new sense (see Corollary 2.6.4), and conversely, a function which is positive definite w.r.t. $\mathrm{L}^2(\mathbb{R}^d)$ can be written, in some sense, as a convolution square (see Corollary 2.6.5). Using Theorem 2.5.1, we also show that the even reflections of locally integrable, completely monotone functions are positive definite w.r.t. $\mathrm{L}^2_0(\mathbb{R})$, the set of functions in $\mathrm{L}^2(\mathbb{R})$ with compact essential support (see Corollary 2.7.6). This result provides many examples of functions which have a singularity at zero and are positive definite in the extended sense. The findings outlined in this section have been accepted to appear in Mathematica Pannonica in the form of the paper [26].

The structure of this section is as follows. We begin with an overview of the positive definite functions as defined in the classical literature, see Section 2.1. In Section 2.3 we introduce the ideas and discuss the main results of [14]. The proof of Theorem 2.5.1 is given in Section 2.5, and an alternative proof can be found in Section 2.8. Sections 2.6 and 2.7 present corollaries to Theorem 2.5.1 and their corresponding proofs. We conclude Section 2.7 with several examples of unbounded positive definite functions.

## 2.1 Classical positive definite functions

Positive definite sequences, arising naturally in the context of a problem in complex function theory posed by Carathéodory [13], were first introduced in 1911 by Toeplitz [44]. Motivated by the work of Carathéodory and Toeplitz, Mathias [23] and subsequently, Bochner [9] defined and studied the properties of positive definite functions, specifically their harmonic analysis. Prior to these developments, however, Mercer [24] had studied the more general concept of positive definite kernels in research on integral equations.

The standard definition of a positive definite function is as follows.

**Definition 2.1.1** A function $f : \mathbb{R}^d \to \mathbb{C}$ is *positive definite* if

$$\sum_{i,j=1}^{n} f(\mathbf{x}_i - \mathbf{x}_j)\, v_i\, \overline{v_j} \geq 0 \tag{5}$$

for all $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ and $v_1, v_2, \ldots, v_n \in \mathbb{C}$, with any $n \in \mathbb{N}$.

In other words, $f$ is positive definite if the matrix $\left[f(\mathbf{x}_i - \mathbf{x}_j)\right]_{i,j=1}^{n}$ is non-negative definite for all $n \in \mathbb{N}$ and $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$. We shall denote the set of classical positive definite functions on $\mathbb{R}^d$ by $\mathrm{P}_{\mathrm{C},d}$. A simple example of a function in $\mathrm{P}_{\mathrm{C},1}$ is $f = e^{\pm i(\cdot)}$, since for any $x_1, x_2, \ldots, x_n \in \mathbb{R}$ and $v_1, v_2, \ldots, v_n \in \mathbb{C}$,

$$\sum_{i,j=1}^{n} f(x_i - x_j)\, v_i\, \overline{v_j} = \left| \sum_{j=1}^{n} e^{\pm i\, x_j} v_j \right|^2 \geq 0.$$

Setting $n = 2,\ \mathbf{x}_1 = \mathbf{0},\ \mathbf{x}_2 = \mathbf{x},\ v_1 = 1$ and $v_2 = v$ in (5) gives

$$\left(1 + |v|^2\right) f(\mathbf{0}) + v f(\mathbf{x}) + \overline{v} f(-\mathbf{x}) \geq 0 \tag{6}$$

for any $v \in \mathbb{C}$. Hence, $v f(\mathbf{x}) + \overline{v} f(-\mathbf{x})$ is real for any $v \in \mathbb{C}$, and thus, in particular, both $f(\mathbf{x}) + f(-\mathbf{x})$ and $i(f(\mathbf{x}) - f(-\mathbf{x}))$ are real. It therefore follows that functions in $\mathrm{P}_{\mathrm{C},d}$ are conjugate symmetric, that is $f(\mathbf{x}) = \overline{f(-\mathbf{x})}$ for all $\mathbf{x} \in \mathbb{R}^d$. Choosing $v$ in (6) such that $v f(\mathbf{x}) = -|f(\mathbf{x})|$ gives $|f(\mathbf{x})| \leq f(\mathbf{0})$ for all $\mathbf{x} \in \mathbb{R}^d$. Hence, positive definite functions by the standard definition are bounded, taking their maximum absolute value at the origin.

However, a positive definite function in the classical sense need not be positive or continuous; for example, the function $f(\mathbf{x}) = 1$ if $\mathbf{x} = \mathbf{0},\ f(\mathbf{x}) = 0$ otherwise ($\mathbf{x} \in \mathbb{R}^d$), is positive definite, but not continuous; the cosine function, $\cos x = (e^{ix} + e^{-ix})/2$, is in $\mathrm{P}_{\mathrm{C},1}$, but not non-negative. The positive definiteness of the cosine function follows from property ii. below, which we list amongst two other simple properties of functions in $\mathrm{P}_{\mathrm{C},d}$.

i. $f \in \mathrm{P}_{\mathrm{C},d}$ if and only if $\overline{f} \in \mathrm{P}_{\mathrm{C},d}$.

ii. If $f_1, f_2, \ldots, f_n \in \mathrm{P}_{\mathrm{C},d}$ and $c_i \geq 0$ for all $i = 1, \ldots, n$, then $\sum_{i=1}^{n} c_i f_i \in \mathrm{P}_{\mathrm{C},d}$.

iii. If $f_n \in \mathrm{P}_{\mathrm{C},d}$ for all $n \in \mathbb{N}$ and the pointwise limit, $\lim_{n\to\infty} f_n(\mathbf{x}) = f(\mathbf{x})$, exists for all $\mathbf{x} \in \mathbb{R}^d$, then $f \in \mathrm{P}_{\mathrm{C},d}$.

These properties follow directly from Definition 2.1.1. A simple consequence of the first two properties is that if $f$ is positive definite, then so is $\mathrm{Re}\,(f) = (f + \overline{f})/2$.

Another interesting result of the class $\mathrm{P}_{\mathrm{C},d}$ is that it is closed under pointwise products. This property is stated in the following proposition and proved directly below.

**Proposition 2.1.1** *If $f$, $g \in \mathrm{P}_{\mathrm{C},d}$, then $fg \in \mathrm{P}_{\mathrm{C},d}$.*

*Proof.* Let $n \in \mathbb{N}$ and $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$. The Schur (or Hadamard) product of two $n \times n$ matrices $A = (a_{i\,j}) = \left[f(\mathbf{x}_i - \mathbf{x}_j)\right]_{i,j=1}^n$ and $B = (b_{i\,j}) = \left[g(\mathbf{x}_i - \mathbf{x}_j)\right]_{i,j=1}^n$ is the $n \times n$ matrix $C$ with entries $c_{i\,j} = a_{i\,j}\, b_{i\,j} = fg\,(\mathbf{x}_i - \mathbf{x}_j)$. We now show that since $A$ and $B$ are non-negative definite, then $C$ is also non-negative definite. This result is often referred to as Schur's product theorem [40, Th. VII].

Since $f \in \mathrm{P}_{\mathrm{C},d}$, then $A$ is Hermitian, that is $a_{i\,j} = \overline{a_{j\,i}}$, and non-negative definite. Hence, $A$ has a Cholesky decomposition of the form $A = L\,L^*$, where $L$ is a lower triangular matrix with non-negative diagonal entries, and $L^*$ denotes the conjugate transpose of $L$. Thus, for any $v_1, v_2, \ldots, v_n \in \mathbb{C}$,

$$\begin{aligned}\sum_{i,j=1}^{n} fg\,(\mathbf{x}_i - \mathbf{x}_j)\, v_i\, \overline{v_j} &= \sum_{i,j=1}^{n} \left( \sum_{k=1}^{n} l_{i\,k}\, \overline{l_{j\,k}} \right) b_{i\,j}\, v_i\, \overline{v_j} \\ &= \sum_{k=1}^{n} \left( \sum_{i,j=1}^{n} b_{i\,j}\, (v_i\, l_{i\,k})\, \overline{(v_j\, l_{j\,k})} \right) \geq 0. \;\square\end{aligned}$$

Hence, if $f$ is positive definite, then so is $|f|^2 = \overline{f}f$.

For real-valued functions we can use the following alternative definition of positive definiteness.

**Definition 2.1.2** A function $f : \mathbb{R}^d \to \mathbb{R}$ is positive definite if $f(\mathbf{x}) = f(-\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^d$ and the inequality in (5) holds for all $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ and $v_1, v_2, \ldots, v_n \in \mathbb{R}$, with any $n \in \mathbb{N}$.

In particular, the evenness of the function is now stipulated, since it no longer follows automatically as it does in Definition 2.1.1. This is also true in the case of positive definite matrices; for example, the real-valued $2 \times 2$ matrix

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

is non-symmetric yet positive definite, since for any $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$, $\mathbf{v} A\, \mathbf{v}^T = (v_1 + v_2)^2 \geq 0$.

It can easily be shown that the functions defined in Definition 2.1.2 are immediately positive definite as in Definition 2.1.1. For instance, for $n \in \mathbb{N}$, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ and $v_1, v_2, \ldots, v_n \in \mathbb{C}$,

$$\begin{aligned} \sum_{j,k=1}^{n} f(\mathbf{x}_j - \mathbf{x}_k)\, v_j\, \overline{v_k} &= \sum_{j,k=1}^{n} f(\mathbf{x}_j - \mathbf{x}_k)(a_j + i\, b_j)\overline{(a_k + i\, b_k)} \\ &= \sum_{j,k=1}^{n} f(\mathbf{x}_j - \mathbf{x}_k)\, a_j\, a_k + \sum_{j,k=1}^{n} f(\mathbf{x}_j - \mathbf{x}_k)\, b_j\, b_k \geq 0, \quad (7) \end{aligned}$$

where $a_j = \mathrm{Re}\,(v_j)$ and $b_j = \mathrm{Im}\,(v_j)$.

The notion of positive definite functions can be generalised to functions and kernels defined on arbitrary topological spaces, groups and semigroups, see e.g. [2, 35, 43]. For example, a real-valued, positive definite kernel on $\mathbb{R}^d \times \mathbb{R}^d$ can be defined as follows.

**Definition 2.1.3** A kernel $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is said to be positive definite if $k$ is symmetric, that is $k\,(\mathbf{x}, \mathbf{y}) = k\,(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, and

$$\sum_{i,j=1}^{n} k\,(\mathbf{x}_i, \mathbf{x}_j)\, v_i\, v_j \geq 0 \tag{8}$$

for all $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ and $v_1, v_2, \ldots, v_n \in \mathbb{R}$, with any $n \in \mathbb{N}$.

We mainly restrict our attention to the positive definite functions defined in Definition 2.1.1. Henceforth, unless clearly stated otherwise, when referring to classically positive definite functions or functions in $\mathrm{P}_{\mathrm{C},d}$, we mean those defined in Definition 2.1.1.

One of the central results on the subject of positive definite functions is the following theorem of Bochner.

**Theorem 2.1.1** (Bochner, [9, Chapter IV.20]). *A function $f : \mathbb{R}^d \to \mathbb{C}$ is continuous and positive definite if and only if it is the (inverse) Fourier transform of a finite, non-negative measure $\mu$ on $\mathbb{R}^d$, i.e.*

$$f(\mathbf{x}) = \check{\mu}(\mathbf{x}) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{i\mathbf{x}\cdot\mathbf{z}}\, \mu(d\mathbf{z}) \qquad (\mathbf{x} \in \mathbb{R}^d). \tag{9}$$

One direction is trivial, for if $f$ has the above form, then for any $n \in \mathbb{N}$, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}$ and $v_1, v_2, \ldots, v_n \in \mathbb{C}$,

$$\begin{aligned}(2\pi)^{\frac{d}{2}} \sum_{j,k=1}^{n} f(\mathbf{x}_j - \mathbf{x}_k)\, v_j\, \overline{v_k} &= \sum_{j,k=1}^{n} \left( \int_{\mathbb{R}^d} e^{\,i\mathbf{x}_j\cdot\mathbf{z}}\, \overline{e^{\,i\mathbf{x}_k\cdot\mathbf{z}}}\, \mu(d\mathbf{z}) \right) v_j\, \overline{v_k} \\ &= \int_{\mathbb{R}^d} \left| \sum_{j=1}^{n} e^{\,i\mathbf{x}_j\cdot\mathbf{z}}\, v_j \right|^2 \mu(d\mathbf{z}) \geq 0.\end{aligned}$$

For a proof of the reverse implication see e.g [20, p. 150], [9, Chapter IV.20].

Hence, Bochner's theorem provides an equivalent characterisation of whether or not a given continuous function $f$ is positive definite. For example, the continuous functions $f_1(x) = \sqrt{2}\, e^{-x^2}$ and $f_2(x) = \sqrt{2}\,(1 + x^2)^{-1}$ ($x \in \mathbb{R}$) are in $\mathrm{P}_{\mathrm{C},1}$, since they are the (inverse) Fourier transforms of $\hat{f}_1 = e^{-\frac{x^2}{4}}$ and $\hat{f}_2 = \sqrt{\pi}\, e^{-|x|}$, respectively.

For a continuous function $f \in \mathrm{P}_{\mathrm{C},d}$, the finiteness of the measure in Bochner's theorem emphasizes the fact that $f$ is non-singular at the origin, since $f(\mathbf{0}) = \mu(\mathbb{R}^d)$. However, there exist functions such as $f = |\cdot|^{-\alpha}$ ($0 < \alpha < 1$), which are unbounded at the origin, yet still exhibit properties similar to those of positive definite functions. Such functions arise naturally in potential theory (see, e.g. [8], [21] and [33]), and recently appeared in the context of extremal measures ([36], [37]). Functions of a real variable which are unbounded at the origin and positive definite in an extended sense were studied by Cooper [14], see Section 2.3 for details.

## 2.2 Completely monotone functions

The concept of complete monotonicity was first introduced by Bernstein [6], who studied functions on intervals of the real line having positive derivatives of all orders. In conjunction to presenting the definition and basic properties of completely monotone functions, we highlight two theorems of note. In Section 2.7 we look to partially extend the second of these theorems, Theorem 2.2.2, to functions with a singularity at zero.

**Definition 2.2.1** A function $f : (0, \infty) \to [0, \infty)$ is *completely monotone* if $f \in \mathrm{C}^{\infty}((0, \infty))$ and

$$(-1)^n\, f^{(n)} \geq 0 \;\text{ on } (0,\, \infty)$$

for all $n \in \mathbb{N}_0$ [35, Def. 1.3].

In particular, any completely monotone function is non-negative and non-increasing. The family of all completely monotone functions is denoted by CM. Functions in CM can be bounded or unbounded at zero; for example, both $f_1(x) = e^{-x}$ and $f_2(x) = x^{-\frac{1}{2}}$ $(x \in (0, \infty))$ are completely monotone. If $f$ is a bounded completely monotone function, then it can be extended continuously to $[0, \infty)$ by taking $f(0) := f(0+) = \lim_{x \to 0} f(x)$ [35, p. 28].

The following theorem, given without proof, characterises functions in CM as Laplace transforms of non-negative measures.

**Theorem 2.2.1** (Bernstein, [35, Th. 1.4], [7]). *Let $f : (0, \infty) \to [0, \infty)$ be a completely monotone function. Then it is the Laplace transform of a unique, non-negative measure $\mu$ on $[0, \infty)$, i.e. for all $x > 0$,*

$$f(x) = \mathcal{L}(\mu\,; x) = \int_{[0,\infty)} e^{-xt}\, \mu(dt). \tag{10}$$

*Conversely, whenever $\mathcal{L}(\mu\,; x) < \infty$ for any $x > 0$, $x \mapsto \mathcal{L}(\mu\,; x)$ is completely monotone.*

Theorem 2.2.1 holds for all completely monotone functions, bounded and unbounded. It follows from (10) that $f(0+) = \mu([0, \infty))$. Hence, the finiteness of the measure is directly related to whether or not the function is bounded. In particular, if $f \in$ CM is bounded, then $\mu$ is finite; conversely, if $f \in$ CM is unbounded, then $\mu([0, \infty)) = +\infty$.

We list some useful properties of completely monotone functions below.

i. If $f_1, f_2, \ldots, f_n \in$ CM and $c_i \geq 0$ for all $i = 1, \ldots, n$, then $\sum_{i=1}^{n} c_i f_i \in$ CM.

ii. If $f_n \in$ CM for all $n \in \mathbb{N}$ and the pointwise limit, $\lim_{n \to \infty} f_n(x) = f(x)$, exists for all $x > 0$, then $f \in$ CM.

iii. If $f$, $g \in$ CM, then $fg \in$ CM.

The first property follows directly from Definition 2.2.1, whilst properties ii. and iii. can be proved using Theorem 2.2.1, see e.g. [35, Corollary 1.6].

The following theorem belongs to Schoenberg, along with a number of other theorems on classically positive definite functions, e.g. [35, Prop. 4.4], [35, Th. 12.14], [5, Th. 1.6].

**Theorem 2.2.2** (Schoenberg, [38, Th. 3]). *A function $\psi : [0, \infty) \to [0, \infty)$ is a bounded completely monotone function if and only if for all $d \in \mathbb{N}$, the function $f = \psi(\| \cdot \|^2) : \mathbb{R}^d \to [0, \infty)$ is continuous and positive definite.*

In particular, if $\psi \in \mathrm{CM}$ is bounded, then $f = \psi(\|\cdot\|^2) : \mathbb{R}^d \to [0, \infty)$ is continuous and positive definite for any $d \in \mathbb{N}$. We generalise this observation to potentially unbounded completely monotone functions in Corollary 2.7.1.

## 2.3 Positive definiteness in the extended sense

An extended notion of positive definiteness was introduced by Cooper in the pioneering paper [14]. The definition of a positive definite function in [14] is more general than in Definition 2.1.1, and accounts for functions which are unbounded at the origin. In this section we highlight the main results of [14], where all of the functions considered are defined on the real line.

For continuous functions in $\mathrm{P}_{\mathrm{C},1}$, (5) is equivalent to

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f(x-y)\phi(x)\overline{\phi(y)}\, dxdy \geq 0 \tag{11}$$

for all functions $\phi \in \mathrm{C}_0(\mathbb{R})$, see e.g. [14, p.53]. Motivated by this observation, Cooper [14] used the following definition for functions which are positive definite in an extended sense.

**Definition 2.3.1** A function $f : \mathbb{R} \to \mathbb{C}$ is called *positive definite w.r.t. a set* $J$ of functions if for every $\phi \in J$, the integral in (11) exists (in the Lebesgue sense) and is non-negative [14, p. 54].

Let $\mathrm{P}(J)$ denote the class of all functions which are positive definite w.r.t. the set $J$. In association with [43, Section 6], such functions may also be called *integrally positive definite* for $J$. For certain spaces of functions $J$, Cooper's definition enables us to extend the concept of positive definiteness to functions which have a singularity at 0. In particular, we shall consider the spaces $J = \mathrm{L}^p(\mathbb{R})$ (and their local versions) for various values of $p$. In the following section we will also consider spaces of functions defined on $\mathbb{R}^d$.

We begin with an overview of some basic properties of the positive definite functions studied in [14], analogous to those for the classical case, see [43, p. 412]. In the following, let $J$ be a set of complex-valued measurable functions on $\mathbb{R}$. This includes functions defined on a non-empty, measurable subset of $\mathbb{R}$, which we consider to be extended by zero to the whole real line. Then the following properties follow directly from Definition 2.3.1.

i. $f \in \mathrm{P}(J) \Leftrightarrow f^* \in \mathrm{P}(J)$, where $f^*(x) := \overline{f(-x)}$ $(x \in \mathbb{R})$.

ii. $f \in \mathrm{P}(J) \Leftrightarrow \overline{f} \in \mathrm{P}(J)$ if $J$ is closed under complex conjugation.

iii. If $f_1, f_2, \ldots, f_n \in \mathrm{P}(J)$ and $c_i \geq 0$ $(i = 1, \ldots, n)$, then $\sum_{i=1}^{n} c_i f_i \in \mathrm{P}(J)$.

iv. If $J_1 \subseteq J_2$, then $\mathrm{P}(J_2) \subseteq \mathrm{P}(J_1)$.

Before proceeding to present our new results, we highlight the most relevant results of [14].

For $p \in [1, \infty) \cup \{\infty\}$, let $\mathrm{L}^p_0(\mathbb{R})$ denote the subspace of functions in $\mathrm{L}^p(\mathbb{R})$ with compact essential support. The functions in $\mathrm{P}(\mathrm{L}^1_0(\mathbb{R}))$ are essentially bounded [14, Th. 5] and almost everywhere equal to a continuous, positive definite function in the classical sense [15, Sec. 6]. The functions in $\mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$ need only be locally integrable [14, Lemma 1]. Cooper has the following Bochner-type theorem.

**Theorem 2.3.1** (Cooper, [14, Th. 6]). *For any function $f \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$, there exists a non-negative, non-decreasing function $\rho$, such that for almost all $x$,*

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ixt}\, d\rho(t) \qquad \textit{in } (\mathrm{C},\, 1) \textit{ sense,} \tag{12}$$

*where $\rho(t) = o\,(t)$ as $t \to \pm\infty$.*

In particular, the function $\rho$ need not be bounded, but satisfies a sublinear growth condition at $\pm\infty$. Note also that, unlike Bochner's theorem, the implication here is only in one direction. The qualification "in (C, 1) sense" in (12) means

$$f(x) = \frac{1}{\sqrt{2\pi}} \lim_{\lambda \to \infty} \frac{1}{\lambda} \int_0^{\lambda} \left( \int_{-u}^{u} e^{ivx}\, d\rho(v) \right) du,$$

in analogy to Cesàro summation of divergent series.

The $\mathrm{P}(\mathrm{L}^p_0(\mathbb{R}))$ spaces have the following additional properties.

**Proposition 2.3.1** *If $f \in \mathrm{P}_{\mathrm{C},1}$ is continuous, then $f \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$.*

*Proof.* By Bochner's theorem, there exists a finite, non-negative measure $\mu$ on $\mathbb{R}$ such that for any $\phi \in \mathrm{L}^2_0(\mathbb{R})$,

$$\begin{aligned} \int_{\mathbb{R}} \int_{\mathbb{R}} f(x-y)\phi(x)\overline{\phi(y)}\, dxdy &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{i(x-y)z}\, \mu(dz)\, \phi(x)\overline{\phi(y)}\, dxdy \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left| \int_{\mathbb{R}} e^{ixz} \phi(x)\, dx \right|^2 \mu(dz) \geq 0. \end{aligned}$$

□

**Proposition 2.3.2** *If $f \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$ and $g \in \mathrm{P}_{\mathrm{C},1}$ is continuous, then $fg \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$* [14, Th. 1].

*Proof.* By Bochner's theorem, there exists a finite, non-negative measure $\mu$ on $\mathbb{R}$ such that for any $\phi \in \mathrm{L}^2_0(\mathbb{R})$,

$$\begin{aligned}\int_{\mathbb{R}}\int_{\mathbb{R}} fg\,(x-y)\phi(x)\overline{\phi(y)}\,dxdy& \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\int_{\mathbb{R}} f(x-y)\int_{\mathbb{R}} e^{i(x-y)z}\,\mu(dz)\,\phi(x)\overline{\phi(y)}\,dxdy \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} f(x-y)(e^{ixz}\phi(x))(\overline{e^{iyz}\phi(y)})\,dxdy\,\mu(dz)\geq 0.\end{aligned}$$

□

**Proposition 2.3.3** *For any $p \in [1,2]$, $\mathrm{P}(\mathrm{L}^p_0(\mathbb{R})) \subseteq \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$.*

*Proof.* This follows directly from the fact that $\mathrm{L}^2_0(\mathbb{R}) \subseteq \mathrm{L}^p_0(\mathbb{R})$ ($p \in [1,2]$). □

**Proposition 2.3.4** *For any $q \in (2,\infty]$ and $r \in [0,\infty]$, $\mathrm{P}(\mathrm{L}^2_0(\mathbb{R})) = \mathrm{P}(\mathrm{L}^q_0(\mathbb{R})) = \mathrm{P}(\mathrm{C}^r_0(\mathbb{R}))$.*

*Proof.* Let $q \in (2,\infty]$ and $r \in [0,\infty]$. Since $\mathrm{C}^r_0(\mathbb{R}) \subset \mathrm{L}^q_0(\mathbb{R}) \subset \mathrm{L}^2_0(\mathbb{R})$, it follows directly that $\mathrm{P}(\mathrm{L}^2_0(\mathbb{R})) \subset \mathrm{P}(\mathrm{L}^q_0(\mathbb{R})) \subset \mathrm{P}(\mathrm{C}^r_0(\mathbb{R}))$. For the reverse implication, we shall use the density of $\mathrm{C}^r_0(\mathbb{R})$ in $\mathrm{L}^2_0(\mathbb{R})$. Suppose that $f \in \mathrm{P}(\mathrm{C}^r_0(\mathbb{R}))$. Then, the integral

$$\int_{\mathbb{R}}\int_{\mathbb{R}} f(x-y)\psi(x)\overline{\psi(y)}\,dxdy = \int_{\mathbb{R}} (f^{*}\!*\psi)(y)\overline{\psi(y)}\,dy$$

exists in the Lebesgue sense and is non-negative for all $\psi \in \mathrm{C}^r_0(\mathbb{R})$, where $f^*(y) = \overline{f(-y)}$ ($y \in \mathbb{R}$). Since $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$ by [14, Lemma 1], and the convolution of an element of $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$ with an element of $\mathrm{L}^2_0(\mathbb{R})$ is in $\mathrm{L}^2(\mathbb{R})$, the integral also exists for all $\psi \in \mathrm{L}^2_0(\mathbb{R})$. By a change of variables and the Fubini theorem,

$$\int_{\mathbb{R}}\int_{\mathbb{R}} f(x-y)\psi(x)\overline{\psi(y)}\,dxdy = \int_{\mathbb{R}}\int_{\mathbb{R}} f(z)\psi(z+y)\overline{\psi(y)}\,dydz \quad (\psi \in \mathrm{L}^2_0(\mathbb{R})).$$

Let $\phi \in \mathrm{L}^2_0(\mathbb{R})$; then there is a sequence $(\psi_n)_{n\in\mathbb{N}}$ in $\mathrm{C}^r_0(\mathbb{R})$ such that $\|\phi - \psi_n\|_2 \to 0$ as $n \to \infty$. Now,

$$\begin{aligned}
&\sup_{z\in\mathbb{R}} \left| \int_{\mathbb{R}} \Big( \phi(z+y)\overline{\phi(y)} - \psi_n(z+y)\overline{\psi_n(y)} \Big) dy \right| \\
&\qquad \le \sup_{z\in\mathbb{R}} \int_{\mathbb{R}} |\phi(z+y)||(\phi-\psi_n)(y)|dy + \sup_{z\in\mathbb{R}} \int_{\mathbb{R}} |(\phi-\psi_n)(z+y)||\psi_n(y)|dy \\
&\qquad \le \|\phi\|_2 \, \|\phi-\psi_n\|_2 + \|\phi-\psi_n\|_2 \, \|\psi_n\|_2 \to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

As $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$, it follows that

$$\left| \int_K f(z) \int_{\mathbb{R}} \Big( \phi(z+y)\overline{\phi(y)} - \psi_n(z+y)\overline{\psi_n(y)} \Big) \, dydz \right| \to 0 \quad \text{as } n \to \infty,$$

where $K$ denotes an interval that contains all the supports of the functions $g_n(z) = \int_{\mathbb{R}} (\phi(z+y)\overline{\phi(y)} - \psi_n(z+y)\overline{\psi_n(y)}) \, dy$. Hence the result follows, for

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f(x-y)\phi(x)\overline{\phi(y)} \, dxdy = \lim_{n\to\infty} \int_{\mathbb{R}} \int_{\mathbb{R}} f(x-y)\psi_n(x)\overline{\psi_n(y)} \, dxdy \ \ge 0. \ \square$$

The last two propositions demonstrate that as $p$ increases from 1 to 2, $\mathrm{P}(\mathrm{L}^p_0(\mathbb{R}))$ increases from a smaller class of positive definite functions to a larger such class. As $p$ increases beyond 2, $\mathrm{P}(\mathrm{L}^p_0(\mathbb{R}))$ remains the same. Moreover, roughly speaking, as $p$ increases from 1 to 2, $\mathrm{P}(\mathrm{L}^p_0(\mathbb{R}))$ runs from the class of bounded, continuous positive definite functions (in the standard sense), to a class of functions which are positive definite in a wider sense and need not be bounded or continuous.

It proceeds from the following theorem of Cooper [14] that if a function is bounded and positive definite w.r.t. $\mathrm{L}^2_0(\mathbb{R})$, then it is positive definite w.r.t. $\mathrm{L}^1_0(\mathbb{R})$, and hence almost everywhere equal to a continuous function in $\mathrm{P}_{\mathrm{C},1}$ [15, Sec. 6]. This observation, in juxtaposition to Proposition 2.3.1, clearly demonstrates the connection between continuous, classically positive definite functions and the functions in $\mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$.

**Theorem 2.3.2** (Cooper, [14, Th. 7]). *Let $p \in [1,2]$, $q = p/2(p-1)$. If $f \in \mathrm{L}^q_{\mathrm{loc}}(\mathbb{R})$ and $f \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$, then $f \in \mathrm{P}(\mathrm{L}^p_0(\mathbb{R}))$.*

It follows from Theorem 2.3.2 that for any $1 < p \le 2$, there exist functions which are both singular at the origin and positive definite w.r.t. $\mathrm{L}^p_0(\mathbb{R})$. For example, consider the functions $f_\alpha = |\cdot|^{-\alpha}$, for $0 < \alpha < 1$. In Section 2.7 we show that $f_\alpha \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$ for all $0 < \alpha < 1$, and, for any $1 < p \le 2$, there exists $0 < \alpha < 1$ such that $f_\alpha \in \mathrm{L}^q_{\mathrm{loc}}(\mathbb{R})$ with $q = p/2(p-1)$.

Hence, when considering the classes $\mathrm{P}(\mathrm{L}_0^p(\mathbb{R}))$ for $1 \leq p \leq 2$, Theorem 2.3.2 infers that the transition from bounded to unbounded positive definite functions occurs at $p = 1$. That is, functions which are positive definite w.r.t. $\mathrm{L}_0^p(\mathbb{R})$, for $1 < p \leq 2$, may have a singularity at zero, whilst functions in $\mathrm{P}(\mathrm{L}_0^1(\mathbb{R}))$ are essentially bounded [14, Th. 5] and thus, non-singular at the origin.

## 2.4 Extension to functions in higher dimensions

Motivated by the work of Cooper [14], we introduce a notion of positive definiteness for functions defined on $\mathbb{R}^d$. Using this framework, we demonstrate that some of the results of Section 2.3 still hold true in higher dimensions. We begin by introducing an extended definition of positive definite functions on $\mathbb{R}^d$, which is analogous to Definition 2.3.1.

In the following, let $J$ be a set of complex-valued measurable functions on $\mathbb{R}^d$. This includes functions defined on a non-empty, measurable subset of $\mathbb{R}^d$, which we consider to be extended by zero to the whole of $\mathbb{R}^d$.

**Definition 2.4.1** A function $f : \mathbb{R}^d \to \mathbb{C}$ is called positive definite w.r.t. $J$ if for every $\phi \in J$, the integral

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{x} - \mathbf{y}) \phi(\mathbf{x}) \overline{\phi(\mathbf{y})} \, d\mathbf{x} d\mathbf{y} \tag{13}$$

exists (in the Lebesgue sense) and is non-negative.

Again, $\mathrm{P}(J)$ will denote the class of all functions which are positive definite w.r.t. the set $J$. It follows directly from Definition 2.4.1 that properties i.- iv. of Section 2.3 remain valid. Note that in the first property we now define $f^*(\mathbf{x}) := \overline{f(-\mathbf{x})}$ for all $\mathbf{x} \in \mathbb{R}^d$.

The following proposition demonstrates that, as in the classical case, even, real-valued positive definite functions are automatically positive definite in a complex sense. When referring to a property holding true almost everywhere (a.e.), or, alternatively, for almost all (f.a.a.) $\mathbf{x} \in \mathbb{R}^d$, we mean the property holds everywhere, except on sets of Lebesgue measure zero.

**Proposition 2.4.1** *Let $f : \mathbb{R}^d \to \mathbb{R}$ be such that $f(\mathbf{x}) = f(-\mathbf{x})$ f.a.a. $\mathbf{x} \in \mathbb{R}^d$. Let $\hat{J}$ denote a vector space of complex-valued functions on $\mathbb{R}^d$, such that if $\phi \in \hat{J}$, then $\overline{\phi} \in \hat{J}$ and $|\phi| \in \hat{J}$. Let $\hat{J}_{\mathbb{R}} := \left\{ \phi \in \hat{J} \mid \phi \text{ is real-valued} \right\}$. Then, $f \in \mathrm{P}(\hat{J}_{\mathbb{R}})$ if and only if $f \in \mathrm{P}(\hat{J})$.*

*Proof.* One direction is clear since $\hat{J}_{\mathbb{R}} \subseteq \hat{J}$. For the reverse implication, consider the following. Let $\psi \in \hat{J}$ and suppose $f \in \mathrm{P}(\hat{J}_{\mathbb{R}})$. First we prove the existence of the integral

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\psi(\mathbf{x})\overline{\psi(\mathbf{y})}\, d\mathbf{x} d\mathbf{y}. \tag{14}$$

Indeed,

$$\begin{aligned}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(\mathbf{x}-\mathbf{y})\psi(\mathbf{x})\overline{\psi(\mathbf{y})}|\, d\mathbf{x} d\mathbf{y} &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(\mathbf{x}-\mathbf{y})|\, |\psi(\mathbf{x})|\, |\psi(\mathbf{y})|\, d\mathbf{x} d\mathbf{y} \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(\mathbf{x}-\mathbf{y})\, \tilde{\psi}(\mathbf{x})\, \tilde{\psi}(\mathbf{y})|\, d\mathbf{x} d\mathbf{y} \quad (15)\end{aligned}$$

where $\tilde{\psi} = |\psi| \in \hat{J}_{\mathbb{R}}$. The integral in (15) exists since $f \in \mathrm{P}(\hat{J}_{\mathbb{R}})$. Hence, it follows that the integral in (14) exists in the Lebesgue sense. Next, we show the non-negativity of (14). $\psi$ can be re-written as

$$\psi = \mathrm{Re}\,(\psi) + i\, \mathrm{Im}\,(\psi)$$

where $\mathrm{Re}\,(\psi) : \mathbb{R}^d \to \mathbb{R}$ and $\mathrm{Im}\,(\psi) : \mathbb{R}^d \to \mathbb{R}$. Moreover,

$$\mathrm{Re}\,(\psi) = \frac{\psi + \overline{\psi}}{2} \in \hat{J} \quad \text{and} \quad \mathrm{Im}\,(\psi) = \frac{\psi - \overline{\psi}}{2i} \in \hat{J}.$$

Thus, $\mathrm{Re}\,(\psi), \mathrm{Im}\,(\psi) \in \hat{J}_{\mathbb{R}}$. Let $a := \mathrm{Re}\,(\psi)$, $b := \mathrm{Im}\,(\psi)$ and

$$t\,[u, v] := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\, u(\mathbf{x})\, \overline{v(\mathbf{y})}\, d\mathbf{x} d\mathbf{y} \qquad (u, v \in \hat{J}).$$

Then,

$$t\,[a, b] + t\,[b, a] = t\,[a + b, a + b] - t\,[a, a] - t\,[b, b]$$

and

$$-i\,(t\,[a, b] - t\,[b, a]) = t\,[\psi, \psi] - t\,[a, a] - t\,[b, b] \tag{16}$$

are finite, since $f \in \mathrm{P}(\hat{J}_{\mathbb{R}})$. Hence, $t\,[a, b]$ and $t\,[b, a]$ individually exist since both the sum $t\,[a, b] + t\,[b, a]$, and the difference $t\,[a, b] - t\,[b, a]$, exist. This allows us to use the Fubini theorem, which in conjunction with the evenness of $f$ gives

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\, a(\mathbf{x})\, b(\mathbf{y})\, d\mathbf{x} d\mathbf{y} = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\, b(\mathbf{x})\, a(\mathbf{y})\, d\mathbf{x} d\mathbf{y}.$$

Hence, $t\,[a,b] = t\,[b,a]$, and it follows from (16) that

$$t\,[\psi,\psi] = t\,[a,a] + t\,[b,b] \geq 0.$$

$\square$

The propositions presented in Section 2.3 extend naturally to the following results concerning functions defined on $\mathbb{R}^d$.

**Proposition 2.4.2** *If $f \in \mathrm{P}_{\mathrm{C},d}$ is continuous, then $f \in \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$.*

**Proposition 2.4.3** *If $f \in \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$ and $g \in \mathrm{P}_{\mathrm{C},d}$ is continuous, then $fg \in \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$.*

**Proposition 2.4.4** *For any $p \in [1,2]$, $\mathrm{P}(\mathrm{L}_0^p(\mathbb{R}^d)) \subseteq \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$.*

**Proposition 2.4.5** *For any $q \in (2,\infty]$ and $r \in [0,\infty]$, $\mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d) = \mathrm{P}(\mathrm{L}_0^q(\mathbb{R}^d)) = \mathrm{P}(\mathrm{C}_0^r(\mathbb{R}^d))$.*

The proofs of the above propositions are analogous to those of Propositions 2.3.1, 2.3.2, 2.3.3 and 2.3.4, respectively.

Next, we demonstrate that under certain conditions on our function, Definitions 2.1.1 and 2.4.1 coincide. In particular, a continuous function is classically positive definite if and only if it is positive definite w.r.t. $\mathrm{C}_0(\mathbb{R}^d)$.

**Proposition 2.4.6** *Let $f : \mathbb{R}^d \to \mathbb{C}$ be continuous. Then, $f \in \mathrm{P}_{\mathrm{C},d}$ if and only if $f \in \mathrm{P}(\mathrm{C}_0(\mathbb{R}^d))$.*

*Proof.* Let $f \in \mathrm{P}_{\mathrm{C},d}$. By Propositions 2.4.2 and 2.4.5, it follows directly that $f \in \mathrm{P}(\mathrm{C}_0(\mathbb{R}^d))$.

For the reverse implication, consider the following. Let $\psi : \mathbb{R} \to \mathbb{R}$ denote the bump function

$$\psi(x) = \begin{cases} c_0 \exp\left(\frac{1}{|x|^2-1}\right), & |x| < 1 \\ 0, & |x| \geq 1, \end{cases}$$

where $c_0 > 0$ is the constant chosen so that $\int_{\mathbb{R}} \psi(x)\,dx = 1$.

For any $\mathbf{x}=(x_1,x_2,\dots,x_d)\in\mathbb{R}^d$, define $\Psi(\mathbf{x}):=\psi(x_1)\psi(x_2)\dots\psi(x_d)$ and

$$\Psi_n(\mathbf{x}) := n^d\,\Psi(n\,\mathbf{x}) \qquad (n\in\mathbb{N}). \tag{17}$$

For any $n\in\mathbb{N}$, $\Psi_n\in \mathrm{C}_0^\infty(\mathbb{R}^d)$ is even and has compact support $[-\frac{1}{n},\,\frac{1}{n}]^d$. Moreover, $\int_{\mathbb{R}^d}\Psi_n(\mathbf{x})\,d\mathbf{x}=1$ for all $n\in\mathbb{N}$.

For any $N\in\mathbb{N}$ and any $\mathbf{x}_1,\mathbf{x}_2,\dots,\mathbf{x}_N\in\mathbb{R}^d$, define

$$\Phi_n(\mathbf{x}) := \sum_{i=1}^N \xi_i\,\Psi_n(\mathbf{x}-\mathbf{x}_i) \qquad (\mathbf{x}\in\mathbb{R}^d, n\in\mathbb{N}),$$

where $\xi_1,\xi_2,\dots,\xi_N\in\mathbb{C}$. Then, $\Phi_n\in\mathrm{C}_0^\infty(\mathbb{R}^d)$ and since $f$ is continuous,

$$\begin{aligned}\sum_{i,j=1}^N f(\mathbf{x}_i-\mathbf{x}_j)\xi_i\overline{\xi_j} &= \lim_{n\to\infty}\sum_{i,j=1}^N\int_{\mathbf{x}_j+[-\frac{1}{n},\frac{1}{n}]^d}\int_{\mathbf{x}_i+[-\frac{1}{n},\frac{1}{n}]^d} f(\mathbf{x}-\mathbf{y})\Phi_n(\mathbf{x})\overline{\Phi_n(\mathbf{y})}d\mathbf{x}d\mathbf{y}\\ &= \lim_{n\to\infty}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\Phi_n(\mathbf{x})\overline{\Phi_n(\mathbf{y})}\,d\mathbf{x}d\mathbf{y}\ge 0.\end{aligned}$$

□

We now show that functions which are positive definite w.r.t. $\mathrm{C}_0(\mathbb{R}^d)$ are locally integrable. We adapt and clarify the proof given by Cooper [14] in [14, Lemma 1].

**Proposition 2.4.7** *If $f\in \mathrm{P}(\mathrm{C}_0(\mathbb{R}^d))$, then $f\in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$.*

*Proof.* Let $\mathrm{K}\subset\mathbb{R}^d$ be any compact set and $\mathrm{I}=[a_1,b_1]\times[a_2,b_2]\times\dots\times[a_d,b_d]\subset\mathbb{R}^d$ be such that $\mathrm{K}\subset\mathrm{I}$. Let $c=\max\{|a_1|,|b_1|,|a_2|,\dots,|b_d|\}>0$.

Let $\psi\in\mathrm{C}_0(\mathbb{R})$ be such that $\psi$ is positive on $[-2c,\,2c]$. For any $\mathbf{x}=(x_1,x_2,\dots,x_d)\in\mathbb{R}^d$, define $\Psi(\mathbf{x}):=\psi(x_1)\psi(x_2)\dots\psi(x_d)$. Then $\Psi\in\mathrm{C}_0(\mathbb{R}^d)$ is positive on $[-2c,\,2c]^d$ and

$$g(\mathbf{z}) = \int_{[-c,\,c]^d}\Psi(\mathbf{z}+\mathbf{y})\Psi(\mathbf{y})\,d\mathbf{y} = \prod_{i=1}^d\int_{-c}^c \psi(z_i+y_i)\psi(y_i)\,dy_i \qquad (\mathbf{z}\in\mathbb{R}^d)$$

is positive and continuous on $[-c,\,c]^d$. Thus,

$$\begin{aligned}\inf_{\mathbf{x}\in[-c,\,c]^d} g(\mathbf{x})\int_{[-c,\,c]^d}|f(\mathbf{z})|\,d\mathbf{z} &\le \int_{[-c,\,c]^d}|f(\mathbf{z})g(\mathbf{z})|\,d\mathbf{z}\\ &\le \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|f(\mathbf{z})\Psi(\mathbf{z}+\mathbf{y})\Psi(\mathbf{y})|\,d\mathbf{y}d\mathbf{z}\\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|f(\mathbf{x}-\mathbf{y})\Psi(\mathbf{x})\Psi(\mathbf{y})|\,d\mathbf{x}d\mathbf{y} \qquad (18)\end{aligned}$$

where (18) follows from the Fubini theorem. Since $f \in \mathrm{P}(\mathrm{C}_0(\mathbb{R}^d))$, the integral in (18) exists. Hence,

$$\int_{\mathrm{K}} |f(\mathbf{z})|\, d\mathbf{z} \le \int_{\mathrm{I}} |f(\mathbf{z})|\, d\mathbf{z} \le \int_{[-c,\, c]^d} |f(\mathbf{z})|\, d\mathbf{z} < \infty.$$

$\square$

*Remark.* By Proposition 2.4.5, we can replace $\mathrm{C}_0(\mathbb{R}^d)$ in Propositions 2.4.6 and 2.4.7 with $\mathrm{L}_0^q(\mathbb{R}^d)$ for any $q \in [2, \infty]$, or $\mathrm{C}_0^r(\mathbb{R}^d)$ for any $r \in [1, \infty]$. Moreover, in Proposition 2.4.7, we can also replace $\mathrm{C}_0(\mathbb{R}^d)$ with a more general space $J$ of functions defined on $\mathbb{R}^d$, provided that for any $c > 0$, $J$ contains a function which is positive almost everywhere on $[-c,\, c]^d$.

## 2.5 An extension of Bochner's theorem to unbounded positive definite functions

We use Definition 2.4.1 with $J = \mathrm{L}^2(\mathbb{R}^d)$. Note that $\mathrm{L}^2(\mathbb{R}^d)$ is a smaller class of functions than those considered by Cooper [14], since $\mathrm{P}(\mathrm{L}^2(\mathbb{R}^d)) \subset \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$. However, for $\mathrm{L}^2(\mathbb{R}^d)$ as opposed to the space of compactly supported functions $\mathrm{L}_0^2(\mathbb{R}^d)$ of Theorem 2.3.1 (with $d = 1$), we obtain the following Bochner-type theorem.

**Theorem 2.5.1** *Let $f \in \mathrm{L}^1(\mathbb{R}^d)$. Then*

$$f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d)) \text{ if and only if } \hat{f} \ge 0,$$

*where $\hat{f}$ denotes the Fourier transform of $f$.*

We remark that under the hypothesis of Theorem 2.5.1, $f$ will correspond to a regular, in particular tempered, distribution, and hence Schwartz's version of Bochner's theorem applies, see Theorem 2.8.1. Nevertheless, with regard to applications where both $f$ and its Fourier transform are functions, the above generalised form of Bochner's theorem in Cooper's framework seems of interest, along with its more elementary proof and the further consequences shown in Sections 2.6 and 2.7 below.

The proof of Theorem 2.5.1 will be based upon the following two lemmas.

**Lemma 2.5.1** *Let $p \in [1, 2]$ and $q = p/2(p-1)$. Let $f \in \mathrm{L}^q(\mathbb{R}^d)$ and $\phi \in \mathrm{L}^p(\mathbb{R}^d)$. Then the integral in* (13) *exists, and*

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{x} - \mathbf{y}) \phi(\mathbf{x}) \overline{\phi(\mathbf{y})}\, d\mathbf{x} d\mathbf{y} = \int_{\mathbb{R}^d} f(\mathbf{z}) (\phi * \phi^*)(\mathbf{z})\, d\mathbf{z}, \tag{19}$$

*where $\phi^*(\mathbf{z}) = \overline{\phi(-\mathbf{z})}$ $(\mathbf{z} \in \mathbb{R}^d)$.*

*Proof.* Since the convolution of two elements of $\mathrm{L}^p(\mathbb{R}^d)$ is in $\mathrm{L}^r(\mathbb{R}^d)$ with $r = p/(2-p)$, and $1/q + 1/r = 1$, it follows that

$$\begin{aligned}\left|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\phi(\mathbf{x})\overline{\phi(\mathbf{y})}\,d\mathbf{x}d\mathbf{y}\right| &= \left|\int_{\mathbb{R}^d} f(\mathbf{z})\int_{\mathbb{R}^d}\phi(\mathbf{x})\overline{\phi(\mathbf{x}-\mathbf{z})}\,d\mathbf{x}d\mathbf{z}\right| \\ &= \left|\int_{\mathbb{R}^d} f(\mathbf{z})(\phi*\phi^*)(\mathbf{z})\,d\mathbf{z}\right| \le \|f\|_q\,\|\phi*\phi^*\|_r.\end{aligned}$$

□

*Remark.* Lemma 2.5.1 also holds true for $f \in \mathrm{L}^q_{\mathrm{loc}}(\mathbb{R}^d)$ and $\phi \in \mathrm{L}^p_0(\mathbb{R}^d)$. To see this, let $\phi \in \mathrm{L}^p_0(\mathbb{R}^d)$ and $K \subset \mathbb{R}^d$ denote the compact support of $\phi * \phi^*$. Then, in the final line of the above proof, we obtain

$$\left|\int_{\mathbb{R}^d} f(\mathbf{z})(\phi*\phi^*)(\mathbf{z})\,d\mathbf{z}\right| \le \left(\int_K |f(\mathbf{z})|^q\,d\mathbf{z}\right)^{\frac{1}{q}} \|\phi*\phi^*\|_r. \tag{20}$$

The following result is proved in a similar fashion to Proposition 2.3.4.

**Lemma 2.5.2** *Let $f \in \mathrm{L}^1(\mathbb{R}^d)$. Then $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ if and only if $f \in \mathrm{P}(\mathcal{S}(\mathbb{R}^d))$, where $\mathcal{S}(\mathbb{R}^d)$ denotes the Schwartz space of rapidly decreasing functions on $\mathbb{R}^d$.*

*Proof.* Since $\mathcal{S}(\mathbb{R}^d) \subset \mathrm{L}^2(\mathbb{R}^d)$, it follows directly that $\mathrm{P}(\mathrm{L}^2(\mathbb{R}^d)) \subset \mathrm{P}(\mathcal{S}(\mathbb{R}^d))$. For the reverse implication, we shall use the density of $\mathcal{S}(\mathbb{R}^d)$ in $\mathrm{L}^2(\mathbb{R}^d)$. Presume, $f \in \mathrm{P}(\mathcal{S}(\mathbb{R}^d))$. Then, the integral

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\psi(\mathbf{x})\overline{\psi(\mathbf{y})}\,d\mathbf{x}d\mathbf{y} = \int_{\mathbb{R}^d}(f^* * \psi)\,\overline{\psi} \tag{21}$$

exists in the Lebesgue sense and is non-negative for all $\psi \in \mathcal{S}(\mathbb{R}^d)$. Since $f \in \mathrm{L}^1(\mathbb{R}^d)$ and the convolution of an element of $\mathrm{L}^1(\mathbb{R}^d)$ with an element of $\mathrm{L}^2(\mathbb{R}^d)$ is in $\mathrm{L}^2(\mathbb{R}^d)$, the integral also exists for all $\psi \in \mathrm{L}^2(\mathbb{R}^d)$.

Let $\phi \in \mathrm{L}^2(\mathbb{R}^d)$. Then, similarly to as in the proof of Proposition 2.3.4, there is a sequence $(\psi_n)_{n\in\mathbb{N}}$ in $\mathcal{S}(\mathbb{R}^d)$ such that $\|\phi-\psi_n\|_2 \to 0$ as $n\to\infty$, and

$$\sup_{\mathbf{z}\in\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\Big(\phi(\mathbf{z}+\mathbf{y})\overline{\phi(\mathbf{y})} - \psi_n(\mathbf{z}+\mathbf{y})\overline{\psi_n(\mathbf{y})}\Big)d\mathbf{y}\right| \to 0 \quad \text{as } n\to\infty.$$

As $f \in \mathrm{L}^1(\mathbb{R}^d)$, it follows that

$$\left|\int_{\mathbb{R}^d} f(\mathbf{z})\int_{\mathbb{R}^d}\Big(\phi(\mathbf{z}+\mathbf{y})\overline{\phi(\mathbf{y})} - \psi_n(\mathbf{z}+\mathbf{y})\overline{\psi_n(\mathbf{y})}\Big)\,d\mathbf{y}d\mathbf{z}\right| \to 0 \quad \text{as } n\to\infty,$$

and hence

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\phi(\mathbf{x})\overline{\phi(\mathbf{y})}\,d\mathbf{x}d\mathbf{y} = \lim_{n\to\infty}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\psi_n(\mathbf{x})\overline{\psi_n(\mathbf{y})}\,d\mathbf{x}d\mathbf{y} \geq 0.$$

□

*Remark.* Similarly, it can be shown that if $f \in \mathrm{L}^1(\mathbb{R}^d)$, then $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ if and only if $f \in \mathrm{P}(\mathrm{C}_0(\mathbb{R}^d))$. Thus, by Proposition 2.4.5 it follows that for functions which are integrable over the whole of $\mathbb{R}^d$, $\mathrm{P}(\mathrm{L}^2(\mathbb{R}^d)) = \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}^d))$.

We now use Lemmas 2.5.1 and 2.5.2 to prove Theorem 2.5.1.

*Proof of Theorem 2.5.1.* Since $f \in \mathrm{L}^1(\mathbb{R}^d)$, the integral in (21) exists for all $\psi \in \mathcal{S}(\mathbb{R})$. Since the space of Schwartz functions is closed under convolution [31, Th. IX.3 (a)], $\psi * \psi^* \in \mathcal{S}(\mathbb{R}^d)$ for all $\psi \in \mathcal{S}(\mathbb{R}^d)$, where $\psi^*(\mathbf{z}) = \overline{\psi(-\mathbf{z})}$ ($\mathbf{z} \in \mathbb{R}^d$). Hence, for any $\mathbf{z} \in \mathbb{R}^d$ and $\psi \in \mathcal{S}(\mathbb{R}^d)$,

$$\begin{aligned}(\psi * \psi^*)(\mathbf{z}) &= (2\pi)^{-\frac{d}{2}}\int_{\mathbb{R}^d} (\psi * \psi^*)\check{}(\mathbf{x})\,e^{-i\mathbf{x}\cdot\mathbf{z}}\,d\mathbf{x} = \int_{\mathbb{R}^d} \check{\psi}(\mathbf{x})\,\check{\psi}^*(\mathbf{x})\,e^{-i\mathbf{x}\cdot\mathbf{z}}\,d\mathbf{x}\\ &= \int_{\mathbb{R}^d} |\check{\psi}(\mathbf{x})|^2\,e^{-i\mathbf{x}\cdot\mathbf{z}}\,d\mathbf{x},\end{aligned}$$

since

$$\int_{\mathbb{R}^d} \overline{\psi(-\mathbf{u})}\,e^{i\mathbf{u}\cdot\mathbf{x}}\,d\mathbf{u} = \overline{\int_{\mathbb{R}^d} \psi(\mathbf{u})\,e^{i\mathbf{u}\cdot\mathbf{x}}\,d\mathbf{u}} \qquad (\mathbf{x} \in \mathbb{R}^d). \tag{22}$$

Hence, $\check{\psi}^* = \overline{\check{\psi}}$. By Lemma 2.5.1,

$$\begin{aligned}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\psi(\mathbf{x})\overline{\psi(\mathbf{y})}\,d\mathbf{x}d\mathbf{y} &= \int_{\mathbb{R}^d} f(\mathbf{z})(\psi * \psi^*)(\mathbf{z})\,d\mathbf{z}\\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{z})\,|\check{\psi}(\mathbf{x})|^2\,e^{-i\mathbf{x}\cdot\mathbf{z}}d\mathbf{x}d\mathbf{z} = (2\pi)^{\frac{d}{2}}\int_{\mathbb{R}^d} \hat{f}(\mathbf{x})\,|\check{\psi}(\mathbf{x})|^2 d\mathbf{x}. \quad (23)\end{aligned}$$

From (23) it is clear that if $\hat{f} \geq 0$, then $f \in \mathrm{P}(\mathcal{S}(\mathbb{R}^d))$. By Lemma 2.5.2 it follows that $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$.

Conversely, suppose that $\hat{f}(\mathbf{z}) < 0$ at some point $\mathbf{z} \in \mathbb{R}^d$ (reductio ad absurdum). $\hat{f}$ is continuous and bounded because $f \in \mathrm{L}^1(\mathbb{R}^d)$. It follows that there exists $\delta > 0$ such that $\hat{f}(\mathbf{x}) < 0$ for all $\|\mathbf{x}-\mathbf{z}\| < \delta$, where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^d$. Let

$$\psi_1(\mathbf{x}) = \begin{cases} \exp\left[\left(\|\mathbf{x}-\mathbf{z}\|^2 - \delta^2\right)^{-1}\right] & \text{if } \|\mathbf{x}-\mathbf{z}\| < \delta \\ 0 & \text{otherwise}\end{cases} \qquad (\mathbf{x} \in \mathbb{R}^d).$$

Then $\psi_1 \in \mathrm{C}_0^\infty(\mathbb{R}^\mathrm{d}) \subset \mathcal{S}(\mathbb{R}^\mathrm{d})$. For $\psi_2 := \hat{\psi}_1 \in \mathcal{S}(\mathbb{R}^d)$, it follows by (23) that

$$\begin{aligned}0 &\leq \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\psi_2(\mathbf{x})\overline{\psi_2(\mathbf{y})}\,d\mathbf{x}d\mathbf{y} = (2\pi)^{\frac{d}{2}}\int_{\mathbb{R}^d} \hat{f}(\mathbf{x})\,|\check{\psi}_2(\mathbf{x})|^2\,d\mathbf{x}\\ &= (2\pi)^{\frac{d}{2}}\int_{\|\mathbf{x}-\mathbf{z}\|<\delta} \hat{f}(\mathbf{x})\,|\psi_1(\mathbf{x})|^2\,d\mathbf{x} < 0,\end{aligned}$$

which is a contradiction. □

*Remark.* It follows from Theorem 2.5.1 that if $f \in \mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$, then $f = f^*$ almost everywhere. Indeed, $\hat{f^*} = \hat{f} \geq 0$ by (22), and thus $f = f^*$ almost everywhere by the uniqueness of the Fourier transform on $\mathrm{L}^1(\mathbb{R}^d)$ [11, Th. 34].

An alternative notion of *$L^2$-positive definiteness* was introduced in [12], where Definition 2.4.1 is used with $J = \mathrm{L}^2(\mathbb{R})$ and $f$ replaced with a two-variable kernel $k \in \mathrm{L}^2(\mathbb{R}^2)$. Buescu et al. were primarily interested in the spectral properties of the integral operator

$$\int_{\mathbb{R}} k(x,y)\phi(x)\,dx \quad (\phi \in \mathrm{L}^2(\mathbb{R})),$$

where $k$ is such that $\int_{\mathbb{R}}\int_{\mathbb{R}} k(x,y)\phi(x)\overline{\phi(y)}\,dxdy \geq 0$ for all $\phi \in \mathrm{L}^2(\mathbb{R})$. In our case, we define the operator for a given function $f \in \mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$, associated with the sesquilinear form

$$q[\phi,\psi] := \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\phi(\mathbf{x})\overline{\psi(\mathbf{y})}\,d\mathbf{x}d\mathbf{y} \quad (\phi,\psi \in \mathrm{L}^2(\mathbb{R}^d)),$$

by

$$T(\phi) := \int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\phi(\mathbf{x})\,d\mathbf{x} = f^* * \phi \quad (\phi \in \mathrm{L}^2(\mathbb{R}^d)).$$

By Young's inequality, $T : \mathrm{L}^2(\mathbb{R}^d) \to \mathrm{L}^2(\mathbb{R}^d)$. The following facts proceed immediately from the standard theory on forms and operators, see e.g. [19, 46]. Since $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$, the quadratic form $q[\phi] := q[\phi,\phi]$ is non-negative and thus, $q$ is symmetric; that is, $q[\phi,\psi] = \overline{q[\psi,\phi]}$ for all $\phi,\psi \in \mathrm{L}^2(\mathbb{R}^d)$. $q$ is bounded below by zero and its upper bound is $\|f\|_1$, since for any $\phi \in \mathrm{L}^2(\mathbb{R}^d)$,

$$0 \leq q[\phi,\phi] \leq \|f^* * \phi\|_2\|\phi\|_2 \leq \|f^*\|_1\|\phi\|_2^2 = \|f\|_1\|\phi\|_2^2,$$

by the Cauchy-Schwarz inequality and Young's inequality, respectively. Our operator $T$ is non-negative, bounded and self-adjoint. Moreover, for any $\phi \in \mathcal{S}(\mathbb{R}^d)$, $f^* * \phi \in \mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{L}^2(\mathbb{R}^d)$ by Young's inequality, and thus

$$T(\phi) = f^* * \phi = \mathcal{F}^{-1}\mathcal{F}\big(f^* * \phi\big) \quad (\phi \in \mathcal{S}(\mathbb{R}^d)),$$

where $\mathcal{F}$ denotes the Fourier transform on $\mathrm{L}^2(\mathbb{R}^d)$. Since $\mathcal{F}(\cdot)$ and $\hat{(\cdot)}$, the Fourier transform on $\mathrm{L}^1(\mathbb{R}^d)$, coincide on $\mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{L}^2(\mathbb{R}^d)$, it follows that

$$T(\phi) = \mathcal{F}^{-1}\big((2\pi)^{\frac{d}{2}} \hat{f}^* \, \mathcal{F}(\phi)\big) \quad (\phi \in \mathcal{S}(\mathbb{R}^d)).$$

By equation (22), $\hat{f}^* = \hat{f} \geq 0$, and thus $T = (2\pi)^{\frac{d}{2}} \mathcal{F}^{-1} \mathcal{M}_{\hat{f}} \mathcal{F}$, where $\mathcal{M}_{\hat{f}}$ denotes the multiplication operator $\mathcal{M}_{\hat{f}}(\cdot) := \hat{f} \times (\cdot)$.

$\mathcal{M}_{\hat{f}}$ is a bounded linear operator on $\mathrm{L}^2(\mathbb{R}^d)$, for $\|\mathcal{M}_{\hat{f}}(\phi)\|_2 \leq \|\hat{f}\|_\infty \|\phi\|_2$ for any $\phi \in \mathrm{L}^2(\mathbb{R}^d)$, and thus, so is $T$. Since $\mathcal{S}(\mathbb{R}^d)$ is dense in $\mathrm{L}^2(\mathbb{R}^d)$, for any $\phi \in \mathrm{L}^2(\mathbb{R}^d)$ there exists a sequence of functions $(\phi_n)_{n \in \mathbb{N}}$ such that $\phi_n \in \mathcal{S}(\mathbb{R}^d)$ and $\|\phi - \phi_n\|_2 \to 0$ as $n \to \infty$. Thus, for any $\psi \in \mathrm{L}^2(\mathbb{R}^d)$, there exists a sequence of functions $(\psi_n)_{n \in \mathbb{N}}$ such that $\psi_n \in \mathcal{S}(\mathbb{R}^d)$,

$$\|T(\psi) - T(\psi_n)\|_2 = \|T(\psi - \psi_n)\|_2 \leq \|T\|_{\mathrm{op}} \|\psi - \psi_n\|_2 \to 0 \quad \text{as } n \to \infty,$$

where $\|T\|_{\mathrm{op}}$ denotes the operator norm of $T$, and

$$\|f^* * \psi - f^* * \psi_n\|_2 = \|f^* * (\psi - \psi_n)\|_2 \leq \|f^*\|_1 \|\psi - \psi_n\|_2 \to 0 \quad \text{as } n \to \infty.$$

Hence,

$$T(\phi) = (2\pi)^{\frac{d}{2}} \mathcal{F}^{-1} \mathcal{M}_{\hat{f}} \mathcal{F}(\phi) = f^* * \phi \quad (\phi \in \mathrm{L}^2(\mathbb{R}^d)).$$

Since $\mathcal{F}$ is unitary on $\mathrm{L}^2(\mathbb{R}^d)$, it follows that the spectrum of $T$, $\sigma(T)$, coincides with the spectrum of $\mathcal{M}_{\hat{f}}$, $\sigma(\mathcal{M}_{\hat{f}})$. We know that $\sigma(\mathcal{M}_{\hat{f}}) = \overline{\mathrm{Range}(\hat{f})} = [0, \max(\hat{f})]$, where $\overline{\mathrm{Range}(\hat{f})}$ denotes the closure of $\mathrm{Range}(\hat{f})$, since the operator $\mathcal{M}_{\hat{f}} - \lambda$ is invertible if and only if $\lambda$ is not in this range. Moreover, it is clear that $\lambda \in \sigma(\mathcal{M}_{\hat{f}})$ if and only if there exists a set $K \subset \mathbb{R}^d$, $|K| > 0$, such that $\hat{f}(\mathbf{x}) = \lambda$ for all $\mathbf{x} \in K$, see [46, p. 103] for details.

## 2.6 Approximation by positive definite functions and convolution squares

In this section we present some corollaries to Theorem 2.5.1. In particular, we show that functions in $\mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ can be approximated by continuous, classically positive definite functions. We also establish connections between functions which are positive definite for $\mathrm{L}^2(\mathbb{R}^d)$ and functions which arise as convolution squares. We begin by proving the following technical lemma, a consequence of which is that $\mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ is a closed subset of $\mathrm{L}^1(\mathbb{R}^d)$.

**Lemma 2.6.1** *Let $p \in [1,2]$ and $q = p/2(p-1)$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions such that $f_n \in \mathrm{L}^q(\mathbb{R}^d)$ and $f_n \in \mathrm{P}(\mathrm{L}^p(\mathbb{R}^d))$ $(n \in \mathbb{N})$. If $\lim_{n\to\infty} \|f_n - f\|_q = 0$ for some $f \in \mathrm{L}^q(\mathbb{R}^d)$, then $f \in \mathrm{P}(\mathrm{L}^p(\mathbb{R}^d))$.*

*Proof.* Let $\phi \in \mathrm{L}^p(\mathbb{R})$ and $r = p/(2-p)$. By Lemma 2.5.1,

$$\left|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(f_n(\mathbf{x}-\mathbf{y}) - f(\mathbf{x}-\mathbf{y}))\,\phi(\mathbf{x})\overline{\phi(\mathbf{y})}\,d\mathbf{x}d\mathbf{y}\right| \leq \|f_n - f\|_q\|\phi * \phi^*\|_r \to 0$$

$(n \to \infty)$. Thus,

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\phi(\mathbf{x})\overline{\phi(\mathbf{y})}\,d\mathbf{x}d\mathbf{y} = \lim_{n\to\infty}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f_n(\mathbf{x}-\mathbf{y})\phi(\mathbf{x})\overline{\phi(\mathbf{y})}\,d\mathbf{x}d\mathbf{y} \geq 0.$$

□

Similarly, $\mathrm{P}(\mathrm{L}^2_0(\mathbb{R}^d))$ is a closed subset of $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ (recall, functions in $\mathrm{P}(\mathrm{L}^2_0(\mathbb{R}^d))$ are automatically locally integrable by Proposition 2.4.7). In fact, we have the following lemma, which will be used during the proof of Lemma 3.4.4 (see Section 3.4).

**Lemma 2.6.2** *Let $p \in [1,2]$ and $q = p/2(p-1)$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions such that $f_n \in \mathrm{L}^q_{\mathrm{loc}}(\mathbb{R}^d)$ and $f_n \in \mathrm{P}(\mathrm{L}^p_0(\mathbb{R}^d))$ $(n \in \mathbb{N})$. If $\lim_{n\to\infty}\int_K |f_n(\mathbf{x}) - f(\mathbf{x})|^q\,d\mathbf{x} = 0$ for some $f \in \mathrm{L}^q_{\mathrm{loc}}(\mathbb{R}^d)$ and any compact set $K \subset \mathbb{R}^d$, then $f \in \mathrm{P}(\mathrm{L}^p_0(\mathbb{R}^d))$.*

The proof of Lemma 2.6.2 follows the same steps as the proof of Lemma 2.6.1, using Lemma 2.5.1 and its ensuing remark. Lemmas 2.6.1 and 2.6.2 are analogous to the pointwise convergence property for the classical positive definite functions, see [43, p. 412].

We now present some consequences of Theorem 2.5.1. The first observation is that $\mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ is the closure of $\mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{P}_{\mathrm{C},d}$.

**Corollary 2.6.1** *Let $f \in \mathrm{L}^1(\mathbb{R}^d)$. Then, $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ if and only if there is a sequence $(g_n)_{n\in\mathbb{N}}$ of continuous functions such that $g_n \in \mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{P}_{\mathrm{C},d}$ $(n \in \mathbb{N})$ and $\lim_{n\to\infty}\|g_n - f\|_1 = 0$.*

*Proof.* Suppose $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$. As $f \in \mathrm{L}^1(\mathbb{R}^d)$, its Fourier transform $\hat{f}$ is continuous and bounded, and $\hat{f}(\mathbf{x}) \to 0$ as $\|\mathbf{x}\| \to \infty$ (again, $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^d$). Also, by Theorem 2.5.1, $\hat{f} \geq 0$. Let

$$\eta(\mathbf{u}) = (2\pi)^{-\frac{d}{2}}\, e^{-\|\mathbf{u}\|^2/2} \qquad (\mathbf{u} \in \mathbb{R}^d),$$

and, for $n \in \mathbb{N}$, $\eta_n := n^d \, \eta\,(n\,\cdot)$. Then,

$$\int_{\mathbb{R}^d} |\eta_n(\mathbf{x})| \, d\mathbf{x} = \int_{\mathbb{R}^d} |\eta(\mathbf{x})| \, d\mathbf{x} = \prod_{i=1}^{d} \left( (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} e^{-x_i^2/2} \, dx_i \right) = 1.$$

Define

$$h_n(\mathbf{u}) = \hat{f}(\mathbf{u}) \, e^{-\|\mathbf{u}\|^2/(2n^2)} \geq 0 \qquad (\mathbf{u} \in \mathbb{R}^d).$$

Then, $h_n \in \mathrm{L}^1(\mathbb{R}^d)$. Let $g_n = \check{h}_n$ be the inverse Fourier transform of $h_n$. By Bochner's theorem [9, Chapter IV.20], $g_n$ is continuous and classically positive definite. In particular, it has the property that $|g_n(\mathbf{v})| \leq g_n(0) < \infty$ $(\mathbf{v} \in \mathbb{R}^d)$. By the Fubini theorem, for any $\mathbf{v} \in \mathbb{R}^d$ and $n \in \mathbb{N}$,

$$\begin{aligned}
g_n(\mathbf{v}) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \left( (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} f(\mathbf{z}) \, e^{-i\mathbf{x}\cdot\mathbf{z}} \, d\mathbf{z} \right) e^{-\|\mathbf{x}\|^2/(2n^2)} \, e^{i\mathbf{v}\cdot\mathbf{x}} \, d\mathbf{x} \\
&= \int_{\mathbb{R}^d} f(\mathbf{z}) \left( (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} (2\pi)^{-\frac{d}{2}} \, e^{-\|\mathbf{x}\|^2/(2n^2)} \, e^{-i\mathbf{x}\cdot(\mathbf{z}-\mathbf{v})} \, d\mathbf{x} \right) d\mathbf{z} \\
&= \int_{\mathbb{R}^d} f(\mathbf{z}) \left( \prod_{i=1}^{d} \left( (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} \left( (2\pi)^{-\frac{1}{2}} \, e^{-x_i^2/2n^2} \right) e^{-ix_i(z_i - v_i)} \, dx_i \right) \right) d\mathbf{z} \\
&= \int_{\mathbb{R}^d} f(\mathbf{z}) \left( \prod_{i=1}^{d} n \, (2\pi)^{-\frac{1}{2}} e^{-(n(z_i - v_i))^2/2} \right) d\mathbf{z} \\
&= \int_{\mathbb{R}^d} f(\mathbf{z}) \left( n^d (2\pi)^{-\frac{d}{2}} \, e^{-\|n(\mathbf{v}-\mathbf{z})\|^2/2} \right) d\mathbf{z} = f * \eta_n \, (\mathbf{v}).
\end{aligned}$$

By Young's inequality, $g_n \in \mathrm{L}^1(\mathbb{R}^d)$ $(n \in \mathbb{N})$. Since $f \in \mathrm{L}^1(\mathbb{R}^d)$ and $\eta \in \mathrm{L}^1(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \eta(\mathbf{x}) \, d\mathbf{x} = 1$, it follows that $\lim_{n\to\infty} \|g_n - f\|_1 = 0$ [42, Th. 1.18].

For the reverse direction, we need only show that $g_n \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ $(n \in \mathbb{N})$. Since $g_n \in \mathrm{L}^1(\mathbb{R}^d)$ $(n \in \mathbb{N})$, $g_n$ has a continuous Fourier transform, and it follows from Bochner's theorem that $\hat{g}_n \geq 0$ $(n \in \mathbb{N})$. Thus, $g_n \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ $(n \in \mathbb{N})$ by Theorem 2.5.1. $\square$

We show next that $\mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ is closed under convolution and, under the further assumption of square integrability, under pointwise multiplication as well.

**Corollary 2.6.2** *Let $f, g \in \mathrm{L}^1(\mathbb{R}^d)$. If $f, g \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ then $f * g \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$.*

*Proof.* Suppose $f, g \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$. By Theorem 2.5.1, $\hat{f}$, $\hat{g} \geq 0$. By Young's inequality, $f * g \in \mathrm{L}^1(\mathbb{R}^d)$; moreover

$$\widehat{f * g} = (2\pi)^{\frac{d}{2}} \, \hat{f} \, \hat{g} \geq 0,$$

so $f * g \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ by Theorem 2.5.1. $\square$

**Corollary 2.6.3** *Let $f, g \in \mathrm{L}^1(\mathbb{R}^d) \cap \mathrm{L}^2(\mathbb{R}^d)$. If $f, g \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ then $fg \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$.*

*Proof.* Suppose $f, g \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$. By Theorem 2.5.1, $\hat{f}$, $\hat{g} \geq 0$. By the Cauchy-Schwarz inequality, $fg \in \mathrm{L}^1(\mathbb{R}^d)$; furthermore

$$\widehat{fg} = (2\pi)^{-\frac{d}{2}}\, \hat{f} * \hat{g} \geq 0,$$

hence $fg \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ by Theorem 2.5.1. $\square$

The next statement shows that functions which arise as 'convolution squares' are positive definite in the new sense, note that $p^*(\mathbf{z}) = \overline{p(-\mathbf{z})}$ ($\mathbf{z} \in \mathbb{R}^d$) as before.

**Corollary 2.6.4** *If $f = p * p^*$ for some $p \in \mathrm{L}^1(\mathbb{R}^d)$, then $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$.*

*Proof.* Suppose $f = p * p^*$ with $p \in \mathrm{L}^1(\mathbb{R}^d)$. By Young's inequality, $f \in \mathrm{L}^1(\mathbb{R}^d)$. From (22) it follows that

$$\hat{f} = \widehat{p * p^*} = (2\pi)^{\frac{d}{2}}\, \hat{p}\, \hat{p^*} = (2\pi)^{\frac{d}{2}}\, |\hat{p}|^2 \geq 0.$$

Thus, $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ by Theorem 2.5.1. $\square$

This result is analogous to the classical result that if $f = g * g^*$ for some $g \in \mathrm{L}^2(\mathbb{R})$, then $f$ is continuous and positive definite in the original sense [22, Th. 4.2.4]. Note that in the classical case we have $f \in \mathrm{L}^\infty(\mathbb{R})$, since the convolution of two elements of $\mathrm{L}^2(\mathbb{R})$ is in $\mathrm{L}^\infty(\mathbb{R})$, whereas in our present situation we have $f = p * p^* \in \mathrm{L}^1(\mathbb{R}^d)$, again by Young's inequality.

In Corollary 2.6.5 we show that a version of the converse to Corollary 2.6.4 is also true, *viz.* that a function which is positive definite w.r.t. $\mathrm{L}^2(\mathbb{R}^d)$ can be written, in some sense, as a convolution square. An analogous statement is known for continuous, classically positive definite functions (Khinchine's criterion, [22, Th. 4.2.5]). In particular, *if $f : \mathbb{R} \to \mathbb{C}$ is a characteristic function then there exists a sequence $(g_n)_{n \in \mathbb{N}}$ of complex-valued functions, such that for any $n \in \mathbb{N}$, $\int_{\mathbb{R}} |g_n(x)|^2\, dx = 1$, and $f(t) = \lim_{n\to\infty} g_n * g_n^*(t)$ holds uniformly in every finite t-interval.* Note that a function $f : \mathbb{R} \to \mathbb{C}$ is a characteristic function if and only if $f \in \mathrm{P}_{\mathrm{C},1}$ is continuous and $f(0) = 1$. The final condition can always be achieved via normalisation due to the bounded nature of classical positive definite functions.

**Corollary 2.6.5** *Let $f \in \mathrm{L}^1(\mathbb{R}^d)$. If $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$, then there is a sequence $(p_n)_{n\in\mathbb{N}}$ of functions such that $p_n \in \mathrm{L}^2(\mathbb{R}^d)$, $p_n * p_n^* \in \mathrm{L}^1(\mathbb{R}^d)$ $(n \in \mathbb{N})$, and $\lim_{n\to\infty} \|p_n * p_n^* - f\|_1 = 0$.*

*Proof.* Let $g_n = \check{h}_n$ $(n \in \mathbb{N})$ be the functions constructed in the proof of Corollary 2.6.1: then, $\lim_{n\to\infty} \|g_n - f\|_1 = 0$. Let $n \in \mathbb{N}$. Since $h_n \geq 0$, then $h_n = q_n^2$ for $q_n := \sqrt{h_n} \in \mathrm{L}^2(\mathbb{R}^d)$. The Fourier transformation $\tilde{q}_n = \mathcal{F}(q_n)$ given by

$$\tilde{q}_n(\mathbf{u}) = \underset{R\to\infty}{\mathrm{l.i.m.}}\ (2\pi)^{-\frac{d}{2}} \int_{[-R,\,R]^d} q_n(\mathbf{x})\, e^{-i\mathbf{x}\cdot\mathbf{u}}\, d\mathbf{x} \quad (\mathbf{u} \in \mathbb{R}^d)$$

defines a unitary operator $\mathcal{F}$ on $\mathrm{L}^2(\mathbb{R}^d)$ to $\mathrm{L}^2(\mathbb{R}^d)$. By $\mathrm{l.i.m.}_{R\to\infty}$ we mean the limit in the mean as $R$ tends to infinity. In other words, $\|\tilde{q}_n - \hat{q}_{n,R}\|_2 \to 0$ as $R \to \infty$, where for any $R > 0$, $q_{n,R}$ denotes the integrable function

$$q_{n,R}(\mathbf{x}) = \begin{cases} q_n(\mathbf{x}), & \mathbf{x} \in [-R,\, R]^d \\ 0, & \text{otherwise,} \end{cases}$$

and $\hat{q}_{n,R}$ represents its Fourier transform

$$\hat{q}_{n,\,R}(\mathbf{u}) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} q_{n,R}(\mathbf{x})\, e^{-i\mathbf{x}\cdot\mathbf{u}}\, d\mathbf{x} \quad (\mathbf{u} \in \mathbb{R}^d).$$

The inverse operator $\mathcal{F}^{-1}(\tilde{q}_n)$ is given by

$$q_n(\mathbf{x}) = \underset{R\to\infty}{\mathrm{l.i.m.}}\ (2\pi)^{-\frac{d}{2}} \int_{[-R,\,R]^d} \tilde{q}_n(\mathbf{u})\, e^{i\mathbf{u}\cdot\mathbf{x}}\, d\mathbf{u} \quad (\mathbf{x} \in \mathbb{R}^d). \tag{24}$$

Let $\mathbf{y} \in \mathbb{R}^d$ and define $\tilde{v}_n(\mathbf{x}) := \tilde{q}_n(\mathbf{x}+\mathbf{y})$ for all $\mathbf{x} \in \mathbb{R}^d$. Then, $\tilde{v}_n \in \mathrm{L}^2(\mathbb{R}^d)$, and it follows from (24) that

$$v_n(\mathbf{x}) = \underset{R\to\infty}{\mathrm{l.i.m.}}\ (2\pi)^{-\frac{d}{2}} \int_{[-R,\,R]^d} \tilde{q}_n(\mathbf{u}+\mathbf{y})\, e^{i\mathbf{u}\cdot\mathbf{x}}\, d\mathbf{u} = q_n(\mathbf{x})\, e^{-i\mathbf{y}\cdot\mathbf{x}} \quad (\mathbf{x} \in \mathbb{R}^d). \tag{25}$$

By the Parseval identity [1, Ex. 4.2.9] and (25), we have

$$\int_{\mathbb{R}^d} \tilde{q}_n(\mathbf{x})\overline{\tilde{v}_n(\mathbf{x})}\, d\mathbf{x} = \int_{\mathbb{R}^d} q_n(\mathbf{x})\overline{v_n(\mathbf{x})}\, d\mathbf{x} = \int_{\mathbb{R}^d} q_n^2(\mathbf{x})\, e^{i\mathbf{y}\cdot\mathbf{x}}\, d\mathbf{x}.$$

It follows from a change of variables that

$$\int_{\mathbb{R}^d} \tilde{q}_n(\mathbf{x})\overline{\tilde{q}_n(\mathbf{x}+\mathbf{y})}\, d\mathbf{x} = \int_{\mathbb{R}^d} \tilde{q}_n(-\mathbf{x})\tilde{q}_n^*(\mathbf{x}-\mathbf{y})\, d\mathbf{x},$$

and thus,

$$g_n(\mathbf{y}) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} h_n(\mathbf{x})\, e^{i\mathbf{x}\cdot\mathbf{y}}\, d\mathbf{x} = p_n * p_n^*\,(\mathbf{y})$$

where $p_n = (2\pi)^{-\frac{d}{4}}\, \tilde{q}_n(-\cdot) \in \mathrm{L}^2(\mathbb{R}^d)$. $\square$

*Remark.* Firstly, the converse to Corollary 2.6.5 is true, since for any $n \in \mathbb{N}$, $p_n * p_n^*$ is continuous and classically positive definite by [22, Th. 4.2.4], and thus, $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ by Corollary 2.6.1. Secondly, $p_n \notin \mathrm{L}^1(\mathbb{R}^d)$ in general. The proof of Corollary 2.6.5 is a natural extension of the proof of Theorem 4.2.4 (i) [22], to functions defined on $\mathbb{R}^d$. Note, however, that unlike in Theorem 4.2.4 (ii) [22], we don't have $\int_{\mathbb{R}^d} |p_n(\mathbf{x})|^2\, d\mathbf{x} = 1$, since we do not assume that $h_n$ is the density of a probability measure.

## 2.7 Sufficient criteria for generalised positive definiteness

It proceeds from Theorem 2.5.1 that an integrable function is positive definite for $\mathrm{L}^2(\mathbb{R}^d)$ if its Fourier transform is non-negative. In this section we provide sufficient conditions for this criterion.

For a measurable set $K \subset \mathbb{R}^d$ and $p \in [1, \infty)$, let

$$\mathrm{L}^p(K) = \left\{ f \,:\, K \to \mathbb{C} \;\middle|\; \int_K |f(\mathbf{x})|^p\, d\mathbf{x} < \infty \right\}.$$

Naturally $\mathrm{L}^p(K) \subset \mathrm{L}^p(\mathbb{R}^d)$, extending functions by zero on $\mathbb{R}^d \backslash K$. We always use this embedding by extension in the following.

A direct consequence of Theorem 2.2.2 is that if $\psi \in \mathrm{CM}$ is bounded, then $f = \psi(\|\cdot\|^2) : \mathbb{R}^d \to [0, \infty)$ is continuous and classically positive definite for any $d \in \mathbb{N}$. We now generalise this observation to potentially unbounded completely monotone functions.

**Corollary 2.7.1** *Let $f \in \mathrm{CM}$ and $g = f(\|\cdot\|^2) : \mathbb{R}^d \to [0, \infty)$. If $g \in \mathrm{L}^1(\mathbb{R}^d)$, then $g \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$.*

*Proof.* By Theorem 2.2.1, $f$ is the Laplace transform of a non-negative measure $\mu$ on $[0,\infty)$. By the Fubini theorem, for any $\mathbf{u} \in \mathbb{R}^d$,

$$\begin{aligned}
\hat{g}(\mathbf{u}) &= (2\pi)^{-\frac{d}{2}} \int_{[0,\infty)} \int_{\mathbb{R}^d} e^{-\|\mathbf{x}\|^2 t} e^{-i\mathbf{x}\cdot\mathbf{u}}\, d\mathbf{x}\, \mu(dt) \\
&= \int_{[0,\infty)} \prod_{i=1}^{d} \left( (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} e^{-x_i^2 t} e^{-i x_i u_i}\, dx_i \right) \mu(dt) \\
&= \int_{[0,\infty)} \prod_{i=1}^{d} \left( (2t)^{-\frac{1}{2}} e^{-u_i^2/4t} \right) \mu(dt) = \int_{[0,\infty)} (2t)^{-\frac{d}{2}} e^{-\|\mathbf{u}\|^2/4t} \mu(dt) \geq 0.
\end{aligned}$$

Thus, $g \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ by Theorem 2.5.1. □

The next result is an analogue of Pólya's criterion [22, Th. 4.3.1] for continuous positive definite functions. Our extension also applies to unbounded functions with an integrable singularity at 0. The proof is based on a technique used by Tuck [45], which shows the non-negativity of a certain Fourier transform, and considers the case $d = 1$, i.e., functions defined on the real line.

**Corollary 2.7.2** *Let $f \in \mathrm{L}^1(\mathbb{R})$ be a function with the following three properties.*

*i. $f$ is locally absolutely continuous on $(0,\infty)$, and $f' \in \mathrm{L}^1_{\mathrm{loc}}((0,\infty))$ has a non-positive, non-decreasing representative.*

*ii. $f(x) = f(-x)$ $(x \in \mathbb{R})$.*

*iii. $f \geq 0$.*

*Then, $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}))$.*

*Proof.* By Theorem 2.5.1, we need only show that the Fourier transform $\hat{f} \geq 0$. Since $f$ is even and real-valued, its Fourier transform $\hat{f}$ is given by

$$\hat{f}(\xi) = \sqrt{\frac{2}{\pi}} \int_0^\infty f(x) \cos(x\,\xi)\, dx \qquad (\xi \in \mathbb{R}),$$

a real-valued, even, bounded function. It is immediate from property iii that $\hat{f}(0) \geq 0$. Hence, it suffices to consider $\xi > 0$ in the following. By property i, $f$ is non-decreasing on $(0,\infty)$. Using this property combined with the facts that $f$ is non-negative and integrable, it follows that

$$\lim_{x\to\infty} f(x) = 0. \tag{26}$$

By the Mean Value Theorem, for any $x > 0$ there is $0 < \xi_x < x$ such that

$$\int_0^x f(y)\,dy = x f(\xi_x).$$

Hence, since $f$ is non-increasing on $(0, \infty)$, it follows that

$$0 \le x f(x) \le \int_0^x f(y)\,dy \qquad (x > 0)$$

and consequently,

$$\lim_{x \to 0} x f(x) = 0. \tag{27}$$

Since $f$ is locally absolutely continuous on $(0, \infty)$, we can use integration by parts to obtain

$$\int_{x_1}^{x_2} f(x) \cos(x\,\xi)\,dx = \frac{1}{\xi}\left[f(x)\sin(x\,\xi)\right]_{x_1}^{x_2} - \frac{1}{\xi}\int_{x_1}^{x_2} f'(x)\sin(x\,\xi)\,dx$$

$(0 < x_1 < x_2 < \infty)$, where

$$\frac{1}{\xi}\left[f(x)\sin(x\,\xi)\right]_{x_1}^{x_2} = \frac{1}{\xi} f(x_2)\sin(x_2\,\xi) - x_1 f(x_1)\frac{\sin(x_1\,\xi)}{x_1 \xi}.$$

Since $|\sin(x)|$, $|\frac{\sin(x)}{x}| \le 1$ $(x \in \mathbb{R})$, it follows from (26) and (27) that

$$\lim_{\substack{x_1 \to 0 \\ x_2 \to \infty}} \frac{1}{\xi}\left[f(x)\sin(x\,\xi)\right]_{x_1}^{x_2} = 0.$$

Hence

$$\int_0^\infty f(x)\cos(x\,\xi)\,dx = -\frac{1}{\xi}\int_0^\infty f'(x)\sin(x\,\xi)\,dx.$$

Using the same technique as in [45, Eq. 4] we find

$$\begin{aligned} -\int_0^\infty f'(x)\sin(x\,\xi)\,dx &= -\sum_{j=0}^\infty \int_{\frac{2\pi j}{\xi}}^{\frac{2\pi(j+1)}{\xi}} f'(x)\sin(x\,\xi)\,dx \\ &= \frac{1}{\xi}\sum_{j=0}^\infty \int_0^\pi \left[f'\left(\frac{2\pi j+\theta}{\xi}+\frac{\pi}{\xi}\right) - f'\left(\frac{2\pi j+\theta}{\xi}\right)\right]\sin(\theta)\,d\theta. \end{aligned}$$

Since $\sin(\theta) \ge 0$ on $[0, \pi]$ and $f'$ is non-decreasing, it follows that

$$\int_0^\infty f(x)\cos(x\,\xi)\,dx = -\frac{1}{\xi}\int_0^\infty f'(x)\sin(x\,\xi)\,dx \ge 0.$$

$\square$

Up to this point, we stipulated that the (generalised) positive definite functions must be in $\mathrm{L}^1(\mathbb{R}^d)$. This assumption ensures both the existence of the integral (21) for $\psi \in \mathrm{L}^2(\mathbb{R}^d)$ and the pointwise existence of $\hat{f}$. In the following we show that the generalised definition of positive definiteness can be localised, extending it from $\mathrm{L}^1(\mathbb{R}^d)$ to functions in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ or in $\mathrm{L}^1(K)$ for some closed, bounded set $K \subset \mathbb{R}^d$.

Let $I = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_d, b_d] \subset \mathbb{R}^d$ and define $\|I\| := [-|I_1|, |I_1|] \times [-|I_2|, |I_2|] \times \ldots \times [-|I_d|, |I_d|] \subset \mathbb{R}^d$ where $|I_i| = b_i - a_i$ denotes the length of the interval $I_i$. Let $f \in \mathrm{L}^1(\|I\|)$. Then, similarly to Lemma 2.5.1, for any $\phi \in \mathrm{L}^2(I)$,

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{x} - \mathbf{y}) \phi(\mathbf{x}) \overline{\phi(\mathbf{y})} \, d\mathbf{x} d\mathbf{y} = \int_{\|I\|} f(\mathbf{z}) \, \phi * \phi^*(\mathbf{z}) \, d\mathbf{z}, \tag{28}$$

since $\phi * \phi^*$ has support in $\|I\|$. The existence of the integral is guaranteed by the fact that $f \in \mathrm{L}^1(\|I\|)$. By Theorem 2.5.1, if the Fourier transform of $f\chi_{\|I\|}$ is non-negative, then $f\chi_{\|I\|} \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d)) \subset \mathrm{P}(\mathrm{L}^2(I))$, which in turn shows the non-negativity of the integral in (28).

The next result is a local variant of Corollary 2.7.2, based on the natural embedding of $\mathrm{L}^p(K)$ into $\mathrm{L}^p(\mathbb{R})$. We need a further technical condition at the end-point of the interval.

**Corollary 2.7.3** *Let $I = [a, b] \subset \mathbb{R}$ be any closed, bounded interval, and $|I| = b - a$ its length. Let $f \in \mathrm{L}^1([-|I|, |I|])$ be a function with the following properties.*

*i. $f$ is locally absolutely continuous on $(0, |I|]$, and $f' \in \mathrm{L}^1_{\mathrm{loc}}((0, |I|])$ has a non-positive, non-decreasing representative.*

*ii. $f(x) = f(-x)$ $(x \in [-|I|, |I|])$.*

*iii. $f(x) \geq 0$ $(x \in [-|I|, |I|])$.*

*iv. $f(|I|) = 0$ if $f'(|I|) = 0$.*

*Then, $f \in \mathrm{P}(\mathrm{L}^2(I))$.*

*Proof.* Define

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } |x| \leq |I| \\ f(|I|) e^{\frac{f'(|I|)}{f(|I|)}(|x| - |I|)} & \text{otherwise} \end{cases}$$

if $f(|I|) \neq 0$; if $f(|I|) = 0$, we set $\tilde{f}(x) = 0$ for $|x| > |I|$.

Then, the function $\tilde{f}$ satisfies the hypotheses of Corollary 2.7.2, and hence, is an element of $\mathrm{P}(\mathrm{L}^2(\mathbb{R})) \subset \mathrm{P}(\mathrm{L}^2(I))$. Moreover, $\tilde{f}(x) = f(x)$ ($x \in [-|I|, |I|]$), so $f \in \mathrm{P}(\mathrm{L}^2(I))$. $\square$

*Remark.* If $f'(|I|) = 0$ and $f(|I|) \neq 0$, then it is not possible to find an extension of the function $f$ from $[-|I|, |I|]$ to the whole real line which is continuous, integrable and has a derivative with a non-decreasing representative.

It follows immediately from Corollary 2.7.1 that if $f \in \mathrm{CM}$ and $g = f(\cdot^2) \in \mathrm{L}^1(\mathbb{R})$, then $g$ is positive definite for $\mathrm{L}^2(\mathbb{R})$. We remark that the result of squaring, or taking the square root of, the argument in a completely monotone function will, in general, not be a completely monotone function.

If we do not square the argument, but just extend the completely monotone function to an even function on the line, then the resulting function will satisfy the hypotheses of Corollary 2.7.2, yielding the following corollary, which is similar to Corollary 2.7.1 for functions defined on $\mathbb{R}$. However, the function $g$ of Corollary 2.7.1 (again, with $d = 1$), with a squared argument, does not satisfy property i. in Corollary 2.7.2, since $g'(x) = 2xf'(x^2)$ is not non-decreasing on $(0, \infty)$; hence, for functions defined on the real line, Corollary 2.7.1 cannot be obtained in this simple way.

**Corollary 2.7.4** *Let $f \in \mathrm{CM}$. If $g = f(|\cdot|) \in \mathrm{L}^1(\mathbb{R})$, then $g \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}))$.*

Moreover, we have the following localised versions.

**Corollary 2.7.5** *Let $I \subset \mathbb{R}$ be any closed, bounded interval. Let $f \in \mathrm{CM}$ be non-constant. If $g = f(|\cdot|) \in \mathrm{L}^1([-|I|, |I|])$, then $g \in \mathrm{P}(\mathrm{L}^2(I))$.*

*Proof.* If $f \in \mathrm{CM}$, then by [35, Remark 1.5], $f^{(n)}(x) \neq 0$ for all $n \geq 1$ and all $x > 0$ unless $f$ is identically constant. Thus $g$ satisfies the hypotheses of Corollary 2.7.3. $\square$

**Corollary 2.7.6** *Let $f \in \mathrm{CM}$ be non-constant. If $g = f(|\cdot|) \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$, then $g \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$.*

*Proof.* For any $\phi \in \mathrm{L}^2_0(\mathbb{R})$,

$$\int_{\mathbb{R}} \int_{\mathbb{R}} g(x-y)\phi(x)\overline{\phi(y)}\, dxdy = \int_{\mathrm{I}} \int_{\mathrm{I}} g(x-y)\phi(x)\overline{\phi(y)}\, dxdy, \tag{29}$$

where I denotes a closed, bounded interval which includes the compact support of $\phi$. Since $g \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$, it follows that $g \in \mathrm{L}^1([-|I|, |I|])$, and by Corollary 2.7.5 the integral in (29) is non-negative. $\square$

Completely monotone functions can be obtained as derivatives of Bernstein functions [35, p.18]. Taking functions $f_i$ from the list of Bernstein functions in [35, Chapter 15], the following derived functions $g_i = f_i'(|\cdot|)$ (up to multiplication by a positive normalising constant) are elements of $\mathrm{P}(\mathrm{L}^2_0(\mathbb{R})) \setminus \mathrm{P}_{\mathrm{C},1}$ by Corollary 2.7.6.

$$
\begin{aligned}
g_1(x) &= |x|^{-\alpha},\ 0 < \alpha < 1;\\
g_8(x) &= |x|^{\alpha-1}/(1+|x|)^{\alpha+1},\ 0 < \alpha < 1;\\
g_{11}(x) &= \Big(\alpha|x|^{\alpha-1}(1-|x|^{\beta}) - \beta|x|^{\beta-1}(1-|x|^{\alpha})\Big)/(1-|x|^{\alpha})^2,\\
&\qquad 0 < \alpha < \beta < 1\,;\\
g_{16}(x) &= \big(\alpha_1|x|^{-\alpha_1-1} + \ldots + \alpha_n|x|^{-\alpha_n-1}\big)\,/\,\big(|x|^{-\alpha_1} + \ldots + |x|^{-\alpha_n}\big)^2,\\
&\qquad 0 \le \alpha_1, \ldots, \alpha_n \le 1;\\
g_{19}(x) &= \Big(1 - (\lambda\sqrt{|x|} - 1)e^{-\lambda\sqrt{|x|}}\Big)\,/\sqrt{|x|}\,,\ \lambda > 0;\\
g_{23}(x) &= |x|\,(1+1/|x|)^{1+|x|}\log(1+1/|x|) \qquad\qquad (x \in \mathbb{R}\setminus\{0\}).
\end{aligned}
$$

*Remark.* It is not the case that all functions in $\mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$ are of the form described in Corollary 2.7.6. In other words, there exist functions in $\mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$ which are not the even reflection of a non-constant, completely monotone function. For example, by Proposition 2.3.1, the cosine function is positive definite with respect to $\mathrm{L}^2_0(\mathbb{R})$. The same is true for functions in $\mathrm{P}(\mathrm{L}^2(\mathbb{R}))$. The inverse hyperbolic cosine function has a non-negative Fourier transform, and thus is positive definite for $\mathrm{L}^2(\mathbb{R})$ by Theorem 2.5.1, yet, its second derivative changes sign at $x \simeq 0.8815$.

The following result is a direct consequence of Corollary 2.7.1, and provides a basis for finding examples of functions in $\mathrm{P}(\mathrm{L}^2_0(\mathbb{R}^d))$ with $d \ge 2$, which are unbounded at zero.

**Corollary 2.7.7** *Let $f \in \mathrm{CM} \cap \mathrm{L}^1_{\mathrm{loc}}((0,\infty))$. For any $s > 0$, define*

$$g(\mathbf{x}) = f(\|\mathbf{x}\|^2)e^{-s\|\mathbf{x}\|^2} \quad (\mathbf{x} \in \mathbb{R}^d). \tag{30}$$

*Then, $g \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}^d))$ for any $d \ge 2$.*

*Proof.* Let $s > 0$ and $d \ge 2$. Since the product of completely monotone functions is completely monotone, see Section 2.2, it follows that $fe^{-s}|_{(0,\infty)} \in$

CM, where $e^{-s}|_{(0,\infty)}$ denotes the restriction of $e^{-s}$ to the domain $(0,\infty)$. Moreover, $g \in \mathrm{L}^1(\mathbb{R}^d)$; for a change of variables to polar co-ordinates gives,

$$\begin{aligned}\int_{\mathbb{R}^d} g(\mathbf{x})\,d\mathbf{x} &= \omega_{d-1}\int_0^1 f(r^2)e^{-sr^2}r^{d-1}\,dr + \omega_{d-1}\int_1^\infty f(r^2)e^{-sr^2}r^{d-1}\,dr \\ &\leq \frac{\omega_{d-1}}{2}\int_0^1 f(x)\,x^{\frac{d-2}{2}}\,dx + \omega_{d-1}f(1)\int_1^\infty e^{-sr^2}r^{d-1}\,dr < \infty,\end{aligned}$$

where $\omega_{d-1}$ denotes the volume of the unit $(d-1)$-dimensional ball and $r$ the radius. Thus, $g \in \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$ by Proposition 2.7.1. $\square$

The derivatives of the functions listed in [35, Chapter 15] are completely monotone and locally integrable on $(0,\infty)$. For those with a singularity at the origin, the corresponding functions in $\mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$ ($d \geq 2$) can be constructed using (30).

## 2.8 Positive definite distributions and an alternative proof of Theorem 2.5.1

The concept of positive definite functions was extended to positive definite distributions by L. Schwartz [41, Chapter VII, §9]. We introduce the notion of a positive definite distribution and present Schwartz's analogue of Bochner's theorem, which states that a distribution is positive definite (and tempered) if and only if it is the Fourier transform of a non-negative measure of slow growth, i.e. such that the measure of balls is polynomially bounded in terms of the radius. Using this result, we then provide an alternative proof of Theorem 2.5.1.

Firstly, we introduce the following notation. Let $\mathbb{N}^d$ denote the set of all $d$-tuples of natural numbers and for $f : \mathbb{R}^d \to \mathbb{C}$ and $\mathbf{u} = (u_1, u_2, \ldots, u_d) \in \mathbb{N}^d$, define

$$D^{\mathbf{u}}(f(\mathbf{x})) := \frac{\delta^{|\mathbf{u}|}}{\delta x_1^{u_1} \ldots \delta x_d^{u_d}}\, f(x_1, x_2, \ldots, x_d) \quad (\mathbf{x} \in \mathbb{R}^d),$$

where $|\mathbf{u}| = \sum_{i=1}^d u_i$.

Let $\mathcal{D}(\mathbb{R}^d)$ denote the set $\mathrm{C}_0^\infty(\mathbb{R}^d)$ with the topology usual for the theory of distributions: a sequence of functions $f_n \in \mathcal{D}(\mathbb{R}^d)$ converges to $f$ if and only if the supports of $f$ and all the $f_n$'s lie inside a common compact set $K \subset \mathbb{R}^d$, and $D^{\mathbf{u}}f_n$ converges uniformly to $D^{\mathbf{u}}f$ for each multi-index $\mathbf{u} \in \mathbb{N}^d$, as $n \to \infty$. A distribution (or generalised function) is a continuous linear functional on $\mathcal{D}(\mathbb{R}^d)$. The space of all continuous linear functionals on $\mathcal{D}(\mathbb{R}^d)$

is denoted by $\mathcal{D}'(\mathbb{R}^d)$. Positive definiteness is now defined in the following sense.

**Definition 2.8.1** A distribution $\mathrm{T} \in \mathcal{D}'(\mathbb{R}^d)$ is said to be positive definite if $\mathrm{T}(\phi * \phi^*) \geq 0$ for all $\phi \in \mathcal{D}(\mathbb{R}^d)$, where $\phi^*(\mathbf{x}) = \overline{\phi(-\mathbf{x})}$ for any $\mathbf{x} \in \mathbb{R}^d$.

In order to see why this definition can be considered an extension of Definition 2.1.1, observe the distribution $\mathrm{T}_f$ associated with any locally integrable function $f$;

$$\mathrm{T}_f(\phi) = \int_{\mathbb{R}^d} f(\mathbf{x})\phi(\mathbf{x})\, d\mathbf{x} \qquad (\phi \in \mathcal{D}(\mathbb{R}^d)). \tag{31}$$

In this particular case, for any $\phi \in \mathcal{D}(\mathbb{R}^d)$,

$$\mathrm{T}_f(\phi * \phi^*) = \int_{\mathbb{R}^d} f(\mathbf{x})(\phi * \phi^*)(\mathbf{x})\, d\mathbf{x} = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\phi(\mathbf{x})\overline{\phi(\mathbf{y})}\, d\mathbf{x}d\mathbf{y},$$

by Lemma 2.5.1. Hence, for $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$, $\mathrm{T}_f$ is a positive definite distribution if and only if $f \in \mathrm{P}(\mathrm{C}_0^\infty(\mathbb{R}^d))$. Moreover, by Proposition 2.4.6, if $f$ is continuous, then $\mathrm{T}_f$ is positive definite if and only if $f$ is classically positive definite as in Definition 2.1.1.

Next, we define what it means for a distribution to be tempered. Recall that $\mathcal{S}(\mathbb{R}^d)$ denotes the Schwartz space of rapidly decreasing functions on $\mathbb{R}^d$. That is,

$$\mathcal{S}(\mathbb{R}^d) = \left\{ f \in \mathrm{C}^\infty(\mathbb{R}^d) \,\middle|\, \|f\|_{\mathbf{u},\mathbf{v}} < \infty \text{ for any } \mathbf{u},\mathbf{v} \in \mathbb{N}^d \right\},$$

where

$$\|f\|_{\mathbf{u},\mathbf{v}} = \sup_{\mathbf{x}\in\mathbb{R}^d} \left|\mathbf{x}^{\mathbf{u}} D^{\mathbf{v}}(f(\mathbf{x}))\right| \qquad (\mathbf{u},\mathbf{v} \in \mathbb{N}^d,\ f \in \mathcal{S}(\mathbb{R}^d))$$

and $\mathbf{x}^{\mathbf{u}} = \prod_{i=1}^d x_i^{u_i}$, as in standard multi-index notation.

**Definition 2.8.2** The dual space of $\mathcal{S}(\mathbb{R}^d)$, denoted by $\mathcal{S}'(\mathbb{R}^d)$, is called the space of *tempered distributions.*

For $m, n \in \mathbb{N}$ and $f \in \mathcal{S}(\mathbb{R}^d)$, let

$$|||f|||_{m,n} := \sum_{\substack{|\mathbf{u}|\leq m \\ |\mathbf{v}|\leq n}} \|f\|_{\mathbf{u},\mathbf{v}}.$$

This defines a family of norms on $\mathcal{S}(\mathbb{R}^d)$. For if $f \in \mathcal{S}(\mathbb{R}^d)$ and $|||f|||_{m,n} = 0$ for some $m, n \in \mathbb{N}$, then for each $\mathbf{x} \neq \mathbf{0}$, we have $D^{\mathbf{v}}(f(\mathbf{x})) = 0$ for all

$\mathbf{v} \in \mathbb{N}^d$ such that $|\mathbf{v}| \leq n$. Moreover, since $D^{\mathbf{v}}(f)$ is continuous, it follows that $D^{\mathbf{v}}(f(\mathbf{x})) = 0$ for any $\mathbf{x} \in \mathbb{R}^d$. Hence, $f \in \mathcal{S}(\mathbb{R}^d)$ is a polynomial which vanishes at infinity and therefore, $f = 0$.

For a linear functional T on $\mathcal{S}(\mathbb{R}^d)$ to be in $\mathcal{S}'(\mathbb{R}^d)$ it must be continuous. By Theorem [30, Th. V.2], this is equivalent to requiring $C > 0$ and $m, n \in \mathbb{N}$, such that $|\mathrm{T}(\phi)| \leq C|||\phi|||_{m,n}$ for all $\phi \in \mathcal{S}(\mathbb{R}^d)$. Using this observation, we now present some examples of tempered distributions.

1. Let $f \in \mathrm{L}^1(\mathbb{R}^d)$ and consider the functional

$$\mathrm{T}_f(\phi) = \int_{\mathbb{R}^d} f(\mathbf{x})\phi(\mathbf{x})\, d\mathbf{x} \qquad (\phi \in \mathcal{S}(\mathbb{R}^d)). \tag{32}$$

$\mathrm{T}_f$ is clearly linear and $|\mathrm{T}_f(\phi)| \leq C|||\phi|||_{0,0}$ with $C = \|f\|_1$, for any $\phi \in \mathcal{S}(\mathbb{R}^d)$. Thus, the distribution associated with an integrable function is tempered.

2. Let $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ be such that there exists $n \in \mathbb{N}$ with $\int_{\mathbb{R}^d} |f(\mathbf{x})|/(1+\|\mathbf{x}\|^2)^n\, d\mathbf{x} < \infty$, where $\|\cdot\|$ denotes the Euclidean norm. Consider the functional defined in (31) with test functions in $\mathcal{S}(\mathbb{R}^d)$.

$\mathrm{T}_f$ is linear and for any $\phi \in \mathcal{S}(\mathbb{R}^d)$,

$$|\mathrm{T}_f(\phi)| \leq \int_{\mathbb{R}^d} \frac{|f(\mathbf{x})|}{(1+\|\mathbf{x}\|^2)^n} (1+\|\mathbf{x}\|^2)^n |\phi(\mathbf{x})|\, d\mathbf{x}.$$

Let $\psi \in \mathcal{S}(\mathbb{R}^d)$. Then, for any $\mathbf{x} \in \mathbb{R}^d$,

$$\begin{aligned}
(1+\|\mathbf{x}\|^2)^n|\psi(\mathbf{x})| &= \sum_{i=0}^{n} \binom{n}{i} \|\mathbf{x}\|^{2n}|\psi(\mathbf{x})| \\
&\leq K\big(|\psi(\mathbf{x})| + \|\mathbf{x}\|^2|\psi(\mathbf{x})| + \ldots + \|\mathbf{x}\|^{2n}|\psi(\mathbf{x})|\big) \\
&\leq K \sum_{|\mathbf{u}|\leq 2n} |\mathbf{x}^{\mathbf{u}}\psi(\mathbf{x})| \leq K \sum_{|\mathbf{u}|\leq 2n} \sup_{\mathbf{z}\in\mathbb{R}^d} |\mathbf{z}^{\mathbf{u}}\psi(\mathbf{z})| < \infty,
\end{aligned}$$

where $K$ is chosen so that $K \geq \binom{n}{i}$ for any $i = 0, \ldots, n$, and $\binom{n}{i} = n!/i!(n-i)!$ denotes the standard binomial coefficient. Thus, $|\mathrm{T}_f(\psi)| \leq C|||\psi|||_{m,0}$ with $C = K\int_{\mathbb{R}^d} |f(\mathbf{x})|/(1+\|\mathbf{x}\|^2)^n d\mathbf{x}$ and $m = 2n$.

**Definition 2.8.3** ([41, p.97]). A measure $\mu$ on $\mathbb{R}^d$ is said to be of slow growth, or polynomially bounded, if there exists $n \in \mathbb{N}$ such that

$$\int_{\mathbb{R}^d} \frac{|d\mu(\mathbf{x})|}{(1+\|\mathbf{x}\|^2)^n} < \infty.$$

3. Let $\nu$ denote a non-negative, slow-growing measure on $\mathbb{R}^d$. Similarly to example 2 above, the distribution associated with $\nu$,

$$\mathrm{T}_\nu(\phi) = \int_{\mathbb{R}^d} \phi(\mathbf{x})\, d\nu(\mathbf{x}) \qquad (\phi \in \mathcal{S}(\mathbb{R}^d)),$$

is tempered [41, Th. VII].

The following result proceeds directly from Lemmas 2.5.1 and 2.5.2, and demonstrates that the tempered distribution associated with an integrable function in $\mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ is positive definite.

**Proposition 2.8.1** *Let $f \in \mathrm{L}^1(\mathbb{R}^d)$. Then, $\mathrm{T}_f$, as defined in* (32)*, is positive definite if and only if $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$.*

Next, we define the Fourier transform on $\mathcal{S}'(\mathbb{R}^d)$.

**Definition 2.8.4** Let $\mathrm{T} \in \mathcal{S}'(\mathbb{R}^d)$. Then, the Fourier transform of T, denoted by $\hat{\mathrm{T}}$, is the tempered distribution defined by $\hat{\mathrm{T}}(\phi) = \mathrm{T}(\hat{\phi})$ $(\phi \in \mathcal{S}(\mathbb{R}^d))$.

*Remark.* For $f \in \mathrm{L}^1(\mathbb{R}^d)$ and $\mathrm{T}_f$ as defined in (32), it follows directly from the Fubini theorem that $\hat{\mathrm{T}}_f = \mathrm{T}_{\hat{f}}$. For, for any $\phi \in \mathcal{S}(\mathbb{R}^d)$,

$$\mathrm{T}_f(\hat{\phi}) = \int_{\mathbb{R}^d} f(\mathbf{x}) \left( (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \phi(\mathbf{z})\, e^{-i\mathbf{x}\cdot\mathbf{z}}\, d\mathbf{z} \right) d\mathbf{x} = \int_{\mathbb{R}^d} \hat{f}(\mathbf{z}) \phi(\mathbf{z})\, d\mathbf{z} = \mathrm{T}_{\hat{f}}(\phi).$$

The central result on the theory of positive definite distributions is the Bochner-Schwartz theorem, which characterises positive definite, tempered distributions as Fourier transforms of non-negative, slow-growing measures.

**Theorem 2.8.1** (Schwartz, [41, Th. XVIII]). *A distribution $T \in \mathcal{D}'(\mathbb{R}^d)$ is positive definite if and only if $T \in \mathcal{S}'(\mathbb{R}^d)$ and $T$ is the Fourier transform of a non-negative, slow-growing measure.*

Theorem 2.8.1 infers that certain distributions in $\mathcal{D}'(\mathbb{R}^d)$, namely those which are positive definite, are necessarily tempered. Moreover, the result can be used to prove Theorem 2.5.1 as follows.

*Proof of Theorem 2.5.1.* Since $f \in \mathrm{L}^1(\mathbb{R}^d)$, then $\mathrm{T}_f(\phi) = \int_{\mathbb{R}^d} f(\mathbf{x})\phi(\mathbf{x})\, d\mathbf{x}$ $(\phi \in \mathcal{S}(\mathbb{R}^d))$ is a tempered distribution, see example 1 above.

Suppose, $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$. Then, $\mathrm{T}_f$ is positive definite by Proposition 2.8.1 and thus, by Theorem 2.8.1,

$$\hat{\mathrm{T}}_f(\phi) = \mathrm{T}_{\hat{f}}(\phi) = \int_{\mathbb{R}^d} \hat{f}(\mathbf{x})\phi(\mathbf{x})\, d\mathbf{x} = \int_{\mathbb{R}^d} \phi(\mathbf{x})\, d\mu(\mathbf{x}) = \mathrm{T}_\mu(\phi) \qquad (33)$$

for any $\phi \in \mathcal{S}(\mathbb{R}^d)$, where $\mu$ is a non-negative, slow-growing measure on $\mathbb{R}^d$. It follows that $\hat{f} \geq 0$ by proof by contradiction; for if $\hat{f}(\mathbf{z}) < 0$ at some point $\mathbf{z} \in \mathbb{R}^d$, then there exists $\delta > 0$ such that $\hat{f}(\mathbf{x}) < 0$ for all $\|\mathbf{x} - \mathbf{z}\| < \delta$, since $\hat{f}$ is continuous. Let

$$\psi(\mathbf{x}) = \begin{cases} \exp\left[\left(\|\mathbf{x}-\mathbf{z}\|^2 - \delta^2\right)^{-1}\right] & \text{if} \quad \|\mathbf{x}-\mathbf{z}\| < \delta \\ 0 & \text{otherwise} \end{cases} \qquad (\mathbf{x} \in \mathbb{R}^d).$$

Then $\psi \in \mathrm{C}_0^\infty(\mathbb{R}^\mathrm{d}) \subset \mathcal{S}(\mathbb{R}^\mathrm{d})$ and it follows by (33) that

$$0 \leq \int_{\mathbb{R}^d} \psi(\mathbf{x})\, d\mu(\mathbf{x}) = \int_{\|\mathbf{x}-\mathbf{z}\|<\delta} \hat{f}(\mathbf{x})\psi(\mathbf{x})\, d\mathbf{x} < 0,$$

which is a contradiction.

Conversely, suppose $\hat{f} \geq 0$. Since $f \in \mathrm{L}^1(\mathbb{R}^d)$, then $\hat{f}$ is bounded on $\mathbb{R}^d$. Take $\nu$ to be the measure whose density is $\hat{f}$. Then, $\nu$ is non-negative and slow-growing on $\mathbb{R}^d$, and

$$\hat{\mathrm{T}}_f(\phi) = \mathrm{T}_{\hat{f}}(\phi) = \int_{\mathbb{R}^d} \phi(\mathbf{x})\, d\nu(\mathbf{x}) = \mathrm{T}_\nu(\phi) \qquad (\phi \in \mathcal{S}(\mathbb{R}^d)).$$

Thus, $\mathrm{T}_f$ is positive definite by Theorem 2.8.1 and hence, $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ by Proposition 2.8.1. $\square$

# 3 On Conditionally Negative Definite Functions With A Singularity At Zero

The renowned Schoenberg theorem [39, Th. 2] establishes a connection between positive definite and conditionally negative definite functions. Using the framework outlined in Sections 2.3 and 2.4 as motivation, we consider the class of functions $\mathrm{CN}(J)$, which are conditionally negative definite with respect to a given set of test functions $J$. For suitably chosen $J$, $\mathrm{CN}(J)$ contains the classical conditionally negative definite functions, which take finite values at zero, as well as functions which are singular at the origin.

Our main result is Theorem 3.4.1, which is a generalisation of Schoenberg's theorem to functions in $\mathrm{P}(J)$ and $\mathrm{CN}(J)$, for $J = \mathrm{L}_0^2(\mathbb{R}^d)$. Several other results concerning the class $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ are also established. For example, we demonstrate that functions in $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ are locally integrable (see Lemma 3.3.1) and that $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ is a closed subset of $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ (see Lemma 3.3.2). Furthermore, we show that real-valued functions which are conditionally negative definite w.r.t. $\mathrm{L}_0^2(\mathbb{R}^d)$ can be approximated, in the $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ sense, by a sequence of infinitely differentiable, classically conditionally negative definite functions (see Lemma 3.4.3). Finally, using Theorem 3.4.1, we indicate how to construct numerous examples of singular functions $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$, such that $-f \notin \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$ (see Section 3.5). The results described in this section have appeared in the Journal of Mathematical Analysis and Applications in the form of the paper [27].

The structure of this section is as follows. We start with an overview of the conditionally negative definite functions as defined in the classical literature, see Section 3.1. In Section 3.3 we extend the definition of conditional negative definiteness to incorporate functions with a singularity at zero, and subsequently, develop the theory of these newly defined functions. Section 3.4 contains the proof of Theorem 3.4.1, split into a series of lemmas. In Section 3.5 we prove two corollaries to Theorem 3.4.1 and provide several algorithmic schemes for constructing functions in $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$.

## 3.1 Classical conditionally negative definite functions

Conditionally negative definite functions arise naturally in the theories of probability and potentials. The standard references for these functions are the monographs [3] and [4], where the term *conditionally* is dropped and the functions are called simply, *negative definite*. In other areas of the literature, however, see e.g. [35, Def. 4.3], negative definite functions are defined differently to those which are conditionally negative definite. We quash this confusion and solely focus our attention on the following definition of conditional negative definiteness.

**Definition 3.1.1** A function $f : \mathbb{R}^d \to \mathbb{C}$ is *conditionally negative definite* if $f$ is conjugate symmetric, that is $f(\mathbf{x}) = \overline{f(-\mathbf{x})}$ for all $\mathbf{x} \in \mathbb{R}^d$, and

$$\sum_{i,j=1}^{n} f(\mathbf{x}_i - \mathbf{x}_j)\, v_i\, \overline{v_j} \leq 0 \tag{34}$$

for all $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ and $v_1, v_2, \ldots, v_n \in \mathbb{C}$ satisfying $\sum_{i=1}^{n} v_i = 0$, with any $n \in \mathbb{N}$.

We shall denote the set of functions defined in Definition 3.1.1 by $\mathrm{CN}_{\mathrm{C},d}$. The conjugate symmetry of functions in $\mathrm{CN}_{\mathrm{C},d}$ is stipulated since unlike functions in $\mathrm{P}_{\mathrm{C},d}$, which are automatically conjugate symmetric by (5), the property no longer follows from the sum in (34) due to the extra constraint on the $v_i$s. For example, the function $f(x) = x$ $(x \in \mathbb{R})$ is non-conjugate symmetric, yet, for any $n \in \mathbb{N}$, $x_1, x_2, \ldots, x_n \in \mathbb{R}$ and $v_1, v_2, \ldots, v_n \in \mathbb{C}$ such that $\sum_{i=1}^{n} v_i = 0$,

$$\sum_{i,j=1}^{n} f(x_i - x_j)\, v_i\, \overline{v_j} = \sum_{i,j=1}^{n} x_i\, v_i\, \overline{v_j} - \sum_{i,j=1}^{n} x_j\, v_i\, \overline{v_j} = 0.$$

The same is true for $g(x) = (a + x)^2$, for any $a \in \mathbb{R}\backslash\{0\}$.

A simple example of a function in $\mathrm{CN}_{\mathrm{C},d}$ is $f = \|\cdot\|^2$, since for any $n \in \mathbb{N}$, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ with component form $\mathbf{x}_i = (x_{i_1}, x_{i_2}, \ldots, x_{i_d})$, and $v_1, v_2, \ldots, v_n \in \mathbb{C}$ such that $\sum_{i=1}^{n} v_i = 0$,

$$\sum_{i,j=1}^{n} \|\mathbf{x}_i - \mathbf{x}_j\|^2 v_i \overline{v_j} = \sum_{k=1}^{d} \sum_{i,j=1}^{n} \big(x_{i_k}^2 - 2\, x_{i_k} x_{j_k} + x_{j_k}^2\big) v_i \overline{v_j} = -2 \sum_{k=1}^{d} \left| \sum_{i=1}^{n} x_{i_k} v_i \right|^2 \leq 0.$$

It is clear that if $f$ is a classically positive definite function, then $-f$ is conditionally negative definite. The converse is not true. For example, $f(\mathbf{x}) = -|\mathbf{x}|^2$ $(\mathbf{x} \in \mathbb{R}^d)$ is not positive definite since $|f(\mathbf{x})| > f(\mathbf{0})$ for any $\mathbf{x} \in \mathbb{R}^d\backslash\{0\}$. Both classically positive definite and conditionally negative definite functions take finite values at zero. However, unlike positive definite functions, functions in $\mathrm{CN}_{\mathrm{C},d}$ can be unbounded away from the origin. Again, the example $f = |\cdot|^2$ demonstrates this. Conditionally negative definite functions need not be negative or continuous; the function $f(\mathbf{x}) = -1$

if $\mathbf{x} = \mathbf{0}$, $f(\mathbf{x}) = 0$ otherwise ($\mathbf{x} \in \mathbb{R}^d$) is in $\mathrm{CN}_{\mathrm{C},d}$, but not continuous; the negative cosine function is in $\mathrm{CN}_{\mathrm{C},1}$, but not non-positive.

The following are simple properties of functions in $\mathrm{CN}_{\mathrm{C},d}$.

i. $f \in \mathrm{CN}_{\mathrm{C},d}$ if and only if $\overline{f} \in \mathrm{CN}_{\mathrm{C},d}$.

ii. If $f_1, f_2, \ldots, f_n \in \mathrm{CN}_{\mathrm{C},d}$ and $c_i \geq 0$ for all $i = 1, \ldots, n$, then $\sum_{i=1}^n c_i f_i \in \mathrm{CN}_{\mathrm{C},d}$.

iii. If $f \in \mathrm{CN}_{\mathrm{C},d}$, then $f + \alpha \in \mathrm{CN}_{\mathrm{C},d}$ for any $\alpha \in \mathbb{C}$.

iv. If $f_n \in \mathrm{CN}_{\mathrm{C},d}$ for all $n \in \mathbb{N}$ and the pointwise limit, $\lim_{n\to\infty} f_n(\mathbf{x}) = f(\mathbf{x})$, exists for all $\mathbf{x} \in \mathbb{R}^d$, then $f \in \mathrm{CN}_{\mathrm{C},d}$.

These properties follow immediately from Definition 3.1.1. A direct consequence of the first two properties is that if $f$ is conditionally negative definite, then so is $\mathrm{Re}\,(f) = (f + \overline{f})/2$.

For real-valued functions we can use the following alternative definition of conditional negative definiteness.

**Definition 3.1.2** A function $f : \mathbb{R}^d \to \mathbb{R}$ is conditionally negative definite if $f(\mathbf{x}) = f(-\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^d$ and the inequality in (34) holds for all $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ and $v_1, v_2, \ldots, v_n \in \mathbb{R}$ satisfying $\sum_{i=1}^n v_i = 0$, with any $n \in \mathbb{N}$.

The functions defined in Definition 3.1.2 are automatically conditionally negative definite as in Definition 3.1.1. This can be seen by using (7) with $v_1, v_2, \ldots, v_n \in \mathbb{C}$ such that $\sum_{i=1}^n v_i = 0$, giving $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i = 0$.

As with positive definite functions, the definition of a conditionally negative definite function can be extended to functions and two-variable kernels on general topological spaces, groups and semigroups, see e.g. [2, 4, 5, 35]. However, we restrict our attention to the conditionally negative definite functions defined in Definition 3.1.1. In later sections we will mainly be interested in real-valued functions, in which case, Definitions 3.1.1 and 3.1.2 are interchangeable. Henceforth, when referring to classically conditionally negative definite functions or functions in $\mathrm{CN}_{\mathrm{C},d}$, we mean those defined in Definition 3.1.1.

The celebrated Schoenberg theorem establishes a relation between positive definite and conditionally negative definite functions.

**Theorem 3.1.1** (Schoenberg, [35, Prop. 4.4]). *A function $f : \mathbb{R}^d \to \mathbb{C}$ is conditionally negative definite if and only if for all $t > 0$, the function $g : \mathbf{x} \mapsto e^{-tf(\mathbf{x})}$ is positive definite.*

Theorem 3.1.1 stems from [39, Th. 2] where it is stipulated that $f$ is continuous and non-negative, and vanishes at the origin. Proofs of the result can be found in [2, Th. C.3.2] and [4, p.74], as well as [35, Prop. 4.4] and [39, Th. 2]. In Section 3.4 we generalise Theorem 3.1.1 to the case when the functions $f$ and $g$ can be singular at the origin.

## 3.2 Bernstein functions

The notion of a Bernstein function is thought to have originated in the potential theory school of A. Beurling and J. Deny. The name *Bernstein function* is not universally accepted in the literature, for example, Bochner [10] referred to Bernstein functions as *completely monotone mappings*, and many probabilists still choose to call them *Laplace exponents.* Bernstein functions are closely related to completely monotone functions, in fact, Schoenberg [38] defined them as primitives of functions in CM. We introduce the definition of the Bernstein functions presented in [35, Chapter 3] and provide a short introduction into their theory.

**Definition 3.2.1** A function $f : (0, \infty) \to [0, \infty)$ is a *Bernstein function* if $f \in \mathrm{C}^\infty((0, \infty))$ and

$$(-1)^{n-1} f^{(n)} \geq 0 \text{ on } (0, \infty)$$

for all $n \in \mathbb{N}$ [35, Def. 3.1].

In particular, any Bernstein function is non-negative and non-decreasing. Unlike completely monotone functions, Bernstein functions are always bounded at zero, however, they may or may not be bounded away from the origin. For example, both $f_1(x) = \sqrt{x}$ and $f_2(x) = x/(x + 1)$ $(x \in (0, \infty))$ are Bernstein functions. The family of all Bernstein functions is denoted by BF. Similarly to bounded completely monotone functions, functions in BF can be extended continuously to $[0, \infty)$. Due to the monotonicity of $f \in \mathrm{BF}$, this can be achieved by taking $f(0) := f(0+) = \lim_{x \to 0} f(x)$ [35, p. 28].

The derivative of a Bernstein function is completely monotone. The converse is only true if the primitive of a completely monotone function is non-negative. This condition fails, for example, for the completely monotone function $f(x) = x^{-\alpha}$ $(0 < \alpha < 1, x > 0)$. However, a non-negative $\mathrm{C}^\infty((0, \infty))$-function $f$, is a Bernstein function if and only if $f'$ is completely monotone. The following theorem of Bochner highlights further connections between Bernstein and completely monotone functions.

**Theorem 3.2.1** (Bochner, [10, p. 83]) *Let $f$ be a positive function on $(0,\infty)$. Then, the following assertions are equivalent.*

*i.* $f \in \mathrm{BF}$.

*ii.* $g \circ f \in \mathrm{CM}$ *for every* $g \in \mathrm{CM}$.

*iii.* $e^{-tf} \in \mathrm{CM}$ *for every* $t > 0$.

For a recent proof see e.g. [35, Th. 3.6]. Many corollaries to Theorem 3.2.1 may be found in [35, Chapters 3, 4].

The next theorem, given without proof, provides a useful characterisation of Bernstein functions.

**Theorem 3.2.2** ([35, Th. 3.2]). *A function $f : (0,\infty) \to [0,\infty)$ is a Bernstein function if, and only if, it admits the representation*

$$f(x) = a + bx + \int_{(0,\infty)} (1 - e^{-xt})\, \mu(dt), \tag{35}$$

*where $a, b \geq 0$ and $\mu$ is a non-negative measure on $(0,\infty)$ satisfying $\int_{(0,\infty)} \min(1,t)\, \mu(dt) < \infty$. In particular, the triplet $(a,b,\mu)$ determines $f$ uniquely and vice versa.*

Theorem 3.2.2 appears elsewhere in the literature, see e.g. [10, Chapter 4], [3, p. 114]. Equation (35) is often called the *Lévy-Khintchine representation* of $f$. The measure $\mu$ and the triplet $(a,b,\mu)$ are referred to as the Lévy measure and the Lévy triplet of the Bernstein function $f$, respectively, see e.g. [35, Rem. 3.3 (i)], [3, Chapter 4].

We note that the integrability condition $\int_{(0,\infty)} \min(1,t)\, \mu(dt) < \infty$ ensures that the integral in (35) exists for all $x > 0$ [35, Rem. 3.3 (iii)]. It follows from (35) that $a = f(0+)$, and using the dominated convergence theorem it can be shown that $b = \lim_{x\to\infty} f(x)/x$, [35, Rem. 3.3 (iv)].

Some properties of Bernstein functions are listed below.

i. If $f_1, f_2, \ldots, f_n \in \mathrm{BF}$ and $c_i \geq 0$ for all $i = 1, \ldots, n$, then $\sum_{i=1}^{n} c_i f_i \in \mathrm{BF}$.

ii. If $f_n \in \mathrm{BF}$ for all $n \in \mathbb{N}$ and the pointwise limit, $\lim_{n\to\infty} f_n(x) = f(x)$, exists for all $x > 0$, then $f \in \mathrm{BF}$.

iii. If $f_1, f_2 \in \mathrm{BF}$, then $f_1 \circ f_2 \in \mathrm{BF}$.

iv. If $f \in \mathrm{BF}$, then $x \mapsto f(x)/x$ is in CM.

v. $f \in \mathrm{BF}$ is bounded if, and only if, in (35) $b = 0$ and $\mu(0, \infty) < \infty$.

vi. Let $f_1$, $f_2 \in \mathrm{BF}$ and $\alpha$, $\beta \in (0,1)$ be such that $\alpha + \beta \leq 1$. Then, $x \mapsto f_1(x^\alpha) f_2(x^\beta)$ is in BF.

The first property follows directly from Definition 3.2.1 or, alternatively, from the representation in (35). The remaining properties can be proved using Theorems 3.2.1 and 3.2.2, see e.g. [35, Cor. 3.7].

Another theorem belonging to Schoenberg, which links Bernstein functions and conditionally negative definite functions in a similar fashion to the bounded completely monotone functions and positive definite functions in Theorem 2.2.2, is as follows.

**Theorem 3.2.3** (Schoenberg, [38, Eq. 5.14]). *A function $\psi : [0, \infty) \to [0, \infty)$ is a Bernstein function if and only if for all $d \in \mathbb{N}$, the function $f = \psi(\|\cdot\|^2) : \mathbb{R}^d \to [0, \infty)$ is continuous and conditionally negative definite.*

In particular, if $\psi \in \mathrm{BF}$, then $f = \psi(\|\cdot\|^2) : \mathbb{R}^d \to [0, \infty)$ is continuous and conditionally negative definite for any $d \in \mathbb{N}$. In Corollary 3.5.1 we show that such functions are also conditionally negative definite in an extended sense. For a recent proof of Theorem 3.2.3 see e.g. [35, Th. 12.14], [10, p. 99].

## 3.3 Conditional negative definiteness in the extended sense

We now introduce a definition of conditional negative definiteness which allows for functions with a singularity at the origin. As in Section 2.4, let $J$ be a set of complex-valued measurable functions on $\mathbb{R}^d$. Again, this includes functions defined on a non-empty, measurable subset of $\mathbb{R}^d$, which we consider to be extended by zero to the whole of $\mathbb{R}^d$. Motivated by Definition 2.4.1 we define an extended notion of conditionally negative definite functions as follows.

**Definition 3.3.1** A function $f : \mathbb{R}^d \to \mathbb{C}$ is called *conditionally negative definite w.r.t. $J$* if $f$ is conjugate symmetric a.e., that is $f(\mathbf{x}) = \overline{f(-\mathbf{x})}$ f.a.a. $\mathbf{x} \in \mathbb{R}^d$, and for every $\phi \in J$ satisfying $\int_{\mathbb{R}^d} \phi(\mathbf{x})\, d\mathbf{x} = 0$, the integral

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{x} - \mathbf{y}) \phi(\mathbf{x}) \overline{\phi(\mathbf{y})}\, d\mathbf{x} d\mathbf{y} \tag{36}$$

exists (in the Lebesgue sense) and is non-positive.

Let $\mathrm{CN}(J)$ denote the class of all functions which are conditionally negative definite w.r.t. the set $J$. Similarly to as in Sections 2.3 and 2.4, for certain spaces of functions $J$, Definition 3.3.1 enables us to extend the concept of conditional negative definiteness to functions which have a singularity at zero. Again, we shall mainly consider the spaces $J = \mathrm{L}^p(\mathbb{R}^d)$ (and their local versions) for various values of $p$.

The following properties proceed directly from Definition 3.3.1.

i. $f \in \mathrm{CN}(J) \Leftrightarrow \overline{f} \in \mathrm{CN}(J)$ if $J$ is closed under complex conjugation.

ii. If $f_1, f_2, \ldots, f_n \in \mathrm{CN}(J)$ and $c_i \geq 0$ $(i = 1, \ldots, n)$, then $\sum_{i=1}^n c_i f_i \in \mathrm{CN}(J)$.

iii. If $f \in \mathrm{P}(J)$, then $-f \in \mathrm{CN}(J)$.

iv. If $f \in \mathrm{CN}(J)$, then $f + \alpha \in \mathrm{CN}(J)$ for any $\alpha \in \mathbb{C}$.

v. If $J_1 \subseteq J_2$, then $\mathrm{CN}(J_2) \subseteq \mathrm{CN}(J_1)$.

In Corollary 3.5.1 we show that under certain conditions on our function, Definitions 3.1.1 and 3.3.1 coincide. In particular, a real-valued, continuous function is classically conditionally negative definite if and only if it is conditionally negative definite w.r.t. $\mathrm{C}_0(\mathbb{R}^d)$.

Recall that for $p \in [1, \infty) \cup \{\infty\}$, $\mathrm{L}_0^p(\mathbb{R}^d)$ denotes the subspace of functions in $\mathrm{L}^p(\mathbb{R}^d)$ with compact essential support, and by $f_n \xrightarrow{\mathrm{L}^p_{\mathrm{loc}}} f$ as $n \to \infty$, we mean that $f_n$ converges to $f$ in the $\mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^d)$ sense as $n \to \infty$; that is,

$$\lim_{n \to \infty} \int_K |f_n(\mathbf{x}) - f(\mathbf{x})|^p \, d\mathbf{x} = 0$$

for any compact set $K \subset \mathbb{R}^d$. Next, we demonstrate that as $p$ increases from 1 to 2, $\mathrm{CN}(\mathrm{L}_0^p(\mathbb{R}^d))$ increases from a smaller class of functions to a larger such class; but for all $p \geq 2$, $\mathrm{CN}(\mathrm{L}_0^p(\mathbb{R}^d))$ remains the same.

**Proposition 3.3.1** *For any $p \in [1, 2]$, $\mathrm{CN}(\mathrm{L}_0^p(\mathbb{R}^d)) \subseteq \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$.*

*Proof.* This follows directly from property v. above. □

**Proposition 3.3.2** *For any $p \in (2, \infty]$ and $r \in [0, \infty]$, $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d)) = \mathrm{CN}(\mathrm{L}_0^p(\mathbb{R}^d)) = \mathrm{CN}(\mathrm{C}_0^r(\mathbb{R}^d))$.*

*Proof.* Let $p \in (2, \infty]$, $r \in [0, \infty]$. Since $\mathrm{C}_0^r(\mathbb{R}^d) \subset \mathrm{L}_0^p(\mathbb{R}^d) \subset \mathrm{L}_0^2(\mathbb{R}^d)$, it follows immediately that $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d)) \subset \mathrm{CN}(\mathrm{L}_0^p(\mathbb{R}^d)) \subset \mathrm{CN}(\mathrm{C}_0^r(\mathbb{R}^d))$.

For the reverse implication, consider the following. Let $\phi \in \mathrm{L}_0^2(\mathbb{R}^d)$ be such that $\int_{\mathbb{R}^d} \phi(\mathbf{x})\, d\mathbf{x} = 0$. For $n \in \mathbb{N}$, let $\Psi_n$ denote the functions defined in (17), and set

$$\psi_n := \phi * \Psi_n \quad (n \in \mathbb{N}).$$

Then, $\psi_n \in \mathrm{C}_0^r(\mathbb{R}^d) \subset \mathrm{L}_0^p(\mathbb{R}^d)$ for all $r \in [0, \infty]$, $p \in [2, \infty]$, $n \in \mathbb{N}$, and by the Fubini theorem,

$$\begin{aligned}\int_{\mathbb{R}^d} \psi_n(\mathbf{x})\, d\mathbf{x} = \int_{\mathbb{R}^d} (\phi * \Psi_n)(\mathbf{x})\, d\mathbf{x} &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \phi(\mathbf{x}-\mathbf{y})\, \Psi_n(\mathbf{y})\, d\mathbf{y} d\mathbf{x} \\ &= \int_{\mathbb{R}^d} \phi(\mathbf{z})\, d\mathbf{z} \int_{\mathbb{R}^d} \Psi_n(\mathbf{y})\, d\mathbf{y} = 0.\end{aligned}$$

Moreover, $\psi_n \xrightarrow{\mathrm{L}^2_{\mathrm{loc}}} \phi$ as $n \to \infty$, by [42, Th. 1.18].

Let $r \in [0, \infty]$ and suppose $f \in \mathrm{CN}(\mathrm{C}_0^r(\mathbb{R}^d))$. Provided $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$, which it is by Lemma 3.3.1 below, the integral

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\phi(\mathbf{x})\overline{\phi(\mathbf{y})}\, d\mathbf{x} d\mathbf{y} = \int_{\mathbb{R}^d} (f^* * \phi)(\mathbf{y})\overline{\phi(\mathbf{y})}\, d\mathbf{y}$$

exists as a Lebesgue integral. Similarly to as in the proof of Proposition 2.3.4, it can be shown that

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\phi(\mathbf{x})\overline{\phi(\mathbf{y})}\, d\mathbf{x} d\mathbf{y} = \lim_{n\to\infty}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\psi_n(\mathbf{x})\overline{\psi_n(\mathbf{y})}\, d\mathbf{x} d\mathbf{y} \leq 0,$$

and the result follows, for $\mathrm{CN}(\mathrm{C}_0^r(\mathbb{R}^d)) \subset \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$. □

As Propositions 3.3.1 and 3.3.2 suggest, $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ is a wide and interesting class of functions. In particular, functions in $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ need not be bounded at the origin, which will be demonstrated in later examples (see Section 3.5), they need only be locally integrable. This fact is proved in the following lemma.

**Lemma 3.3.1** *If $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$, then $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$.*

*Proof.* Let $\mathrm{K} \subset \mathbb{R}^d$ be any compact set and $\mathrm{I} = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_d, b_d] \subset \mathbb{R}^d$ be such that $\mathrm{K} \subset \mathrm{I}$. Let $c = \max\{|a_1|, |b_1|, |a_2|, \ldots, |b_d|\} > 0$.

Let $\psi \in \mathrm{L}_0^2(\mathbb{R})$ be such that $\psi$ is positive and continuous on $[-2c, 2c]$, and $\int_{\mathbb{R}} \psi(x)\, dx = 0$. For any $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$, define $\Psi(\mathbf{x}) :=$

$\psi(x_1)\psi(x_2)\ldots\psi(x_d)$. Then, $\Psi \in \mathrm{L}_0^2(\mathbb{R}^d)$, $\int_{\mathbb{R}^d} \Psi(\mathbf{x})\, d\mathbf{x} = 0$ and, using the same steps as in the proof of Proposition 2.4.7, it can be shown that

$$\int_{\mathrm{K}} |f(\mathbf{z})|\, d\mathbf{z} \le \int_{\mathrm{I}} |f(\mathbf{z})|\, d\mathbf{z} \le \int_{[-c,\, c]^d} |f(\mathbf{z})|\, d\mathbf{z} < \infty.$$

□

*Remark.* By Proposition 3.3.2, we can replace $\mathrm{L}_0^2(\mathbb{R}^d)$ in Lemma 3.3.1 with $\mathrm{L}_0^p(\mathbb{R}^d)$ for any $p \in (2, \infty]$, or $\mathrm{C}_0^r(\mathbb{R}^d)$ for any $r \in [0, \infty]$. We can also replace $\mathrm{L}_0^2(\mathbb{R}^d)$ with a more general space $J$ of functions defined on $\mathbb{R}^d$, provided that for any $c > 0$, $J$ contains a function $h$, which is positive almost everywhere on $[-c,\, c]^d$ and $\int_{\mathbb{R}^d} h(\mathbf{x})\, d\mathbf{x} = 0$.

The following result is analogous to Lemma 2.6.2 and demonstrates, as a particular case, that $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ is a closed subset of $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ (recall, functions in $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ are necessarily locally integrable by Lemma 3.3.1).

**Lemma 3.3.2** *Let $p \in [1, 2]$ and $q = p/2(p-1)$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions such that $f_n \in \mathrm{L}^q_{\mathrm{loc}}(\mathbb{R}^d)$ and $f_n \in \mathrm{CN}(\mathrm{L}_0^p(\mathbb{R}^d))$ $(n \in \mathbb{N})$. If $f_n \xrightarrow{\mathrm{L}^q_{\mathrm{loc}}} f$ as $n \to \infty$, for some $f \in \mathrm{L}^q_{\mathrm{loc}}(\mathbb{R}^d)$, then $f \in \mathrm{CN}(\mathrm{L}_0^p(\mathbb{R}^d))$.*

*Proof.* Let $\phi \in \mathrm{L}_0^p(\mathbb{R}^d)$, $K \subset \mathbb{R}^d$ denote the compact support of $\phi * \phi^*$, and $r = p/(2-p)$. By equations (19) and (20),

$$\left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (f_n(\mathbf{x}-\mathbf{y}) - f(\mathbf{x}-\mathbf{y}))\phi(\mathbf{x})\overline{\phi(\mathbf{y})}\, d\mathbf{x}d\mathbf{y} \right| \le \left( \int_K |f_n(\mathbf{z}) - f(\mathbf{z})|^q\, d\mathbf{z} \right)^{\frac{1}{q}} \|\phi * \phi^*\|_r \to 0$$

$(n \to \infty)$. Thus,

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\phi(\mathbf{x})\overline{\phi(\mathbf{y})}\, d\mathbf{x}d\mathbf{y} = \lim_{n\to\infty} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f_n(\mathbf{x}-\mathbf{y})\phi(\mathbf{x})\overline{\phi(\mathbf{y})}\, d\mathbf{x}d\mathbf{y} \ge 0.$$

□

In the next section, for reasons which will be discussed, we will mainly be interested in real-valued functions $f$. For any set of functions $J$, let $J_{\mathbb{R}} := \{\psi \in J \mid \psi \text{ is real-valued}\}$. Consider the following definition for real-valued, conditionally negative definite functions.

**Definition 3.3.2** A real-valued function $f : \mathbb{R}^d \to \mathbb{R}$ is called conditionally negative definite w.r.t. $J$ if $f$ is even a.e., that is $f(\mathbf{x}) = f(-\mathbf{x})$ f.a.a. $\mathbf{x} \in \mathbb{R}^d$, and for every function $\phi \in J_{\mathbb{R}}$ satisfying $\int_{\mathbb{R}^d} \phi(\mathbf{x})\, d\mathbf{x} = 0$, the integral in (36) exists (in the Lebesgue sense) and is non-positive.

Note that in general, $J$ is an arbitrary set of test functions containing both real and complex-valued elements.

Let $\mathrm{CN}(J_{\mathbb{R}})$ denote the class of all real-valued functions which are conditionally negative definite with respect to $J$ by Definition 3.3.2. The next proposition demonstrates the connection between real-valued functions which are conditionally negative definite as in Definitions 3.3.1 and 3.3.2.

**Proposition 3.3.3** *Let $f : \mathbb{R}^d \to \mathbb{R}$ and $\tilde{J}$ denote a vector space of complex-valued functions on $\mathbb{R}^d$, such that if $\phi \in \tilde{J}$, then $\overline{\phi} \in \tilde{J}$. Then, $f \in \mathrm{CN}(\tilde{J}_{\mathbb{R}})$ if $f \in \mathrm{CN}(\tilde{J})$. Moreover, if (36) exists for all $\phi \in \tilde{J}$ such that $\int_{\mathbb{R}^d} \phi(\mathbf{x})\, d\mathbf{x} = 0$, then $f \in \mathrm{CN}(\tilde{J})$ if $f \in \mathrm{CN}(\tilde{J}_{\mathbb{R}})$.*

*Proof.* The first statement is clear since $\tilde{J}_{\mathbb{R}} \subseteq \tilde{J}$. For the second statement, consider the following. Let $\psi \in \tilde{J}$ be such that $\int_{\mathbb{R}^d} \psi(\mathbf{x})\, d\mathbf{x} = 0$, and suppose $f \in \mathrm{CN}(\tilde{J}_{\mathbb{R}})$. As in the proof of Proposition 2.4.1, we write $\psi$ as

$$\psi = \mathrm{Re}\,(\psi) + i\,\mathrm{Im}\,(\psi)$$

where $\mathrm{Re}\,(\psi), \mathrm{Im}\,(\psi) \in \tilde{J}_{\mathbb{R}}$ and $\int_{\mathbb{R}^d} \mathrm{Re}\,(\psi)(\mathbf{x})\, d\mathbf{x} = \int_{\mathbb{R}^d} \mathrm{Im}\,(\psi)(\mathbf{x})\, d\mathbf{x} = 0$. Again, we define $a := \mathrm{Re}\,(\psi)$, $b := \mathrm{Im}\,(\psi)$ and

$$t\,[u,v] := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\, u(\mathbf{x})\, \overline{v(\mathbf{y})}\, d\mathbf{x}d\mathbf{y} \qquad (u, v \in \tilde{J}).$$

Then,

$$t\,[a,b] + t\,[b,a] = t\,[a+b, a+b] - t\,[a,a] - t\,[b,b]$$

and

$$-i\,(t\,[a,b] - t\,[b,a]) = t\,[\psi,\psi] - t\,[a,a] - t\,[b,b] \tag{37}$$

are finite, since $f \in \mathrm{CN}(\tilde{J}_{\mathbb{R}})$ and we know that $t\,[\psi,\psi]$ exists. Hence, both $t\,[a,b]$ and $t\,[b,a]$ exist, and $t\,[a,b] = t\,[b,a]$. Thus, it follows from (37) that

$$t\,[\psi,\psi] = t\,[a,a] + t\,[b,b] \leq 0.$$

□

*Remark.* The above proof is almost identical to that of Proposition 2.4.1. The difference between the two lies in proving the existence of the integral in (36) for the appropriate set of test functions. In Proposition 3.3.3 we simply stipulate that (36) exists for all $\phi \in \tilde{J}$ such that $\int_{\mathbb{R}^d} \phi(\mathbf{x})\, d\mathbf{x} = 0$, whereas in Proposition 2.4.1, the existence of the integral follows under the assumption that $f \in \mathrm{P}(\hat{J}_{\mathbb{R}})$, and since $\hat{J}$ is closed under the operation $|\cdot|$, see (15). A similar approach will not work in the case of Proposition 3.3.3, since for $\psi \in \hat{J}$ such that $\int_{\mathbb{R}^d} \psi(\mathbf{x})\, d\mathbf{x} = 0$, although we have $\tilde{\psi} = |\psi| \in \hat{J}_{\mathbb{R}}$, as in (15), it doesn't necessarily follow that $\int_{\mathbb{R}^d} \tilde{\psi}(\mathbf{x})\, d\mathbf{x} = 0$. Hence, assuming $f \in \mathrm{CN}(\hat{J}_{\mathbb{R}})$ does not guarantee the existence of the integral in (36) for $\psi \in \hat{J}$.

It follows from Proposition 3.3.3 that under certain conditions, real-valued functions are conditionally negative definite as in both Definitions 3.3.2 and 3.3.1. This is the case when $J$ is a vector space of complex-valued functions which is closed under complex conjugation and (36) exists for all $\phi \in J$ such that $\int_{\mathbb{R}^d} \phi(\mathbf{x})\, d\mathbf{x} = 0$. In most of our examples this will be the case, for if $f \in \mathrm{L}^q_{\mathrm{loc}}(\mathbb{R}^d)$, then (36) exists for all $\phi \in \mathrm{L}^p_0(\mathbb{R}^d)$ ($p \in [1, 2]$, $q = p/2(p-1)$), by Lemma 2.5.1. The following proposition proceeds from this observation.

**Proposition 3.3.4** *Let $f : \mathbb{R}^d \to \mathbb{R}$. Then, $f \in \mathrm{CN}(\mathrm{L}^2_0(\mathbb{R}^d))$ if and only if $f \in \mathrm{CN}(\mathrm{L}^2_0(\mathbb{R}^d)_{\mathbb{R}})$.*

*Proof.* One direction is clear since $\mathrm{L}^2_0(\mathbb{R}^d)_{\mathbb{R}} \subset \mathrm{L}^2_0(\mathbb{R}^d)$. For the reverse implication, consider the following. Let $f \in \mathrm{CN}(\mathrm{L}^2_0(\mathbb{R}^d)_{\mathbb{R}})$. Then, $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$, by Lemma 3.3.1 (note that in the proof of Lemma 3.3.1, both $\psi$ and $\Psi$ are real-valued), and thus (36) exists for all $\phi \in \mathrm{L}^2_0(\mathbb{R}^d)$, by Lemma 2.5.1. Hence, $f \in \mathrm{CN}(\mathrm{L}^2_0(\mathbb{R}^d))$. □

## 3.4 An extension of Schoenberg's theorem to conditionally negative definite functions with a singularity at zero

Theorem 3.1.1 establishes a relation between the function classes $\mathrm{P}_{\mathrm{C},d}$ and $\mathrm{CN}_{\mathrm{C},d}$. We generalise this classical result to the classes $\mathrm{P}(\mathrm{L}^2_0(\mathbb{R}^d))$ and $\mathrm{CN}(\mathrm{L}^2_0(\mathbb{R}^d))$, and in doing so, derive a Schoenberg-type theorem for real-valued functions which need not be bounded at the origin.

**Theorem 3.4.1** *Let $f : \mathbb{R}^d \to \mathbb{R}$. If there exist $t_0 > 0$ and $p > 1$ such that $e^{\,t|f|} \in \mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^d)$ for all $0 < t \le t_0$, then*

$$f \in \mathrm{CN}(\mathrm{L}^2_0(\mathbb{R}^d)) \iff e^{-tf} \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}^d)) \;\; (0 < t \le t_0). \tag{38}$$

The proof of Theorem 3.4.1 will be based on the following five lemmas. By Propositions 2.4.1 and 3.3.4, we need only consider real-valued test functions in $\mathrm{L}^2_0(\mathbb{R}^d)$.

**Lemma 3.4.1** *Let $f : \mathbb{R}^d \to \mathbb{R}$ and $\phi \in \mathrm{L}_0^2(\mathbb{R}^d)$ be such that $\int_{\mathbb{R}^d} \phi(\mathbf{x})\, d\mathbf{x} < \infty$. If $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$, then there exists a Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle_f)$ and a mapping $k : \mathbb{R}^d \to \mathcal{H}$, $\mathbf{z} \mapsto k_{\mathbf{z}}$ such that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$,*

$$\|k_{\mathbf{x}} - k_{\mathbf{y}}\|_f^2 = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{u}-\mathbf{v})\phi(\mathbf{x}-\mathbf{u})\phi(\mathbf{y}-\mathbf{v})\, d\mathbf{u}d\mathbf{v} - \mathrm{C} \tag{39}$$

*where*

$$\mathrm{C} = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{s}-\mathbf{t})\phi(\mathbf{s})\phi(\mathbf{t})\, d\mathbf{s}d\mathbf{t} \tag{40}$$

*is independent of $\mathbf{x}$ and $\mathbf{y}$.*

*Proof.* Suppose $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$. Let V be the subset of $\mathrm{L}_0^2(\mathbb{R}^d)$ defined by

$$\mathrm{V} = \left\{ h \in \mathrm{L}_0^2(\mathbb{R}^d) \,\middle|\, h(\mathbf{u}) = \sum_{i=1}^{m} a_i\, \phi(\mathbf{x}_i - \mathbf{u})\ (\mathbf{u} \in \mathbb{R}^d) \text{ for some } m \in \mathbb{N}, \right.$$
$$\left. \mathbf{x}_i \in \mathbb{R}^d,\ a_i \in \mathbb{R};\ \text{s.t.} \int_{\mathbb{R}^d} h(\mathbf{u})\, d\mathbf{u} = 0 \right\}.$$

V is a vector space. For $\Phi, \Psi \in \mathrm{V}$, define

$$\langle \Phi, \Psi \rangle_f := -\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\Phi(\mathbf{x})\Psi(\mathbf{y})\, d\mathbf{x}d\mathbf{y}.$$

Since $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$, the mapping

$$(\Phi, \Psi) \mapsto \langle \Phi, \Psi \rangle_f$$

is a bilinear, symmetric and non-negative form on V. Set

$$\mathrm{V}' = \left\{ h' \in \mathrm{V} \,\middle|\, \langle h', h' \rangle_f = 0 \right\}.$$

V′ is a subspace of V since for any $g', h' \in \mathrm{V}'$,

$$\langle g' + h', g' + h' \rangle_f = 2\,\langle g', h' \rangle_f \le 2\,\langle g', g' \rangle_f^{\frac{1}{2}}\,\langle h', h' \rangle_f^{\frac{1}{2}} = 0$$

by the Cauchy-Schwarz inequality. On the quotient space V/V′, define

$$\langle [g], [h] \rangle_f := \langle g, h \rangle_f, \tag{41}$$

where $[g], [h]$ denote the equivalence classes in V/V′. The inner product in (41) is well-defined since

$$\langle g + g', h + h' \rangle_f = \langle g, h \rangle_f \qquad (g, h \in \mathrm{V},\ g', h' \in \mathrm{V}').$$

To see this, note that

$$\langle g, h' \rangle_f \leq \langle g, g \rangle_f^{\frac{1}{2}} \langle h', h' \rangle_f^{\frac{1}{2}} = 0$$

and

$$-\langle g, h' \rangle_f = \langle g, -h' \rangle_f \leq \langle g, g \rangle_f^{\frac{1}{2}} \langle -h', -h' \rangle_f^{\frac{1}{2}} = 0,$$

by the Cauchy-Schwarz inequality. Thus, $\langle g, h' \rangle_f = 0$ and similarly, $\langle g', h \rangle_f = 0$. Let $(\mathcal{H}, \langle \cdot, \cdot \rangle_f)$ be the Hilbert space completion of $\mathrm{V}/\mathrm{V}'$; then, in particular, $\mathrm{V}/\mathrm{V}'$ is dense in $\mathcal{H}$.

Let $\tilde{\mathbf{x}} \in \mathbb{R}^d$. For any $\mathbf{z} \in \mathbb{R}^d$, set $k_{\mathbf{z}} := [\phi(\mathbf{z} - \cdot) - \phi(\tilde{\mathbf{x}} - \cdot)] \in \mathcal{H}$. Then, for any $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^d$,

$$\begin{aligned}\|k_{\mathbf{x}} - k_{\mathbf{y}}\|_f^2 &= \langle k_{\mathbf{x}} - k_{\mathbf{y}}, k_{\mathbf{x}} - k_{\mathbf{y}} \rangle_f \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{u} - \mathbf{v}) \phi(\mathbf{x} - \mathbf{u}) \phi(\mathbf{y} - \mathbf{v}) \, d\mathbf{u} d\mathbf{v} \\ &\quad - \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{u} - \mathbf{v}) \phi(\mathbf{x} - \mathbf{u}) \phi(\mathbf{x} - \mathbf{v}) \, d\mathbf{u} d\mathbf{v} \\ &\quad - \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{u} - \mathbf{v}) \phi(\mathbf{y} - \mathbf{u}) \phi(\mathbf{y} - \mathbf{v}) \, d\mathbf{u} d\mathbf{v},\end{aligned} \tag{42}$$

since

$$\begin{aligned}\langle k_{\mathbf{x}}, k_{\mathbf{x}} \rangle_f =& - \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{u} - \mathbf{v}) \phi(\mathbf{x} - \mathbf{u}) \phi(\mathbf{x} - \mathbf{v}) \, d\mathbf{u} d\mathbf{v} \\ &+ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{u} - \mathbf{v}) \phi(\mathbf{x} - \mathbf{u}) \phi(\tilde{\mathbf{x}} - \mathbf{v}) \, d\mathbf{u} d\mathbf{v} \\ &- \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{u} - \mathbf{v}) \phi(\tilde{\mathbf{x}} - \mathbf{u}) \phi(\tilde{\mathbf{x}} - \mathbf{v}) \, d\mathbf{u} d\mathbf{v},\end{aligned}$$

$$\begin{aligned}2\langle k_{\mathbf{x}}, k_{\mathbf{y}} \rangle_f =& - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{u} - \mathbf{v}) \phi(\mathbf{x} - \mathbf{u}) \phi(\mathbf{y} - \mathbf{v}) \, d\mathbf{u} d\mathbf{v} \\ &+ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{u} - \mathbf{v}) \phi(\tilde{\mathbf{x}} - \mathbf{u}) \phi(\mathbf{y} - \mathbf{v}) \, d\mathbf{u} d\mathbf{v} \\ &+ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{u} - \mathbf{v}) \phi(\mathbf{x} - \mathbf{u}) \phi(\tilde{\mathbf{x}} - \mathbf{v}) \, d\mathbf{u} d\mathbf{v} \\ &- \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{u} - \mathbf{v}) \phi(\tilde{\mathbf{x}} - \mathbf{u}) \phi(\tilde{\mathbf{x}} - \mathbf{v}) \, d\mathbf{u} d\mathbf{v}\end{aligned}$$

and

$$\begin{aligned}\langle k_{\mathbf{y}}, k_{\mathbf{y}}\rangle_f = &-\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{u}-\mathbf{v})\phi(\mathbf{y}-\mathbf{u})\phi(\mathbf{y}-\mathbf{v})\,d\mathbf{u}d\mathbf{v}\\ &+\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{u}-\mathbf{v})\phi(\tilde{\mathbf{x}}-\mathbf{u})\phi(\mathbf{y}-\mathbf{v})\,d\mathbf{u}d\mathbf{v}\\ &-\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{u}-\mathbf{v})\phi(\tilde{\mathbf{x}}-\mathbf{u})\phi(\tilde{\mathbf{x}}-\mathbf{v})\,d\mathbf{u}d\mathbf{v}.\end{aligned}$$

By a simple change of variables, each of the last two integrals in (42) is equal to $-\mathrm{C}/2$, where C is defined by (40). Thus, formula (39) follows. □

We will refer to $\mathcal{H}$ as the Hilbert space associated with $f$ and $\phi$, and to $k$ as the mapping similarly associated.

Lemma 3.4.1 can be considered as a generalised version of the *GNS construction*, which is a widely celebrated technique in the literature, see e.g. [2, Th. C.2.3]. In fact, (39) is a direct extension of [2, Th. C.2.3 (i)]. Note that in [2], it is assumed that conditionally negative definite functions vanish at the origin. The following result is analogous to [2, Lemma C.3.1] and [4, Chapter 3, Lemma 2.1].

**Lemma 3.4.2** *Let $f : \mathbb{R}^d \to \mathbb{R}$ be such that $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ and $\phi \in \mathrm{L}_0^2(\mathbb{R}^d)$ be such that $\int_{\mathbb{R}^d}\phi(\mathbf{x})\,d\mathbf{x} < \infty$. Let $\mathcal{H}$ and $k$ denote the associating Hilbert space and mapping respectively. Fix $\mathbf{x}_0 \in \mathbb{R}^d$. The kernel*

$$g(\mathbf{x},\mathbf{y}) = \|k_{\mathbf{x}} - k_{\mathbf{x}_0}\|_f^2 + \|k_{\mathbf{y}} - k_{\mathbf{x}_0}\|_f^2 - \|k_{\mathbf{x}} - k_{\mathbf{y}}\|_f^2 \qquad (\mathbf{x},\mathbf{y}\in\mathbb{R}^d) \tag{43}$$

*is classically positive definite as in Definition 2.1.3.*

*Proof.* $g$ is clearly symmetric, that is $g(\mathbf{x},\mathbf{y}) = g(\mathbf{y},\mathbf{x})$ for all $\mathbf{x}$, $\mathbf{y}\in\mathbb{R}^d$. As in [2, p. 373], a straightforward calculation gives

$$g(\mathbf{x},\mathbf{y}) = 2\langle k_{\mathbf{x}} - k_{\mathbf{x}_0}, k_{\mathbf{y}} - k_{\mathbf{x}_0}\rangle_f \qquad (\mathbf{x},\mathbf{y}\in\mathbb{R}^d).$$

Therefore, for any $n\in\mathbb{N}$, $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n\in\mathbb{R}^d$ and $v_1,v_2,\ldots,v_n\in\mathbb{R}$,

$$\sum_{i,j=1}^{n} g(\mathbf{x}_i,\mathbf{x}_j)\,v_i\,v_j = 2\left\|\sum_{i=1}^{n} v_i(k_{\mathbf{x}_i} - k_{\mathbf{x}_0})\right\|_f^2 \geq 0.$$

□

For a real-valued function $f\in\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ and $\phi\in\mathrm{L}_0^2(\mathbb{R}^d)$ such that $\int_{\mathbb{R}^d}\phi(\mathbf{x})\,d\mathbf{x}<\infty$, with associating Hilbert space $(\mathcal{H},\langle\cdot,\cdot\rangle_f)$ and mapping $k$, define

$$\mathrm{K}(\mathbf{x},\mathbf{y}) := \|k_{\mathbf{x}} - k_{\mathbf{y}}\|_f^2 \qquad (\mathbf{x},\mathbf{y}\in\mathbb{R}^d).$$

It follows from (39) that $\mathrm{K}(\mathbf{x}+\mathbf{a}, \mathbf{y}+\mathbf{a}) = \mathrm{K}(\mathbf{x}, \mathbf{y})$ for any $\mathbf{a} \in \mathbb{R}^d$, and hence, $\mathrm{K}(\mathbf{x}, \mathbf{y}) = \mathrm{K}(\mathbf{x}-\mathbf{y}, \mathbf{0})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. Define $\tilde{f}(\mathbf{z}) := \mathrm{K}(\mathbf{z}, \mathbf{0})$ for all $\mathbf{z} \in \mathbb{R}^d$. Then, $\mathrm{K}(\mathbf{x}, \mathbf{y}) = \tilde{f}(\mathbf{x}-\mathbf{y})$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$.

The kernel $g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, defined in Lemma 3.4.2, is positive definite. Hence, so is $t^n g^n$ for any $t > 0$, $n \in \mathbb{N}$, since the product of positive definite kernels is also positive definite, see e.g. [2, Prop. C.1.6 (iv)]. Consequently, $e^{tg}$ is a classically positive definite kernel as in Definition 2.1.3.

Let $t > 0$. It follows from (43) that

$$e^{-t\mathrm{K}(\mathbf{x},\mathbf{y})} = e^{tg(\mathbf{x},\mathbf{y})} \times \left( e^{-t\mathrm{K}(\mathbf{x},\mathbf{x}_0)}\, e^{-t\mathrm{K}(\mathbf{y},\mathbf{x}_0)} \right) \qquad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^d).$$

The kernel $e^{-t\mathrm{K}(\,\cdot\,,\mathbf{x}_0)}\, e^{-t\mathrm{K}(\,\cdot\,,\mathbf{x}_0)} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is positive definite, since

$$\sum_{i,j=1}^{n} \left( e^{-t\mathrm{K}(\mathbf{x}_i,\mathbf{x}_0)}\, e^{-t\mathrm{K}(\mathbf{x}_j,\mathbf{x}_0)} \right) v_i\, v_j = \left( \sum_{i=1}^{n} v_i\, e^{-t\mathrm{K}(\mathbf{x}_i,\mathbf{x}_0)} \right)^2 \geq 0$$

for any $n \in \mathbb{N}$, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ and $v_1, v_2, \ldots, v_n \in \mathbb{R}$; as in [2, p. 374]. Hence, $e^{-t\mathrm{K}}$ is a positive definite kernel and therefore $e^{-t\tilde{f}}$ is a classically positive definite function as in Definition 2.1.2. By Theorem 3.1.1, it follows that $\tilde{f}$ is conditionally negative definite as in Definition 3.1.2. Thus, by the properties of functions in $\mathrm{CN}_{\mathrm{C},d}$ (see Section 3.1), it proceeds that $\tilde{f} + \alpha$ is conditionally negative definite for any $\alpha \in \mathbb{R}$.

The next lemma highlights a connection between classically conditionally negative definite functions and functions which are conditionally negative definite with respect to $\mathrm{L}_0^2(\mathbb{R}^d)$. In particular, we observe that for real-valued functions, $\mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^d) \cap \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ is the closure of $\mathrm{C}^\infty(\mathbb{R}^d) \cap \mathrm{CN}_{\mathrm{C},d}$, where $\mathrm{C}^\infty(\mathbb{R}^d)$ denotes the space of infinitely differentiable functions on $\mathbb{R}^d$.

**Lemma 3.4.3** *Let $p \in [1, \infty)$ and $f : \mathbb{R}^d \to \mathbb{R}$ be such that $f \in \mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^d) \cap \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$. Then, there is a sequence $(f_n)_{n\in\mathbb{N}}$ of infinitely differentiable, classically conditionally negative definite functions such that $f_n \xrightarrow{\mathrm{L}^p_{\mathrm{loc}}} f$ as $n \to \infty$.*

*Proof.* Recall the functions $\Psi_n$ defined in the proof of Proposition 2.4.6; i.e. let $\psi : \mathbb{R} \to \mathbb{R}$ denote the bump function

$$\psi(x) = \begin{cases} c_0 \exp\left(\frac{1}{|x|^2-1}\right), & |x| < 1 \\ 0, & |x| \geq 1, \end{cases}$$

where $c_0 > 0$ is the constant chosen such that $\int_{\mathbb{R}} \psi(x)\, dx = 1$, and for any $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$, define $\Psi(\mathbf{x}) := \psi(x_1)\psi(x_2)\ldots\psi(x_d)$ and

$$\Psi_n(\mathbf{x}) := n^d\, \Psi(n\, \mathbf{x}) \qquad (n \in \mathbb{N}). \tag{44}$$

Then, for any $n \in \mathbb{N}$, $\Psi_n \in \mathrm{C}_0^\infty(\mathbb{R}^d)$ is even and has compact support $[-\frac{1}{n}, \frac{1}{n}]^d$. Moreover, $\int_{\mathbb{R}^d} \Psi_n(\mathbf{x})\, d\mathbf{x} = 1$ for all $n \in \mathbb{N}$.

Applying Lemma 3.4.1 to the functions $\Psi_n$ and $f$, we find that for any $n \in \mathbb{N}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$,

$$\|k_{n,\mathbf{x}} - k_{n,\mathbf{y}}\|_f^2 = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{u} - \mathbf{v}) \Psi_n(\mathbf{x} - \mathbf{u}) \Psi_n(\mathbf{y} - \mathbf{v})\, d\mathbf{u} d\mathbf{v} - \mathrm{C}_n, \tag{45}$$

where

$$\mathrm{C_n} = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{s} - \mathbf{t}) \Psi_n(\mathbf{s}) \Psi_n(\mathbf{t})\, d\mathbf{s} d\mathbf{t} \in \mathbb{R}$$

and $k_{n,\mathbf{z}}$ is the equivalence class $[\Psi_n(\mathbf{z} - \cdot) - \Psi_n(\tilde{\mathbf{x}} - \cdot)]$ in $\mathcal{H}$ ($\mathbf{z}$, $\tilde{\mathbf{x}} \in \mathbb{R}^d$). Let

$$\tilde{f}_n(\mathbf{x} - \mathbf{y}) = \|k_{n,\mathbf{x}} - k_{n,\mathbf{y}}\|_f^2 \qquad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^d).$$

From Lemma 3.4.2 and the ensuing remarks, it follows that $\tilde{f}_n + \alpha \in \mathrm{CN}_{\mathrm{C},d}$ for any $\alpha \in \mathbb{R}$, $n \in \mathbb{N}$. In particular,

$$\begin{aligned} f_n(\mathbf{z}) &:= \tilde{f}_n(\mathbf{z}) + \mathrm{C_n} \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{u} - \mathbf{v}) \Psi_n(\mathbf{z} - (\mathbf{u} - \mathbf{y})) \Psi_n(\mathbf{z} - (\mathbf{x} - \mathbf{v}))\, d\mathbf{u} d\mathbf{v} \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{z} - \mathbf{t} - \mathbf{s}) \Psi_n(\mathbf{z} - \mathbf{s}) \Psi_n(\mathbf{z} - \mathbf{t})\, d\mathbf{s} d\mathbf{t} \quad (\mathbf{z} \in \mathbb{R}^d) \end{aligned} \tag{46}$$

defines a classically conditionally negative definite function. Note that we have used the evenness of $f$ and $\Psi_n$ in order to arrive at the above equation. On rewriting (46), again by using the fact that $f$ is even, we obtain

$$\begin{aligned} f_n(\mathbf{z}) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(\mathbf{t} - (\mathbf{z} - \mathbf{s})) \Psi_n(\mathbf{z} - \mathbf{s}) \Psi_n(\mathbf{z} - \mathbf{t})\, d\mathbf{s} d\mathbf{t} \\ &= \int_{\mathbb{R}^d} (f * \Psi_n(\mathbf{t}))\, \Psi_n(\mathbf{z} - \mathbf{t})\, d\mathbf{t}. \end{aligned}$$

Hence, for any $n \in \mathbb{N}$,

$$f_n = f * \eta_n, \tag{47}$$

where $\eta_n = \Psi_n * \Psi_n = n^d (\Psi * \Psi)(n\,\cdot)$. Using the properties of $\Psi_n$, it follows that for any $n \in \mathbb{N}$, $\eta_n \in \mathrm{C}_0^\infty(\mathbb{R}^d)$ has compact support $[-\frac{2}{n}, \frac{2}{n}]^d$ and

$\int_{\mathbb{R}^d} \eta_n(\mathbf{x})\, d\mathbf{x} = 1$. Hence, $f_n \in \mathrm{C}^\infty(\mathbb{R}^d)$ for all $n \in \mathbb{N}$, and $f_n \xrightarrow{\mathrm{L}^p_{\mathrm{loc}}} f$ as $n \to \infty$ by [42, Th. 1.18]. $\square$

Corollary 3.5.2 demonstrates that the converse to Lemma 3.4.3 is also true (see Section 3.5). In the following result we establish one direction of the equivalence (38) in Theorem 3.4.1.

**Lemma 3.4.4** *Let $f : \mathbb{R}^d \to \mathbb{R}$ be such that $f \in \mathrm{CN}(\mathrm{L}^2_0(\mathbb{R}^d))$. If there exist $t_0 > 0$ and $p > 1$ such that $e^{t|f|} \in \mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^d)$ for any $0 < t \le t_0$, then $e^{-tf} \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}^d))$ for all $0 < t \le t_0$.*

*Proof.* Suppose $t_0 > 0$ and $p > 1$ are such that $e^{t|f|} \in \mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^d)$ for all $0 < t \le t_0$. Then, it follows that $e^{-tf} \in \mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^d)$ for all $0 < t \le t_0$ and $f \in \mathrm{L}^q_{\mathrm{loc}}(\mathbb{R}^d)$ for any $1 \le q < \infty$.

The functions $f_n$, as defined in the proof of Lemma 3.4.3, are conditionally negative definite in the sense of Definition 3.1.2. Thus, by Theorem 3.1.1, $e^{-tf_n}$ is positive definite for any $t > 0$, $n \in \mathbb{N}$. Moreover, $e^{-tf_n}$ is continuous for all $t > 0$, $n \in \mathbb{N}$, since $f_n$ is continuous for any $n \in \mathbb{N}$. By Proposition 2.4.2, it follows that $e^{-tf_n} \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}^d))$ for any $t > 0$, $n \in \mathbb{N}$.

By Lemma 2.6.2, we need only show that $e^{-tf_n} \xrightarrow{\mathrm{L}^1_{\mathrm{loc}}} e^{-tf}$ as $n \to \infty$. Let $\epsilon > 0$, $0 < t \le t_0$ and $K \subset \mathbb{R}^d$ be a compact set. Let $n_0 \in \mathbb{N}$ and define $\hat{K} := K + \left[-\frac{2}{n_0}, \frac{2}{n_0}\right]^d$. By Lemma 3.4.3, there exists $n_* \in \mathbb{N}$ such that for all $n \ge n_*$,

$$\|f_n - f\|_{q,K} = \left(\int_K |f_n(\mathbf{x}) - f(\mathbf{x})|^q\, d\mathbf{x}\right)^{\frac{1}{q}} < \frac{\epsilon}{2t\, \|e^{t|f|}\|_{p,\hat{K}}} \tag{48}$$

where $\frac{1}{p} + \frac{1}{q} = 1$. W.l.o.g., we assume $n_* \ge n_0$.

Let $n > n_*$. We partition $K = K_1 \cup K_2$, where $K_1 = \{\mathbf{x} \in K \,|\, f(\mathbf{x}) \le f_n(\mathbf{x})\}$ and $K_2 = \{\mathbf{x} \in K \,|\, f(\mathbf{x}) > f_n(\mathbf{x})\}$. By the Mean Value Theorem,

$$\left|\left(e^{-tf_n} - e^{-tf}\right)(\mathbf{x})\right| = t\, e^{-t\,\xi_n(\mathbf{x})}\, |f_n(\mathbf{x}) - f(\mathbf{x})| \qquad \text{f. a. a. } \mathbf{x} \in \mathbb{R}^d,$$

where $\xi_n(\mathbf{x})$ lies between $f_n(\mathbf{x})$ and $f(\mathbf{x})$. Therefore,

$$\begin{aligned}
\int_K \left|\left(e^{-tf_n} - e^{-tf}\right)(\mathbf{x})\right| dx &= \int_{K_1} t e^{-t\xi_n(\mathbf{x})} |(f_n - f)(\mathbf{x})|\, d\mathbf{x} + \int_{K_2} t e^{-t\xi_n(\mathbf{x})} |(f_n - f)(\mathbf{x})|\, d\mathbf{x} \\
&\le \int_{K_1} t e^{-tf(\mathbf{x})} |(f_n - f)(\mathbf{x})|\, d\mathbf{x} + \int_{K_2} t e^{-tf_n(\mathbf{x})} |(f_n - f)(\mathbf{x})|\, d\mathbf{x} \\
&\le \int_{K} t e^{-tf(\mathbf{x})} |(f_n - f)(\mathbf{x})|\, d\mathbf{x} + \int_{K} t e^{-tf_n(\mathbf{x})} |(f_n - f)(\mathbf{x})|\, d\mathbf{x} \\
&\le t\left(\|e^{-tf}\|_{p,K} + \|e^{-tf_n}\|_{p,K}\right) \|f_n - f\|_{q,K},
\end{aligned} \tag{49}$$

using Hölder's inequality in the last step. Next, for almost all $\mathbf{x} \in \mathbb{R}^d$,

$$0 \leq e^{-tf(\mathbf{x})} = \sum_{j=0}^{\infty} \frac{(-1)^j \, t^j \, f(\mathbf{x})^j}{j!} \leq e^{t\,|f(\mathbf{x})|},$$

and thus, since $K \subset \hat{K}$,

$$\|e^{-tf}\|_{p,K} \leq \|e^{t\,|f|}\|_{p,K} \leq \|e^{t\,|f|}\|_{p,\hat{K}}. \tag{50}$$

By Jensen's inequality, see e.g. [25, Th. 1.8.1], and (47), it follows that for any $\mathbf{x} \in \mathbb{R}^d$,

$$e^{-tf_n(\mathbf{x})} = \exp\left(-t \int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{z})\eta_n(\mathbf{z})d\mathbf{z}\right) \leq \int_{\mathbb{R}^d} e^{-tf(\mathbf{x}-\mathbf{z})}\eta_n(\mathbf{z})d\mathbf{z} = e^{-tf} * \eta_n(\mathbf{x}),$$

and hence,

$$\|e^{-tf_n}\|_{p,K} \leq \|e^{-tf} * \eta_n\|_{p,K}.$$

Moreover,

$$\big(e^{-tf} * \eta_n\big)(\mathbf{x})\, \chi_K(\mathbf{x}) \leq \big((e^{-tf}\chi_{\hat{K}}) * \eta_n\big)(\mathbf{x}) \qquad (\mathbf{x} \in \mathbb{R}^d).$$

This follows since for $\mathbf{x} \in K$ and $n \geq n_0$,

$$\begin{aligned}
\big((e^{-tf}\chi_{\hat{K}}) * \eta_n\big)(\mathbf{x}) &= \int_{[-\frac{2}{n},\frac{2}{n}]^d} \Big(e^{-tf(\mathbf{x}-\mathbf{y})}\chi_{\hat{K}}(\mathbf{x}-\mathbf{y})\Big)\, \eta_n(\mathbf{y})\, d\mathbf{y} \\
&= \int_{[-\frac{2}{n},\frac{2}{n}]^d} e^{-tf(\mathbf{x}-\mathbf{y})}\eta_n(\mathbf{y})\, d\mathbf{y} = \big(e^{-tf} * \eta_n\big)(\mathbf{x}).
\end{aligned}$$

Thus, using Young's inequality, it proceeds that

$$\|e^{-tf} * \eta_n\|_{p,K} \leq \|e^{-tf}\|_{p,\hat{K}}\, \|\eta_n\|_1 = \|e^{-tf}\|_{p,\hat{K}} \leq \|e^{t\,|f|}\|_{p,\hat{K}},$$

and thus,

$$\|e^{-tf_n}\|_{p,K} \leq \|e^{t\,|f|}\|_{p,\hat{K}}. \tag{51}$$

From (48), (49), (50) and (51), we conclude that

$$\int_K \Big|\big(e^{-tf_n} - e^{-tf}\big)(\mathbf{x})\Big|\, d\mathbf{x} < \epsilon.$$

□

*Remark.* We take $p > 1$ in the above Lemma so that $q < \infty$ in (48) and we can apply Lemma 3.4.3. This condition allows us to compensate for functions with a singularity at zero.

By considering complex-valued functions in Lemmas 3.4.1, 3.4.2, 3.4.3 and 3.4.4, we obtain the following results. Equation (39) becomes

$$\|k_{\mathbf{x}} - k_{\mathbf{y}}\|_f^2 = \mathrm{Re}\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{u}-\mathbf{v})\phi(\mathbf{x}-\mathbf{u})\overline{\phi(\mathbf{y}-\mathbf{v})}\, d\mathbf{u}d\mathbf{v}\right) - \mathrm{C}, \tag{52}$$

where

$$\mathrm{C} = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{s}-\mathbf{t})\phi(\mathbf{s})\overline{\phi(\mathbf{t})}\, d\mathbf{s}d\mathbf{t} = \overline{\mathrm{C}} \in \mathbb{R}. \tag{53}$$

To see this, consider the same steps as in the proof of Lemma 3.4.1, with complex-valued $f$ and $\phi$. Take V as before and for $\Phi, \Psi \in \mathrm{V}$, define

$$\langle \Phi, \Psi\rangle_f := -\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\Phi(\mathbf{x})\overline{\Psi(\mathbf{y})}\, d\mathbf{x}d\mathbf{y},$$

so that the mapping $(\Phi, \Psi) \mapsto \langle \Phi, \Psi\rangle_f$ is a sesquilinear, Hermitian and non-negative form on V. Recall that $f$ is conjugate symmetric a.e., since $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$. To arrive at (52) and (53), compute $\|k_{\mathbf{x}} - k_{\mathbf{y}}\|_f^2 = \langle k_{\mathbf{x}}, k_{\mathbf{x}}\rangle_f - \langle k_{\mathbf{x}}, k_{\mathbf{y}}\rangle_f - \langle k_{\mathbf{y}}, k_{\mathbf{x}}\rangle_f + \langle k_{\mathbf{y}}, k_{\mathbf{y}}\rangle_f$.

Lemma 3.4.2, as well as the discussion that follows it, holds true for complex-valued $f$ and $\phi$.

The complex analogue of Lemma 3.4.3 gives rise to a sequence $(f_n)_{n\in\mathbb{N}}$ of infinitely differentiable, classically conditionally negative definite functions, such that $f_n \xrightarrow{\mathrm{L}^p_{\mathrm{loc}}} \mathrm{Re}(f)$ as $n \to \infty$. This can be seen by using equations (52) and (53) in the proof of Lemma 3.4.3, so that (45) becomes

$$\|k_{n,\mathbf{x}} - k_{n,\mathbf{y}}\|_f^2 = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathrm{Re}\left(f(\mathbf{u}-\mathbf{v})\right)\Psi_n(\mathbf{x}-\mathbf{u})\Psi_n(\mathbf{y}-\mathbf{v})\, d\mathbf{u}d\mathbf{v} - \mathrm{C}_n,$$

for any $n \in \mathbb{N}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. As a direct consequence, Lemma 3.4.4, for complex-valued $f$, roughly states that if $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$, then $e^{-t\mathrm{Re}(f)} \in \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$; thus, yielding information about the real part of $f$ only. It is for this reason that we consider real-valued functions in Theorem 3.4.1.

The following lemma concludes the proof of Theorem 3.4.1.

**Lemma 3.4.5** *Let $f : \mathbb{R}^d \to \mathbb{R}$. If there exists $t_0 > 0$ such that $e^{t|f|} \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ for any $0 < t \le t_0$, then*

$$e^{-tf} \in \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))\ (0 < t \le t_0) \implies f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d)).$$

*Proof.* Suppose $e^{-tf} \in \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$ for all $0 < t \le t_0$. Then, for any $\phi \in \mathrm{L}_0^2(\mathbb{R}^d)$ and $0 < t \le t_0$,

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{-tf(\mathbf{x}-\mathbf{y})}\phi(\mathbf{x})\phi(\mathbf{y})\,d\mathbf{x}d\mathbf{y} \ge 0$$

and hence,

$$-\frac{1}{t}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{-tf(\mathbf{x}-\mathbf{y})}\phi(\mathbf{x})\phi(\mathbf{y})\,d\mathbf{x}d\mathbf{y} \le 0.$$

Let $\psi \in \mathrm{L}_0^2(\mathbb{R}^d)$ be such that $\int_{\mathbb{R}^d}\psi(\mathbf{x})\,d\mathbf{x} = 0$. Then, for any $0 < t \le t_0$,

$$\begin{aligned}
-\frac{1}{t}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{-tf(\mathbf{x}-\mathbf{y})}\psi(\mathbf{x})\psi(\mathbf{y})d\mathbf{x}d\mathbf{y} &= \frac{1}{t}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\psi(\mathbf{x})\psi(\mathbf{y})\,d\mathbf{x}d\mathbf{y} \\
&\quad - \frac{1}{t}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{-tf(\mathbf{x}-\mathbf{y})}\psi(\mathbf{x})\psi(\mathbf{y})\,d\mathbf{x}d\mathbf{y} \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left(\frac{1-e^{-tf(\mathbf{x}-\mathbf{y})}}{t}\right)\psi(\mathbf{x})\psi(\mathbf{y})d\mathbf{x}d\mathbf{y}
\end{aligned}$$

Define $f_t := (1-e^{-tf})/t$ $(0 < t \le t_0)$. It follows that $f_t \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ for all $0 < t \le t_0$. To prove $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ we need only show that $f_t \xrightarrow{\mathrm{L}^1_{\mathrm{loc}}} f$ as $t \to 0$, by Lemma 3.3.2. Let $K \subset \mathbb{R}^d$ be a compact set. Then,

$$\begin{aligned}
\int_K \left| f(\mathbf{x}) - \left(\frac{1-e^{-tf(\mathbf{x})}}{t}\right)\right| d\mathbf{x} &= \int_K \left| t\sum_{j=2}^{\infty}\frac{(-1)^j\,t^{j-2}\,f(\mathbf{x})^j}{j!}\right| d\mathbf{x} \\
&\le \int_K t\sum_{j=2}^{\infty}\frac{t_0^{j-2}\,|f(\mathbf{x})|^j}{j!}\,d\mathbf{x} \\
&= \frac{t}{t_0^2}\int_K \sum_{j=2}^{\infty}\frac{t_0^j\,|f(\mathbf{x})|^j}{j!}\,d\mathbf{x} \\
&= \frac{t}{t_0^2}\int_K\left(e^{t_0|f(\mathbf{x})|} - 1 - t_0|f(\mathbf{x})|\right)d\mathbf{x} \\
&\le \frac{t}{t_0^2}\,\|e^{t_0|f|}\|_{1,K} \to 0 \qquad (t \to 0).
\end{aligned}$$

□

*Remark.* Lemma 3.4.5 is also valid for complex-valued functions $f$. Note, however, we cannot prove analogous results to Lemmas 3.4.4 and 3.4.5 for the function spaces $\mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$ and $\mathrm{CN}(\mathrm{L}^2(\mathbb{R}^d))$, as opposed to $\mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$

and $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$, for we cannot have both $f$ and $e^{-tf}$ in $\mathrm{L}^1(\mathbb{R}^d)$. In fact, it is clear that if $f \in \mathrm{L}^1(\mathbb{R}^d)$, then $e^{-tf} \notin \mathrm{L}^1(\mathbb{R}^d)$. Moreover, $e^{t|f|} \notin \mathrm{L}^p(\mathbb{R}^d)$ ($p \in [1, \infty)$) for $f \in \mathrm{L}^1(\mathbb{R}^d)$.

## 3.5 Corollaries to Theorem 3.4.1 and examples of conditionally negative definite functions with a singularity at zero

We begin this section by proving two results which follow directly from Theorem 3.4.1. Firstly, we demonstrate that functions which are continuous and conditionally negative definite for $\mathrm{L}_0^2(\mathbb{R}^d)$ are conditionally negative definite in the classical sense.

**Corollary 3.5.1** *Let $f : \mathbb{R}^d \to \mathbb{R}$ be continuous. Then, $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ if and only if $f \in \mathrm{CN}_{\mathrm{C},d}$.*

*Proof.* Suppose $f$ is continuous and classically conditionally negative definite. By Theorem 3.1.1, $e^{-tf}$ is classically positive definite for all $t > 0$. Moreover, for any $t > 0$, $e^{-tf}$ is continuous since $f$ is continuous. By Proposition 2.4.2, $e^{-tf} \in \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$ for all $t > 0$. By Theorem 3.4.1, we conclude that $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$.

For the reverse implication, consider the same argument as in the proof of Proposition 2.4.6, with

$$\Phi_n(\mathbf{x}) := \sum_{i=1}^{N} \xi_i \, \Psi_n(\mathbf{x} - \mathbf{x}_i) \quad (\mathbf{x} \in \mathbb{R}^d, n \in \mathbb{N}),$$

for any $N \in \mathbb{N}$ and any $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N \in \mathbb{R}^d$, where $\xi_1, \xi_2, \ldots, \xi_N \in \mathbb{R}$ are such that $\sum_{i=1}^N \xi_i = 0$. Then, $\Phi_n \in \mathrm{L}_0^2(\mathbb{R}^d)$ and $\int_{\mathbb{R}^d} \Phi_n(\mathbf{x}) \, d\mathbf{x} = 0$ for all $n \in \mathbb{N}$. $\square$

*Remark.* It follows directly from Corollary 3.5.1 that the functions $f$ defined in Theorem 3.2.3 are conditionally negative definite with respect to $\mathrm{L}_0^2(\mathbb{R}^d)$.

The following corollary shows that the converse to Lemma 3.4.3 is true. Indeed, if there exists a sequence $(f_n)_{n\in\mathbb{N}}$ of classically conditionally negative definite functions such that $f_n \xrightarrow{\mathrm{L}^p_{\mathrm{loc}}} f$ as $n \to \infty$, then $f$ is conditionally negative definite for $\mathrm{L}_0^2(\mathbb{R}^d)$.

**Corollary 3.5.2** *Let $f : \mathbb{R}^d \to \mathbb{R}$ be such that $f \in \mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^d)$ for some $p \in [1, \infty)$. Then, $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ if and only if there exists a sequence $(f_n)_{n\in\mathbb{N}}$ of infinitely differentiable, classically conditionally negative definite functions, such that $f_n \xrightarrow{\mathrm{L}^p_{\mathrm{loc}}} f$ as $n \to \infty$.*

*Proof.* One direction is proved in Lemma 3.4.3. For the reverse implication, we note that for any $n \in \mathbb{N}$, $f_n \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$ by Corollary 3.5.1. By Lemma 3.3.2, we need only show that $f_n \xrightarrow{\mathrm{L}^1_{\mathrm{loc}}} f$ as $n \to \infty$. This follows directly, since for any compact set $K \subset \mathbb{R}^d$,

$$\int_K |f_n(\mathbf{x}) - f(\mathbf{x})|\, d\mathbf{x} \leq |K|^{\frac{1}{q}} \left( \int_K |f_n(\mathbf{x}) - f(\mathbf{x})|^p \, d\mathbf{x} \right)^{\frac{1}{p}}$$

where $\frac{1}{p} + \frac{1}{q} = 1$ and $|K|$ denotes the Lebesgue measure of $K$. $\square$

We will now demonstrate how to construct examples of functions in $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$, which have a singularity at the origin. Firstly, assume $d = 1$. It is clear that if $f \in \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}))$, then $-f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}))$. Hence, the functions $-g_i$, for $i = 1, 8, 11, 16, 19, 23$, of Section 2.7 are elements of $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}))$. Next, we show how to find examples of singular functions $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}))$ such that $-f \notin \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}))$.

In view of Theorem 3.2.1 ii. with $g(x) = x^{-t}$ ($t > 0$), if $h$ is a Bernstein function, then $u_t = h^{-t}$ is completely monotone for all $t > 0$. Assume $h$ is non-constant, define $v_t := h^{-t}(|\cdot|)$ and let $t_0 > 0$. By Corollary 2.7.6, if $v_t \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$ for all $0 < t \leq t_0$, then $v_t \in \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}))$ for all $0 < t \leq t_0$. Define $f := \log(h(|\cdot|))$, so that $e^{-tf} = v_t$. By Theorem 3.4.1, if $e^{t|f|} \in \mathrm{L}^p_{\mathrm{loc}}(\mathbb{R})$ for any $0 < t \leq t_0$ and some p $> 1$, then $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}))$. It only remains to check the last condition, $-f \notin \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}))$.

Consider the following simple examples. In all three cases we choose p $= 2$ and $t_0 = 1/4$, so that $v_t \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$ and $e^{t|f|} \in \mathrm{L}^p_{\mathrm{loc}}(\mathbb{R})$ for all $0 < t \leq t_0$.

1. Take $h(x) = x$, then $u_t(x) = x^{-t}$, $v_t(x) = |x|^{-t}$ and $f(x) = \log|x|$. All conditions are satisfied, hence $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}))$. It is easy to see that $-f \notin \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}))$. Indeed, take the test function $\phi(x) = 1$ for $0 < x < 8$, $\phi(x) = 0$ otherwise. Then,

$$-\int_{\mathbb{R}}\int_{\mathbb{R}} f(x-y)\phi(x)\phi(y)\, dxdy = -\int_0^8\int_0^8 \log|x-y|\, dxdy = 96(1 - 2\log 2) < 0.$$

2. Take $h(x) = x + \sqrt{x}$, then $u_t(x) = (x + \sqrt{x})^{-t}$, $v_t(x) = (|x| + \sqrt{|x|})^{-t}$ and $f = \log(|\cdot| + \sqrt{|\cdot|}) \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}))$. To prove $-f \notin \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}))$, we take the same test function $\phi$ as in the example above, giving

$$-\int_0^8\int_0^8 \log\left(|x-y| + \sqrt{|x-y|}\right) dxdy \simeq -75.20631216 < 0.$$

3. Take $h(x) = \Gamma(x + \frac{1}{2})/\Gamma(x)$, where $\Gamma$ is the Gamma function. We have $f = \log\Gamma(|\cdot| + \frac{1}{2}) - \log\Gamma(|\cdot|) \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}))$. To prove $-f \notin \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}))$,

again we take the same test function $\phi$ as above, finding

$$-\int_0^8 \int_0^8 \big( \log \Gamma(|x-y| + 1/2) - \log \Gamma(|x-y|) \big) dx dy \simeq -10.83 < 0.$$

All three examples are built on the same principle: if $h$ is a Bernstein function and some regularity conditions are satisfied, then $f = \log h(|\cdot|) \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}))$.

We now indicate how to construct examples of functions in $\mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$, for $d \geq 2$, which are not in $\mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$. Similarly to the above, for any $\mathbf{x} \in \mathbb{R}^d$, define $v_t(\mathbf{x}) := h^{-t}(\|\mathbf{x}\|^2) e^{-t\|\mathbf{x}\|^2}$, and let $t_0 > 0$. By Corollary 2.7.7, if $h^{-t} \in \mathrm{L}^1_{\mathrm{loc}}((0, \infty))$ for all $0 < t \leq t_0$, then $v_t \in \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$ for all $0 < t \leq t_0$. Define

$$f(\mathbf{x}) := \log \big( h(\|\mathbf{x}\|^2) \exp(\|\mathbf{x}\|^2) \big) = \log \big( h(\|\mathbf{x}\|^2) \big) + \|\mathbf{x}\|^2 \qquad (\mathbf{x} \in \mathbb{R}^d),$$

so that $e^{-tf} = v_t$. By Theorem 3.4.1, if $e^{t|f|} \in \mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^d)$ for any $0 < t \leq t_0$ and some p $> 1$, then $f \in \mathrm{CN}(\mathrm{L}_0^2(\mathbb{R}^d))$. It only remains to check the last condition, $-f \notin \mathrm{P}(\mathrm{L}_0^2(\mathbb{R}^d))$.

# 4 Applications in Potential Theory

The development of the classical theory of potentials was motivated by the energy problem: given a function $f : \mathbb{R}^d \to \mathbb{R}$ and a set $\mathcal{M}(\mathcal{X})$ of non-negative unit Borel measures with support in the compact set $\mathcal{X} \subset \mathbb{R}^d$, which measure(s) $\mu \in \mathcal{M}(\mathcal{X})$ minimise(s)/maximise(s) the energy integral

$$I_f(\mu) := \int_{\mathcal{X}} \int_{\mathcal{X}} f(\mathbf{x} - \mathbf{y})\mu(d\mathbf{x})\mu(d\mathbf{y})? \tag{54}$$

Moreover, if $\mu^* \in \mathcal{M}(\mathcal{X})$ minimises $I_f$, then is it unique and thus, optimal?

In logarithmic potential theory, i.e. when $f = -\log|\cdot|$, the energy problem has been solved for various sets $\mathcal{X}$, see e.g. [33, Chapter I.1], where circles, discs and line segments are considered. For example, when $\mathcal{X} \subset \mathbb{R}$ is a segment of length $l$, that is $\mathcal{X} = [-l/2, l/2]$, the minimising measure $\mu^*$ is unique and has arcsine density;

$$\mu^*(dx) = \frac{1}{\pi\sqrt{l^2/4 - x^2}}\, dx \quad (x \in [-l/2, l/2]),$$

see [33, Eq. 1.7]. The more general case (for $d = 1$) when

$$f(x) = \begin{cases} (1 - |x|^{\alpha-1})/(\alpha - 1) & \text{if} \quad \alpha \neq 1, \\ -\log|x| & \text{if} \quad \alpha = 1 \end{cases} \qquad (x \in \mathbb{R}\backslash\{0\})$$

has been considered in [37]. Here it is shown that for any $\alpha \in (0, 2)$, $I_f(\mu)$ is strictly convex on the set of all probability measures on the set of Borel subsets of $[0, 1]$, and that the measure with generalised arcsine density

$$p_{1-\alpha/2}(t) = \frac{\Gamma\big(2 - \alpha\big)\, t^{-\alpha/2}\, (1 - t)^{-\alpha/2}}{\Gamma^2\big(1 - \alpha/2\big)}$$

is the optimal measure for $I_f$, see [37, Th. 2]. The energy problem has also been widely studied for the Riesz kernel $\kappa_\alpha(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|^{\alpha-d}$ ($0 < \alpha < d$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$) and the classical Newtonian kernel, i.e. when $f(\mathbf{x}) = |\mathbf{x}|^{2-d}$ ($d > 2$, $\mathbf{x} \in \mathbb{R}^d$), see e.g. [16], [21], [28], [32]. In the non-singular case, when $f = |\cdot|^\alpha$ ($\alpha > 0$), properties of the maximising measures and their potentials

$$\mathcal{P}_\mu(\mathbf{y}) := \int_{\mathcal{X}} f(\mathbf{x} - \mathbf{y})\, \mu(d\mathbf{x}) \quad (y \in \mathcal{X})$$

have been explored in [8]. The main purpose of this chapter is to study the energy integral $I_{f(|\cdot|)}$ for several singular, completely monotone functions $f$, which have not previously been considered in the field of potential theory, and to solve the corresponding energy problems by means of numerically computing densities of measures which minimise $I_{f(|\cdot|)}$.

It is worth noting that the energy problem in potential theory is very closely related to the problem of finding optimal designs in experimental design. In particular, the functional $I_f(\mu)$ arises as an optimality criterion in the optimal design problem with correlated observations for the location model $y_j = \theta + \epsilon_j$, see e.g. [47, Eq. 5]. The measure $\mu^*$ that minimises $I_f$ on the set of probability measures defined on a compact subset of $\mathbb{R}$, say $[0, 1]$, defines an optimal design for a suitable correlation function $f$. Standard correlation functions are positive definite in the classical sense, however, as in [37, Corollary 1], we extend the optimal design problem to the case when $f$ is positive definite in an extended sense and singular at the origin.

Consider the energy integral $I_{f(|\cdot|)}$ with $f \in \mathrm{CM}$ unbounded at the origin. Using the results of Section 4.4, where we construct discrete optimal measures for $I_g$ with $g$ classically strictly positive definite, and Theorem 4.1.1, which describes a method for approximating singular completely monotone functions by non-singular such functions, we derive our principle result; an algorithm for constructing continuous probability measures which approximate the minimising measure for $I_{f(|\cdot|)}$. We apply this algorithm to the case when $f(|\cdot|) = |\cdot|^{-\alpha}$ ($\alpha \in (0, 1)$) - the Riesz kernel on $\mathbb{R}$, or a compact subset of $\mathbb{R}$ (see Section 4.5), and, later, to a variety of singular, completely monotone functions $f$ (see Section 4.6).

The structure of this section is as follows. We begin by introducing a procedure for approximating a singular completely monotone function by a family of non-singular completely monotone functions, see Section 4.1. Section 4.3 contains some known ideas on finding the optimal measure for a convex, non-negative energy integral. In Section 4.4 we construct discrete optimal measures for $I_f$ in the case when $f$ is classically strictly positive definite. Next, we provide an in-depth analysis of the optimal density for the Riesz energy and, in doing so, construct a general algorithmic scheme for approximating an optimal density for a given energy by a sequence of probability densities, see Section 4.5. We conclude with several examples of approximate minimising measures for $I_{f(|\cdot|)}$ with alternative singular, completely monotone functions $f$, see Section 4.6.

## 4.1 Approximation of singular completely monotone functions by non-singular completely monotone functions

The central result of this section, Theorem 4.1.1, determines a method for approximating a singular completely monotone function $f$ by a family of bounded completely monotone functions $f_\epsilon$ ($\epsilon > 0$), and proves to be very useful when numerically computing densities of measures which minimise the energy $I_{f(|\cdot|)}$ defined in (54), see Sections 4.5 and 4.6 for details.

For $g \in \mathrm{BF}$ of the form (35), let $f = g' \in \mathrm{CM}$. Then, for any $x > 0$,

$$f(x) = b + \int_{(0\,\infty)} te^{-xt}\mu(dt) \tag{55}$$

and the measure $\nu(dt) := t\mu(dt)$ satisfies

$$\int_{(0\,\infty)} \frac{1}{1+t}\,\nu(dt) = \int_{(0\,\infty)} \frac{t}{1+t}\,\mu(dt) \le \int_{(0\,\infty)} \min(1,t)\,\mu(dt) < \infty, \tag{56}$$

see [35, p.18]. The following proposition demonstrates that the converse to the above discussion is also true. That is, a completely monotone function of the form (55) with representing measure satisfying (56) has a primitive in BF. This, in turn, characterises the image of BF under differentiation.

**Proposition 4.1.1** ([35, Prop. 3.4]). *Let $f \in \mathrm{CM}$ be of the form*

$$f(x) = b + \int_{(0\,\infty)} e^{-xt}\nu(dt) \quad (x > 0). \tag{57}$$

*Then, $f$ has primitive $g \in \mathrm{BF}$ if and only if the representing measure $\nu$ satisfies*

$$\int_{(0\,\infty)} (1+t)^{-1}\,\nu(dt) < \infty. \tag{58}$$

*Proof.* Retracing the steps in the above discussion, namely equations (55) and (56), reveals that functions of the form (57), with representing measure satisfying (58), have primitives which can be written in the form of (35). □

Let $f \in \mathrm{CM}$ with representing measure $\nu$ satisfying (58). Then, by Proposition 4.1.1, there exists $g \in \mathrm{BF}$ such that $g(x) = \int_0^x f(t)\,dt$ ($x > 0$). As discussed in Section 2.2, the value $f(0)$ may be undefined, that is, $f$ may be singular at the origin. For any $x > 0$, define the family of functions

$$f_\epsilon(x) = \begin{cases} f(x) = g'(x), & \epsilon = 0 \\ \frac{1}{\epsilon}\int_x^{x+\epsilon} f(t)\,dt = \frac{1}{\epsilon}(g(x+\epsilon) - g(x)), & \epsilon > 0. \end{cases} \tag{59}$$

Next, we show that the functions defined in (59) are completely monotone and that $f_\epsilon(0+) = \lim_{x\to 0+} f_\epsilon(x) < \infty$ for all $\epsilon > 0$ and $\lim_{\epsilon\to 0} f_\epsilon(x) = f(x)$ for all $x > 0$.

**Theorem 4.1.1** *Let $f \in \mathrm{CM}$ with representing measure $\nu$ satisfying* (58). *Consider the family of functions $f_\epsilon$ defined in* (59), *where $g \in \mathrm{BF}$ and $g(x) = \int_0^x f(t)\,dt$ $(x > 0)$. Then,*

*i.* $f_\epsilon(0+) = \lim_{x\to 0+} f_\epsilon(x) < \infty$ *for any* $\epsilon > 0$,

*ii.* $f_\epsilon \in \mathrm{CM}$ *for any* $\epsilon \geq 0$,

*iii.* $f_\delta - f_\epsilon \in \mathrm{CM}$ *for any* $0 \leq \delta < \epsilon$,

*iv.* $\lim_{\epsilon\to 0} f_\epsilon(x) = f(x)$ *for any* $x > 0$.

*Proof.* i. Let $\epsilon > 0$. Then, $f(0) = (g(\epsilon) - g(0))/\epsilon < \infty$ since $g$ is non-negative and increasing by definition.

ii. The case $\epsilon = 0$ is trivial. For $\epsilon > 0$, consider the following. Using the Lévy-Khintchine representation of $g$, see (35), it follows that for any $\epsilon$, $x > 0$,

$$f_\epsilon(x) = \frac{1}{\epsilon}(g(x+\epsilon) - g(x)) = b + \frac{1}{\epsilon}\int_{(0,\infty)} e^{-xt}(1 - e^{-\epsilon t})\mu(dt)$$

or, alternatively,

$$f_\epsilon(x) = \int_{(0,\infty)} e^{-xt}\nu_\epsilon(dt), \tag{60}$$

where $\nu_\epsilon(dt) = b\delta_0(dt) + h_\epsilon(t)\mu(dt)$, $\delta_0(dt)$ denotes the delta (Dirac) measure concentrated at 0 and $h_\epsilon(t) = \epsilon^{-1}(1 - e^{-\epsilon t})$ for any $t > 0$. Since $h_\epsilon(t) > 0$ for any $\epsilon$, $t > 0$ and $f_\epsilon(0+) < \infty$ for any $\epsilon > 0$ by i., it follows from Theorem 2.2.1 that $f_\epsilon \in \mathrm{CM}$ for any $\epsilon > 0$.

iii. Let $0 < \delta < \epsilon$. Using the representations of $f_\epsilon$ and $f_\delta$ described in (60), we have that for any $x > 0$,

$$f_\delta(x) - f_\epsilon(x) = \int_{(0,\infty)} e^{-xt}(\nu_\delta - \nu_\epsilon)(dt), \tag{61}$$

where $(\nu_\delta - \nu_\epsilon)(dt) = (h_\delta - h_\epsilon)(t)\mu(dt)$ for any $t > 0$. The measure $(\nu_\delta - \nu_\epsilon)(dt)$ is non-negative, since $(h_\delta - h_\epsilon)(t) > 0$ for any $t > 0$. Indeed, for fixed $t > 0$, the function $h_{\tilde\epsilon}$, considered as a function of $\tilde\epsilon > 0$, is positive, as shown in the proof of ii., and strictly decreasing. The latter property follows since

$$\frac{\partial h_{\tilde{\epsilon}}(t)}{\partial \tilde{\epsilon}} = \frac{(1+\tilde{\epsilon}t)e^{-\tilde{\epsilon}t}-1}{\tilde{\epsilon}^2}$$

and $(1+s)e^{-s}-1$ is negative for any $s>0$. Thus, since $f_\delta(0+)-f_\epsilon(0+) < \infty$ by i., it follows from Theorem 2.2.1 that $f_\delta - f_\epsilon \in \mathrm{CM}$.

Consider the case $\delta = 0$. It follows from (59) and (55) that for any $x>0$,

$$f_0(x) = g'(x) = b + \int_{(0,\infty)} te^{-xt}\mu(dt) = \int_{(0,\infty)} e^{-xt}\nu_0(dt),$$

where $\nu_0(dt) = b\delta_0(dt) + t\mu(dt)$. Setting $\delta = 0$ in (61) gives the representing measure $(\nu_0 - \nu_\epsilon)(dt) = (t - h_\epsilon(t))\mu(dt)$ $(t>0)$ for $f_0 - f_\epsilon$. Since $t - \epsilon^{-1}(1 - e^{-\epsilon t}) > 0$ for any $t$, $\epsilon > 0$, $f_\epsilon(0+) < \infty$ for any $\epsilon > 0$ and $f_0(x) = f(x) < \infty$ for any $x > 0$, it follows from Theorem 2.2.1 that $f_0 - f_\epsilon \in \mathrm{CM}$ for any $\epsilon > 0$.

iv. Let $x>0$. For any $\epsilon > 0$, $f(x+\epsilon) < f_\epsilon(x)$ since $f$ is decreasing on $(0,\infty)$, and $f_\epsilon(x) < (f(x+\epsilon)+f(x))/2$ since $f$ is convex on $(0,\infty)$. Moreover, there exists $x < \xi < x+\epsilon$ such that

$$f(x) - f(x+\epsilon) = \epsilon f'(\xi),$$

by the Mean Value Theorem. Hence, $f_\epsilon(x) - f(x+\epsilon) = \mathcal{O}(\epsilon)$ as $\epsilon \to 0$ and thus, $\lim_{\epsilon\to 0} f_\epsilon(x) = f(x)$. □

## 4.2 Convexity of the energy functional

Let $f : \mathbb{R}^d \to \mathbb{R}$ and $\mathcal{M}$ denote a general set of signed measures on $\mathbb{R}^d$, or a compact subset of $\mathbb{R}^d$. We define the *energy functional*

$$\Phi_f(\mu) := \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})\mu(d\mathbf{x})\mu(d\mathbf{y}) \quad (\mu \in \mathcal{M}). \tag{62}$$

As follows from the standard definition, $\Phi_f : \mathcal{M} \to \mathbb{R}$ is *convex* on $\mathcal{M}$ if

$$\Phi_f\left((1-\alpha)\,\mu_1 + \alpha\,\mu_2\right) \le (1-\alpha)\,\Phi_f(\mu_1) + \alpha\,\Phi_f(\mu_2) \tag{63}$$

for any $0 \le \alpha \le 1$ and $\mu_1, \mu_2 \in \mathcal{M}$. $\Phi_f$ is said to be *strictly convex* if for any $\mu_1 \ne \mu_2$, the inequality in (63) is replaced with a strict inequality and $0 < \alpha < 1$.

In the field of experimental design, the location model $y_j = \theta + \epsilon_j$ with correlated errors $\epsilon_j$ is often considered. The normalised variance of the least square estimator of $\theta$, i.e. the optimality criterion, is the energy $\Phi_\rho(\mu)$ defined in (62), where $\mu \in \mathcal{M}$ is a design and $\rho$ is the correlation function of the error process. $\mathcal{M}$ is commonly taken to be the set of probability measures on $[0,1]$, and since $\rho$ is classically positive definite by definition, it follows that $\Phi_\rho$ is convex on $\mathcal{M}$, see e.g. [47, Lemma 1]. Due to the bounded nature of $\rho$, $\Phi_\rho$ is finite on the set of all probability measures on $[0,1]$. This is clearly not the case when $\rho$ is replaced with a function $f$ which has a singularity at the origin, for $\Phi_f(\mu) = +\infty$ for any discrete measure $\mu$.

Using the framework developed in Sections 2 and 3, we now provide several examples of alternative functions $f$ (potentially unbounded at 0) which guarantee the convexity of $\Phi_f$ on certain sets of measures $\mathcal{M}$. Although these results will not be used in later sections, we present them for completeness. Firstly, we introduce the notion of conditional positive definiteness, which is simply the "negative" of conditional negative definiteness.

**Definition 4.2.1** A function $f : \mathbb{R}^d \to \mathbb{R}$ is called *conditionally positive definite w.r.t. a set* $J$ of functions, denoted $f \in \mathrm{CP}(J)$, if $-f \in \mathrm{CN}(J)$, as in Definition 3.3.2.

For any $c \in \mathbb{R}$, define $\mathrm{L}^2_{0,c}(\mathbb{R}^d)$ to be the set of functions $\phi \in \mathrm{L}^2_0(\mathbb{R}^d)$ such that $\int_{\mathbb{R}^d} \phi(\mathbf{x})\, d\mathbf{x} = c$. A consequence of the following proposition is that if $f \in \mathrm{CP}(\mathrm{L}^2_0(\mathbb{R}^d))$, then $\Phi_f$ is convex on the set of absolutely continuous signed measures with densities in $\mathrm{L}^2_{0,c}(\mathbb{R}^d)$. Note, in the proceeding results, $\mathcal{M}_i$ will always denote a set of signed measures.

**Proposition 4.2.1** *Let $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ be even a.e., $c \in \mathbb{R}$ and $\mathcal{M}_1 = \{\mu \mid \mu$ absolutely continuous with density $\phi \in \mathrm{L}^2_{0,c}(\mathbb{R}^d)\}$. Then, $\Phi_f : \mathcal{M}_1 \to \mathbb{R}$, as defined in* (62)*, is convex if and only if $f \in \mathrm{CP}(\mathrm{L}^2_0(\mathbb{R}^d))$.*

*Proof.* For any $0 \le \alpha \le 1$ and $\mu_1, \mu_2 \in \mathcal{M}_1$ with densities $\phi_1$ and $\phi_2$, respectively, we rewrite the left hand side in (63) as follows,

$$\Phi_f\left((1-\alpha)\,\mu_1 + \alpha\,\mu_2\right) = (1-\alpha)\,\Phi_f\left(\mu_1\right) + \alpha\,\Phi_f\left(\mu_2\right) - \alpha\,(1-\alpha)\,\Psi_f(\mu_1,\mu_2), \tag{64}$$

where

$$\begin{aligned}\Psi_f(\mu_1,\mu_2) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})(\mu_1-\mu_2)(d\mathbf{x})(\mu_1-\mu_2)(d\mathbf{y}) \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(\mathbf{x}-\mathbf{y})(\phi_1-\phi_2)(\mathbf{x})(\phi_1-\phi_2)(\mathbf{y})\, d\mathbf{x}d\mathbf{y}, \end{aligned} \tag{65}$$

see e.g. [29, p. 71]. It follows that $\Phi_f$ is convex if and only if (65) is non-negative for all $\phi_1, \phi_2 \in \mathrm{L}^2_{0,c}(\mathbb{R}^d)$.

One direction is clear, since if $\phi_1, \phi_2 \in \mathrm{L}^2_{0,c}(\mathbb{R}^d)$, then $\int_{\mathbb{R}^d}(\phi_1-\phi_2)(\mathbf{x})\,d\mathbf{x} = 0$. Thus, if $f \in \mathrm{CP}(\mathrm{L}^2_0(\mathbb{R}^d))$, then $\Phi_f$ is convex on $\mathcal{M}_1$. For the reverse implication, let $\phi \in \mathrm{L}^2_0(\mathbb{R}^d)$ be such that $\int_{\mathbb{R}^d}\phi(\mathbf{x})\,d\mathbf{x} = 0$, and let $K$ denote the compact support of $\phi$. Take $\phi_1(\mathbf{x}) = \phi(\mathbf{x}) + c$ for $\mathbf{x} \in K$, $\phi_1(\mathbf{x}) = 0$ otherwise, and $\phi_2(\mathbf{x}) = c$ for $\mathbf{x} \in K$, $\phi_2(\mathbf{x}) = 0$ otherwise. Then, $\phi_1$, $\phi_2 \in \mathrm{L}^2_{0,c}(\mathbb{R}^d)$ and $\phi = \phi_1 - \phi_2$. $\square$

Hence, for the functions $-\log(|x|)$, $-\log(|x| + \sqrt{|x|})$ and $\log\Gamma(|x|) - \log\Gamma(|x| + \frac{1}{2})$ $(x \in \mathbb{R} \setminus \{0\})$, see examples 1, 2 and 3 of Section 3.5, the corresponding energy functionals on $\mathcal{M}_1$ are convex.

*Remark.* It is clear that if $f \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}^d))$, then $\Phi_f : \mathcal{M}_2 \to \mathbb{R}$ is convex on $\mathcal{M}_2 := \{\mu \mid \mu$ absolutely continuous with density $\phi \in \mathrm{L}^2_0(\mathbb{R}^d)\}$. Similarly, if $f \in \mathrm{P}(\mathrm{L}^2(\mathbb{R}^d))$, then $\Phi_f : \mathcal{M}_3 \to \mathbb{R}$ is convex on $\mathcal{M}_3 := \{\mu \mid \mu$ absolutely continuous with density $\phi \in \mathrm{L}^2(\mathbb{R}^d)\}$.

Using the above observations in conjunction with Corollaries 2.7.1, 2.7.4, 2.7.6 and 2.7.7, respectively, we present the following results, which are given without proof.

**Corollary 4.2.1** *Let $g \in \mathrm{CM}$ and $f = g(\|\cdot\|^2) : \mathbb{R}^d \to [0,\infty)$. If $f \in \mathrm{L}^1(\mathbb{R}^d)$, then $\Phi_f : \mathcal{M}_3 \to \mathbb{R}$ defined in* (62) *is convex on $\mathcal{M}_3$.*

**Corollary 4.2.2** *Let $g \in \mathrm{CM}$. If $f = g(|\cdot|) \in \mathrm{L}^1(\mathbb{R})$, then $\Phi_f : \mathcal{M}_3 \to \mathbb{R}$ is convex on $\mathcal{M}_3$ $(d = 1)$.*

**Corollary 4.2.3** *Let $g \in \mathrm{CM}$ be non-constant. If $f = g(|\cdot|) \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$, then $\Phi_f : \mathcal{M}_2 \to \mathbb{R}$ is convex on $\mathcal{M}_2$ $(d = 1)$.*

**Corollary 4.2.4** *Let $g \in \mathrm{CM} \cap \mathrm{L}^1_{\mathrm{loc}}((0,\,\infty))$. For any $s > 0$, define*

$$f(\mathbf{x}) = g(\|\mathbf{x}\|^2)e^{-s\|\mathbf{x}\|^2} \quad (\mathbf{x} \in \mathbb{R}^d),$$

*similarly to* (30)*. Then, $\Phi_f : \mathcal{M}_2 \to \mathbb{R}$ is convex on $\mathcal{M}_2$ $(d \geq 2)$.*

It follows directly from Corollary 4.2.3 that the functions $g_i$, for $i = 1, 8, 11, 16, 19, 23$, of Section 2.7 give rise to convex energy functionals $\Phi_{g_i}$ on $\mathcal{M}_2$. In the case when a general functional $\Phi$ is convex and bounded from below on a given set of measures $\mathcal{M}$, there exists at least one measure $\mu^* = \arg\min_{\mu\in\mathcal{M}} \Phi(\mu)$ which minimises $\Phi$.

## 4.3 Optimality criterion

In this short section we compute the directional derivative of $\Phi_f$ along a measure in a given set $\mathcal{M}$ and introduce the notion of the potential of a measure. Moreover, we establish a criterion for finding the optimal measure which minimises a strictly convex energy functional which is bounded from below and show that, in some cases, this measure is a probability measure.

Consider the energy $\Phi_f$ defined in (62), where $f$ is even a.e. and $\mathcal{M}$ denotes a general set of signed measures on a compact subset $\mathcal{X} \subset \mathbb{R}^d$. We assume that $\Phi_f$ is finite on $\mathcal{M}$. Let $\nu \in \mathcal{M}$. Then, by (64), for any $\mu \in \mathcal{M}$, we have for the directional derivative of $\Phi_f$ at $\mu$ in the direction of $\nu$,

$$\begin{aligned}
\mathcal{D}_\nu\big(\Phi_f(\mu)\big) &= \lim_{\alpha\to 0} \frac{\Phi_f((1-\alpha)\mu + \alpha\nu) - \Phi_f(\mu)}{\alpha} \\
&= 2\left(\int_{\mathcal{X}}\int_{\mathcal{X}} f(\mathbf{x}-\mathbf{y})\mu(d\mathbf{x})\nu(d\mathbf{y}) - \Phi_f(\mu)\right) \\
&= 2\left(\int_{\mathcal{X}} \mathcal{P}_\mu(\mathbf{y})\nu(d\mathbf{y}) - \Phi_f(\mu)\right),
\end{aligned}$$

where

$$\mathcal{P}_\mu(\mathbf{y}) := \int_{\mathcal{X}} f(\mathbf{x}-\mathbf{y})\mu(d\mathbf{x}) \quad (y \in \mathcal{X}) \tag{66}$$

is the *potential* of $\mu$ at $\mathbf{y}$, see e.g. [8, p. 256], [34, p. 21].

The following theorem provides a criterion for finding the unique measure which minimises a given energy functional $\Phi_f$ which is both bounded below and strictly convex. We call such a measure *optimal* or the *minimum-energy measure*. Terms such as *equilibrium measure*, see e.g. [34, p. 24], *minimal distribution*, see e.g. [8, p. 256], and, in the area of experimental design, *optimal design*, see e.g. [47], are also widely used in the literature.

**Theorem 4.3.1** ([47, Th. 1]). *Let $f : \mathbb{R}^d \to \mathbb{R}$ be even a.e. and $\mathcal{M}$ denote a set of signed measures on a compact subset $\mathcal{X} \subset \mathbb{R}^d$, with total mass 1. Let $\Phi_f$ denote the energy functional defined in (62), which we assume to be finite, bounded below and strictly convex on $\mathcal{M}$. Then, a measure $\mu^*$ is optimal if the potential $\mathcal{P}_{\mu^*}$ is constant, that is $\mathcal{P}_{\mu^*}(\mathbf{x}) = \Phi_f(\mu^*)$ for all $\mathbf{x} \in \mathcal{X}$.*

## 4.4 Construction of optimal measures for $\Phi_f$ with $f$ classically strictly positive definite

On describing a real-valued function on $\mathbb{R}$ as being classically *strictly positive definite* we mean that the inequality in Definition 2.1.2 is strictly positive for all $x_1, x_2, \ldots, x_n \in \mathbb{R}$ and $v_1, v_2, \ldots, v_n \in \mathbb{R}\backslash\{0\}$, with any $n \in \mathbb{N}$. Such functions are considered at this stage, as opposed to those which are solely positive definite, so that the matrix $\left[f(x_i - x_j)\right]_{i,j=1}^{n}$ is (strictly) positive definite for all $n \in \mathbb{N}$ and $x_1, x_2, \ldots, x_n \in \mathbb{R}$, and hence, invertible. Moreover, roughly speaking, a strictly positive definite function $f$ generates a strictly convex energy $\Phi_f$, see e.g. [47, Lemma 1], which, in turn, gives rise to a unique and therefore, optimal minimising measure.

For a real-valued, strictly positive definite function $f$ we can construct the optimal measure for the energy $\Phi_f$ using the following approach. Note, for simplicity in our numerical algorithms, we consider functions defined on $\mathbb{R}$, or some compact subset of $\mathbb{R}$, in the remaining sections of this chapter.

Firstly, consider the discrete case. Assume $\mathcal{X} \subset \mathbb{R}$ is of the form $\mathcal{X} = \mathcal{X}_N = \{x_1, \ldots, x_N\}$ and $\mathcal{M}$ is a set of signed measures on $\mathcal{X}_N$ with total mass 1. Let $\mathbf{1} = (1, 1, \ldots, 1)^T$ be the vector of ones of size $N$ and $\mathbf{f} = \left[f(x_i - x_j)\right]_{i,j=1}^{N}$. For $\mathbf{w} = (w_1, \ldots, w_N)^T$, the vector of weights assigned to the points $x_k$ $(k = 1, \ldots, N)$ by a measure $\mu \in \mathcal{M}$, the energy is $\Phi_f(\mu) = \Phi_f(\mathbf{w}) = \mathbf{w}^T \mathbf{f}\, \mathbf{w}$. The vector of optimal weights can be easily computed, for

$$\frac{\partial \Phi_f(\mathbf{w})}{\partial w_j} = \mathbf{e}_j^T \mathbf{f}\, \mathbf{w} + \mathbf{w}^T \mathbf{f}\, \mathbf{e}_j = 2\, \mathbf{e}_j^T \mathbf{f}\, \mathbf{w} \quad (j = 1, \ldots, N),$$

by the product rule, where $\mathbf{e}_j^T = (0, \ldots, 0, 1, 0, \ldots, 0)$ denotes the $j^{th}$ standard basis vector, and, by the method of Lagrange multipliers,

$$\frac{\partial \mathcal{L}(\lambda, \mathbf{w})}{\partial w_j} = \frac{\partial \left(\Phi_f(\mathbf{w}) - \lambda \left(\sum_{i=1}^{N} w_i - 1\right)\right)}{\partial w_j} = 0 \quad (j = 1, \ldots, N)$$

when $\lambda = 2\, \mathbf{e}_j^T \mathbf{f}\, \mathbf{w}$ for all $j = 1, \ldots, N$: in other words, when $\mathbf{w} = \frac{\lambda}{2} \mathbf{f}^{-1} \mathbf{1}$. Hence,

$$\sum_{i=1}^{N} w_i = \mathbf{1}^T \mathbf{w} = \frac{\lambda}{2} \mathbf{1}^T \mathbf{f}^{-1} \mathbf{1} = 1$$

and the vector of optimal weights $\mathbf{w}^*$ is

$$\mathbf{w}^* = \mathbf{f}^{-1} \mathbf{1} / \left(\mathbf{1}^T \mathbf{f}^{-1} \mathbf{1}\right), \tag{67}$$

giving

$$\Phi(\mathbf{w}^*) = \min_{\mathbf{w}} \Phi_f(\mathbf{w}) = 1 / \left(\mathbf{1}^T \mathbf{f}^{-1} \mathbf{1}\right),$$

where the minimum is taken over all vectors $\mathbf{w} = (w_1, \ldots, w_N)^T$ such that $\mathbf{1}^T \mathbf{w} = \sum_{i=1}^{N} w_i = 1$. The potential of the optimal measure $\mu^* \in \mathcal{M}$ with weights $\mathbf{w}^*$ is the vector $\mathcal{P}_{\mathbf{w}^*} = \mathbf{f}\, \mathbf{w}^* = \mathbf{1} / \left(\mathbf{1}^T \mathbf{f}^{-1} \mathbf{1}\right)$.

Note, by [17, Th. 5.3], if $f$ is convex on $\mathcal{X}$, in addition to being (strictly) positive definite, then the components of the vector of optimal weights $\mathbf{w}^* = (w_1^*, \ldots, w_N^*)^T$ are all non-negative, meaning that $\mu^* \in \mathcal{M}$ is automatically a probability measure.

Next, consider the general case. For arbitrary $\mathcal{X} \subset \mathbb{R}$ we approximate $\mathcal{X}$ with a discrete set $\mathcal{X}_N = \{x_1, \ldots, x_N\}$ and, in doing so, we approximate the original problem of finding the optimal measure for the energy $\Phi_f$ defined in (62) with the discrete problem of optimising the energy $\Phi_f(\mathbf{w}) = \mathbf{w}^T \mathbf{f}\, \mathbf{w}$. As discussed above, this discrete problem has the unique solution $\mathbf{w}^*$, see (67). In the main cases of interest, both the continuous optimal measure $\mu^*$ and the discrete optimal measure with weights $\mathbf{w}^*$ are probability measures. Thus, we can easily construct the continuous optimal measure from the discrete one, building piece-wise constant or continuous piece-wise linear approximations to the optimal density, for example.

### 4.5 Approximations to the optimal measure for the Riesz energy

Using the results of Section 4.1 we now demonstrate how to construct accurate approximations to the optimal measure of the Riesz energy.

The Riesz kernel of order $\alpha \in (0,1)$ on $[0,1]$ given by $\kappa_\alpha(x,y) = |x-y|^{-\alpha}$, $x, y \in [0,1]$, $x \neq y$, was first studied by Riesz [32], see also [21]. We define $f_\alpha = |\cdot|^{-\alpha}$ to be the function associated with $\kappa_\alpha$, so that

$$\kappa_\alpha(x,y) = f_\alpha(x-y) = |x-y|^{-\alpha} \quad (x,\, y \in [0,1],\ x \neq y,\ 0 < \alpha < 1). \tag{68}$$

Consider the Riesz energy, i.e. the energy functional $\Phi_f : \mathcal{M} \to [0, +\infty]$ defined in (62) where $\mathcal{M}$ denotes the set of probability measures on $[0,1]$ and $f = f_\alpha$. In this case the optimal measures $\mu_\alpha^*$ are known, see e.g. [37, Corollary 1]; in fact, they are probability measures with densities

$$\phi_\alpha(t) = \frac{\Gamma\big(\alpha+1\big)}{\Gamma^2\big(\frac{\alpha+1}{2}\big)} \big(t(1-t)\big)^{(\alpha-1)/2} \quad (t \in [0,1]). \tag{69}$$

That is, $\mu_\alpha^*$ have densities of a Beta distribution on $[0,1]$ with parameters $(\alpha+1)/2$ and $(\alpha+1)/2$. Note, the strict convexity of the energy $\Phi_{f_\alpha}$ follows from [37, Th. 3], where only measures $\mu_1, \mu_2 \in \mathcal{M}$ such that $\Phi_{f_\alpha}(\mu_1) < \infty$ and $\Phi_{f_\alpha}(\mu_2) < \infty$ are considered in (63).

For any $\alpha \in (0,1)$ and $y \in [0,1]$, $y \neq x$, we have for the Riesz potential of $\mu_\alpha^*$ at $y$, see [37, Prop. 1],

$$\mathcal{P}_{\mu_\alpha^*}(y) = \int_0^1 |y-x|^{-\alpha} \phi_\alpha(x)\, dx = \frac{\Gamma\big(\frac{1-\alpha}{2}\big)\Gamma\big(\alpha+1\big)}{\Gamma\big(\frac{\alpha+1}{2}\big)} = \Phi_{f_\alpha}^*, \tag{70}$$

where $\Phi^*_{f_\alpha} = \min_{\mu\in\mathcal{M}} \Phi_{f_\alpha}(\mu)$. Values of $\Phi^*_{f_\alpha}$ are plotted in Fig. 1 (left). Normalised values of $\left(\Phi^*_{f_\alpha}\right)^{1-\alpha}$, see Fig. 1 (right), appear more interesting.

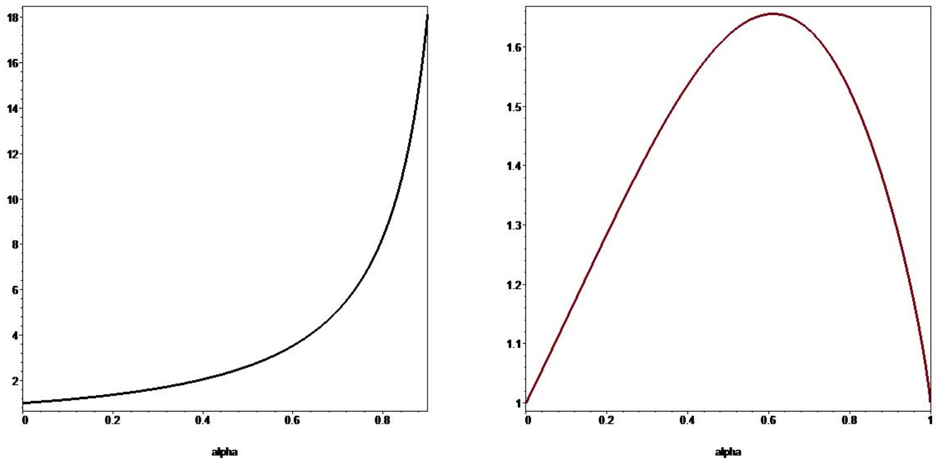



Figure 1: *Left: values of* $\Phi^*_\alpha$ *for* $\alpha \in [0, 0.9]$. *Right: values of* $(\Phi^*_\alpha)^{1-\alpha}$ *for* $\alpha \in [0, 1]$.

Next, we discuss how well the uniform measure minimises the energy $\Phi_{f_\alpha}$ compared to the optimal measures $\mu^*_\alpha$. Let $\mu_0$ denote the uniform probability distribution on $[0, 1]$. For any $\alpha \in (0, 1)$, we define the efficiency of $\mu_0$ as

$$\mathrm{eff}(\mu_0) = \frac{\Phi^*_{f_\alpha}}{\Phi_{f_\alpha}(\mu_0)} = \frac{(1-\alpha)(2-\alpha)\Gamma\left(\frac{1-\alpha}{2}\right)\Gamma\left(\alpha+1\right)}{2\,\Gamma\left(\frac{\alpha+1}{2}\right)}, \tag{71}$$

where

$$\Phi_{f_\alpha}(\mu_0) = \int_0^1\int_0^1 |x-y|^{-\alpha}dxdy = \frac{2}{(1-\alpha)(2-\alpha)}$$

is the energy of the uniform measure. For all $\alpha \in (0, 1)$ the efficiency is reasonably high, see Fig. 2 (left) below. In fact, the lowest value of the efficiency is $\simeq 0.98135$, which is achieved when $\alpha \simeq 0.36253$. This demonstrates that the behaviour of the energy $\Phi_{f_\alpha}$ for the uniform measure $\mu_0$ is indicative of that for the optimal measure $\mu^*_\alpha$.

For $\alpha \in (0, 1)$ and $y \in [0, 1]$, $y \neq x$, the Riesz potential of $\mu_0$ at $y$ is,

$$\mathcal{P}_{\alpha,\mu_0}(y) = \int_0^y (y-x)^{-\alpha}dx + \int_y^1 (x-y)^{-\alpha}dx = \frac{y^{1-\alpha}+(1-y)^{1-\alpha}}{1-\alpha}.$$

This potential, along with its average value $\Phi_{f_\alpha}(\mu_0) = \int_0^1 \mathcal{P}_{\alpha,\mu_0}(y)\,\mu_0(dy)$, is plotted in Fig. 2 (right) for $\alpha = 0.5$. Despite the fact that the uniform measure is highly efficient, as discussed above, there is still scope for improvement in the approximation to the optimal measure. Indeed, the potential of the optimal measure is a constant function.

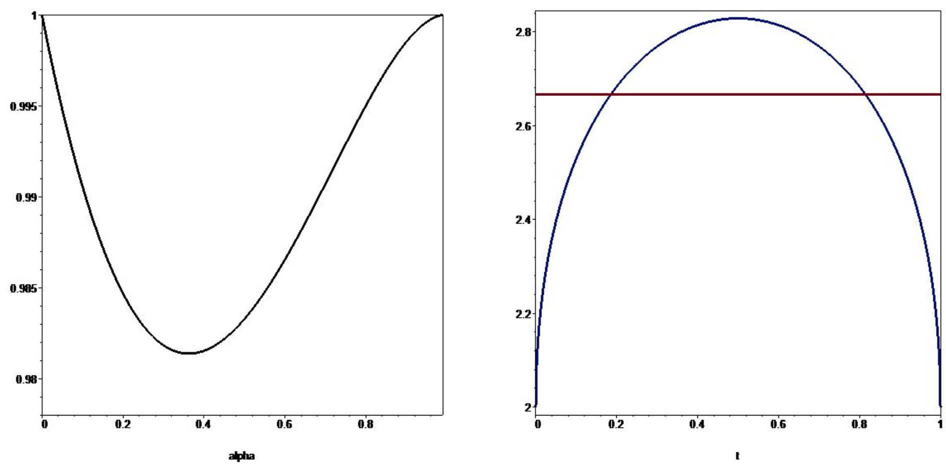



Figure 2: *Left: efficiency of the uniform measure, see* (71)*, for* $\alpha \in [0, 1)$. *Right: potential of the uniform measure* $\mathcal{P}_{\alpha,\mu_0}(t)$ *and its average value* $\Phi_{f_\alpha}(\mu_0)$ *computed for* $\alpha = 0.5$.

The primitive of the completely monotone function $f_\alpha|_{(0,\infty)} = (\cdot)^{-\alpha}$, $0 < \alpha < 1$, is the Bernstein function $g_\alpha(x) = x^{1-\alpha}/(1-\alpha)$ $(x > 0)$. For $\epsilon > 0$, the corresponding family of functions defined in (59) are

$$f_{\alpha,\epsilon}(x) = \frac{(x+\epsilon)^{1-\alpha} - x^{1-\alpha}}{\epsilon\,(1-\alpha)} \quad (x > 0,\ 0 < \alpha < 1).$$

For any $x \in [-1, 1]\backslash\{0\}$ and $0 < \alpha < 1$, let

$$\tilde{f}_{\alpha,\epsilon}(x) := f_{\alpha,\epsilon}(|x|). \tag{72}$$

We now study the quality of the approximations to the energy $\Phi_{f_\alpha}$ (for $f_\alpha$ defined in (68)) by $\Phi_{\tilde{f}_{\alpha,\epsilon}}$ (for $\tilde{f}_{\alpha,\epsilon}$ defined in (72)).

For any $\epsilon > 0$ and $0 < \alpha < 1$, the energy $\Phi_{\tilde{f}_{\alpha,\epsilon}}$ of the uniform measure is

$$\Phi_{\tilde{f}_{\alpha,\epsilon}}(\mu_0) = \int_0^1 \int_0^1 f_{\alpha,\epsilon}(|x-y|)\,dxdy = 2\frac{(1+\epsilon)^{3-\alpha} - \epsilon^{3-\alpha} - (3-\alpha)\epsilon^{2-\alpha} - 1}{\epsilon(1-\alpha)(2-\alpha)(3-\alpha)}. \tag{73}$$

Since $f_{\alpha,\epsilon}(x) < f_\alpha(x)$ for any $\alpha \in (0,1)$, $x > 0$ and $\epsilon > 0$, it follows that $\Phi_{\tilde{f}_{\alpha,\epsilon}}(\mu_0) < \Phi_{f_\alpha}(\mu_0)$ for any $\alpha \in (0,1)$ and $\epsilon > 0$. Values of the ratio $\Phi_{\tilde{f}_{\alpha,\epsilon}}(\mu_0)/\Phi_{f_\alpha}(\mu_0)$ are plotted in Fig. 3 (left). Observe that if $\alpha$ is not too close to 1, that is, if the singularity of $f_\alpha$ is not too strong, then $\Phi_{\tilde{f}_{\alpha,\epsilon}}(\mu_0)$ can be considered as an accurate approximation to $\Phi_{f_\alpha}(\mu_0)$, even for relatively large $\epsilon$.

*Remark.* The case when $f_\alpha$ has a strong singularity, i.e. when $\alpha$ is close to 1, is not overly interesting, since for any $t \in [0,1]$, $\phi_\alpha(t) \to 1$ as $\alpha \to 1$ and hence, $\mu^*_\alpha \to \mu_0$ - the uniform probability distribution on $[0,1]$, as $\alpha \to 1$.

On re-writing the equation in (73) and using Newton's generalised binomial theorem we obtain, for any $0 < \alpha < 1$,

$$\begin{aligned}\Phi_{\tilde{f}_{\alpha,\epsilon}}(\mu_0) &= \Phi_{f_\alpha}(\mu_0)\frac{(1+\epsilon)^{3-\alpha} - \epsilon^{3-\alpha} - (3-\alpha)\epsilon^{2-\alpha} - 1}{\epsilon(3-\alpha)} \\ &= \Phi_{f_\alpha}(\mu_0)\big(1 - \epsilon^{1-\alpha} + \epsilon(1-\alpha/2) + \mathcal{O}(\epsilon^{2-\alpha})\big), \quad \epsilon \to 0;\end{aligned}$$

hence the reason why $\Phi_{\tilde{f}_{\alpha,\epsilon}}(\mu_0)$ provides such a precise approximation to $\Phi_{f_\alpha}(\mu_0)$. Already, the simple estimate

$$\Phi_{\tilde{f}_{\alpha,\epsilon}}(\mu_0)/\Phi_{f_\alpha}(\mu_0) \simeq 1 - \epsilon^{1-\alpha} \quad (\epsilon \simeq 0) \tag{74}$$

demonstrates the accuracy of the approximation, see Fig. 3 (right).

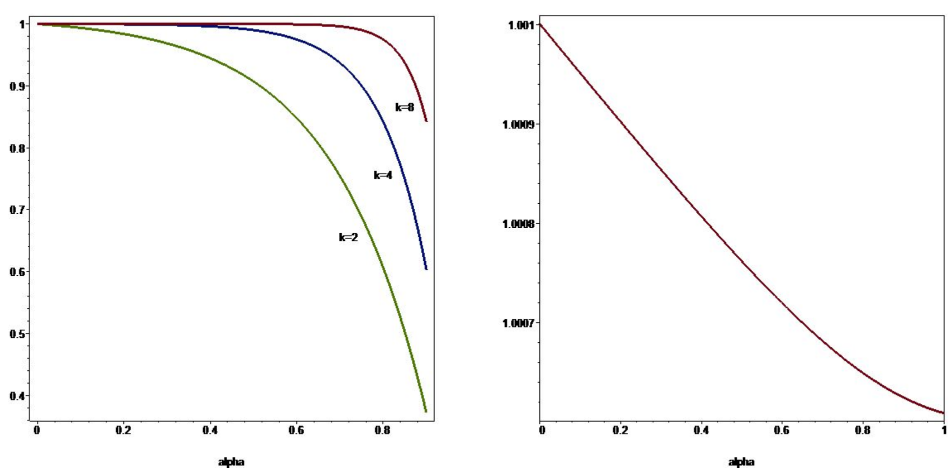



Figure 3: *Left: ratios* $\Phi_{\tilde{f}_{\alpha,\epsilon}}(\mu_0)/\Phi_{f_\alpha}(\mu_0)$ *for* $\epsilon = 10^{-k}$, $k = 2, 4, 8$. *Right: quality of the approximation in* (74) *for* $\epsilon = 0.001$ *and* $\alpha \in [0,1)$.

Next, we demonstrate an application of the methodology outlined in Section 4.4 and construct approximations to the optimal measures for the energies $\Phi_{\tilde{f}_{\alpha,\epsilon}}$ with $\tilde{f}_{\alpha,\epsilon}$ as defined in (72).

Let $N \in \mathbb{N}$ and choose $N+1$ points $0 \le x_0 < x_1 < \ldots < x_N \le 1$ in $[0,1]$; for example, set $x_k = k/N$ for $k = 0,1,\ldots,N$. For $\alpha \in (0,1)$, form the matrix $\tilde{\mathbf{f}}_{\alpha,\epsilon,N} = \big[\tilde{f}_{\alpha,\epsilon}(x_i - x_j)\big]_{i,j=1}^{N}$. It follows from (67) that the optimal weights are given by

$$\mathbf{w}^*_{\alpha,\epsilon,N} = \tilde{\mathbf{f}}^{-1}_{\alpha,\epsilon,N}\mathbf{1}/\big(\mathbf{1}^T\,\tilde{\mathbf{f}}^{-1}_{\alpha,\epsilon,N}\mathbf{1}\big), \tag{75}$$

where $\mathbf{1}$ denotes the vector of ones of size $N+1$.

The discrete energy $\Phi_{\tilde{f}_{\alpha,\epsilon},N}(\mathbf{w}) := \mathbf{w}^T\,\tilde{\mathbf{f}}_{\alpha,\epsilon,N}\,\mathbf{w}$ is minimised when $\mathbf{w} = \mathbf{w}^*_{\alpha,\epsilon,N}$, the vector of optimal weights, in which case

$$\Phi_{\tilde{f}_{\alpha,\epsilon},N}(\mathbf{w}^*_{\alpha,\epsilon,N}) = 1/\big(\mathbf{1}^T\,\tilde{\mathbf{f}}^{-1}_{\alpha,\epsilon,N}\mathbf{1}\big).$$

For fixed $\epsilon > 0$ and $\alpha \in (0,1)$, $\Phi_{\tilde{f}_{\alpha,\epsilon},N}(\mathbf{w}^*_{\alpha,\epsilon,N}) \to \min_{\mu\in\mathcal{M}} \Phi_{\tilde{f}_{\alpha,\epsilon}}(\mu)$ as $N \to \infty$, where $\mathcal{M}$ denotes the set of all probability measures on $[0,1]$. This convergence could be slow however, since for small $\epsilon$, $\tilde{f}_{\alpha,\epsilon}$ is very sharp at zero and approximates a function with a singularity at the origin. Note, also, that if $\epsilon$ is small, the value of $\Phi_{\tilde{f}_{\alpha,\epsilon},N}(\mathbf{w}^*_{\alpha,\epsilon,N})$ can be far away from $\Phi^*_{f_\alpha}$, see (70), since for any $N\in\mathbb{N}$, $\Phi_{\tilde{f}_{\alpha,\epsilon},N}(\mathbf{w}^*_{\alpha,\epsilon,N}) \to \infty$ as $\epsilon \to 0$.

Discrete measures do not provide accurate approximations to the optimal measures for the Riesz energy, since $f_\alpha$ is singular at the origin ($0 < \alpha < 1$). However, using the following general scheme, we construct continuous probability measures which approximate the discrete measures with weights (75) and, in turn, provide precise estimates of the optimal probability measures with densities (69).

Let $0 = x_0 < x_1 < \ldots < x_N = 1$ be the support points of a discrete probability measure and $w_k \ge 0$, $k = 0,\ldots,N$, be the corresponding weights with $\sum_{k=0}^N w_k = 1$. Firstly, define $N+2$ points $z_i$ ($i = 0,\ldots,N+1$) by $z_0 = 0$, $z_{N+1} = 1$ and $z_j = (x_{j-1} + x_j)/2$ for $j = 1,\ldots,N$. Next, partition the interval $[0,1)$ into $N+1$ non-intersecting intervals $I_k = [z_k, z_{k+1})$, $k = 0,\ldots,N$. Denote by $l_k$, the length of the interval $I_k$: that is, $l_k = z_{k+1} - z_k$. We have $l_k > 0$ for all $k = 0,\ldots,N$, and $\sum_{k=0}^N l_k = 1$. Note that if $x_k = k/N$ for $k = 0,1,\ldots,N$, then $l_0 = l_N = 1/(2N)$ and $l_n = 1/N$ for $n = 1,\ldots,N-1$.

Define the piece-wise constant function

$$p_N(t) = \begin{cases} w_k/l_k, & t \in I_k \quad \text{for some } k = 0,1,\ldots,N \\ 0, & t \notin [0,1). \end{cases} \tag{76}$$

Then, $p_N \geq 0$ and $\int_0^1 p_N(t)\,dt = \sum_{k=0}^N w_k = 1$, and therefore, $p_N$ is a probability density function. We shall use $p_N$ as a continuous approximation to the discrete probability measure supported on $0 = x_0 < x_1 < \ldots < x_N = 1$ with weights $w_k$ $(k = 0, \ldots, N)$.

For a measure $\mu_N$ with density $p_N$, see (76), we have for the Riesz energy,

$$
\begin{aligned}
\Phi_{f_\alpha}(\mu_N) &= \int_0^1 \int_0^1 |x-y|^{-\alpha} p_N(x) p_N(y)\,dxdy \\
&= 2\sum_{i=1}^N \frac{w_i}{l_i} \int_{I_i} \left[ \sum_{j=0}^{i-1} \frac{w_j}{l_j} \int_{I_j} (x-y)^{-\alpha} dy \right] dx + \sum_{i=0}^N \frac{w_i^2}{l_i^2} \int_{I_i} \int_{I_i} |x-y|^{-\alpha} dxdy \\
&= 2 \sum_{0 \leq j < i}^N \frac{w_i}{l_i} \frac{w_j}{l_j} \int_{z_i}^{z_{i+1}} \int_{z_j}^{z_{j+1}} (x-y)^{-\alpha}\,dydx \\
&\qquad + \sum_{i=0}^N \frac{w_i^2}{l_i^2} \int_{z_i}^{z_{i+1}} \int_{z_i}^{z_{i+1}} |x-y|^{-\alpha}\,dxdy \quad (0 < \alpha < 1).
\end{aligned}
$$

Moreover, for any $0 \leq a \leq b \leq c \leq d$,

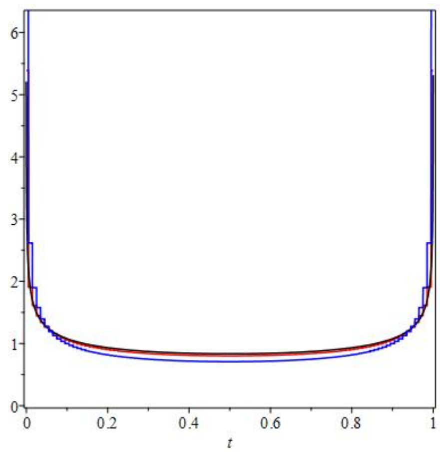
$$
\int_a^b \int_a^b |x-y|^{-\alpha}\,dxdy = \frac{2\,(b-a)^{2-\alpha}}{(1-\alpha)(2-\alpha)}
$$

and

$$
\int_c^d \int_a^b (x-y)^{-\alpha} dydx = \frac{(d-a)^{2-\alpha} + (c-b)^{2-\alpha} - (d-b)^{2-\alpha} - (c-a)^{2-\alpha}}{(1-\alpha)(2-\alpha)}.
$$

Combining the above formulas we obtain an explicit expression for computing $\Phi_{f_\alpha}(\mu_N)$ and, in turn, the efficiency of $\mu_N$, $\text{eff}(\mu_N) = \Phi^*_{f_\alpha}/\Phi_{f_\alpha}(\mu_N)$.

Let $\mu_N$ be the probability measures with densities $p_N$ defined in (76), where $w_k$ are the $k^{th}$ elements of the vector of optimal weights $\mathbf{w}^*_{\alpha,\epsilon,N}$, see (75). In Tables 1 − 5 of Section 5.1 we highlight the efficiency of the measures $\mu_N$ for various values of $\alpha \in (0,1)$, $N \in \mathbb{N}$ and $\epsilon > 0$. Numerous graphs displaying both the optimal densities $\phi_\alpha$, see (69), and the numerically computed densities described above, see (76), can also be found in Section 5.1. By construction, it is clear that for any $\alpha \in (0,1)$, the approximations improve as $\epsilon \to 0$ and $N \to \infty$. However, computing the densities $p_N$ for $N > 250$ is rather time-consuming due to the calculation of the inverse matrix $\tilde{\mathbf{f}}^{-1}_{\alpha,\epsilon,N}$ in our algorithm. For fixed $N \in \mathbb{N}$, chosen so that the computation time is reasonable, we are faced with the problem of finding a "good" $\epsilon > 0$ which gives rise to accurate minimising probability densities. This $\alpha$-dependent $\epsilon$ is not found by simply choosing $\epsilon > 0$ as small as possible. For $N$ reasonably large we need $\epsilon$ small, but not too small, in order to obtain the best approximations. Roughly speaking, we pass both parameters through the limit at the same time, and in proportion to one another. These observations are demonstrated in Tables $1-5$, see Section 5.1, and Figures 4 and 5 below. For $N = 250$ and $\alpha \in (0, 1)$, $\epsilon = 0.001$ produces very accurate approximating densities, see e.g. Figures 20, 26, 32, 38 and 44 of Section 5.1.



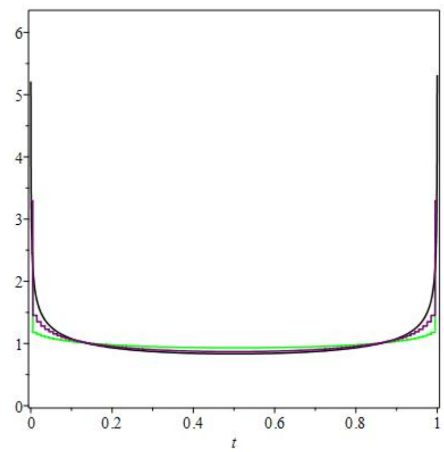
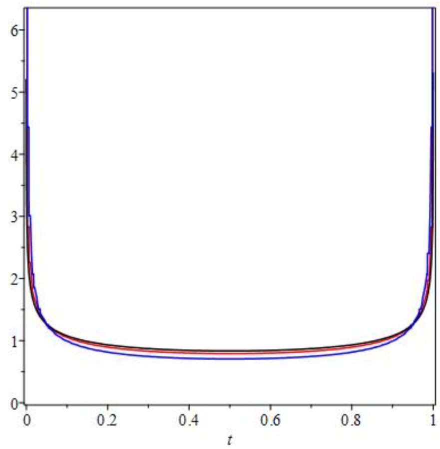




Figure 4: *Optimal density (black), see* (69)*, and numerically computed densities, see* (76)*, on the uniform grid* $x_k = k/N$*,* $k = 0, 1, \ldots, N$*, for* $N = 100$*,* $\alpha = 0.5$*. Left:* $\epsilon = 0.1$ *(blue),* $\epsilon = 0.01$ *(red). Right:* $\epsilon = 0.001$ *(purple),* $\epsilon = 0.0001$ *(green). It is clear that* $\epsilon \simeq 0.01$ *produces the best approximation.*



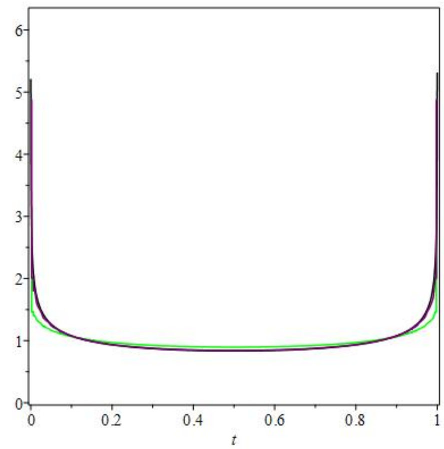




Figure 5: *Optimal density (black), see* (69)*, and numerically computed densities, see* (76)*, on the uniform grid* $x_k = k/N$*,* $k = 0, 1, \ldots, N$*, for* $N = 250$*,* $\alpha = 0.5$*. Left:* $\epsilon = 0.1$ *(blue),* $\epsilon = 0.01$ *(red). Right:* $\epsilon = 0.001$ *(purple),* $\epsilon = 0.0001$ *(green). Hence,* $\epsilon \simeq 0.001$ *now produces the best approximation.*

### 4.6 Numerical approximations to minimising measures for $\Phi_f$ with $f$ unbounded at zero and $f|_{(0,\infty)} \in \mathrm{CM}$

Motivated by the work of Section 4.5 where we constructed precise approximations to the optimal densities for the Riesz energy, we now construct continuous probability measures which accurately estimate the minimising measures for the energy $\Phi_f$, see (62), with alternative singular functions $f$ such that $f|_{(0,\infty)} \in \mathrm{CM}$. Note, $f|_{(0,\infty)}$ denotes the restriction of $f$ to the domain $(0, \infty)$.

The general approach is as follows. Firstly, we take one of the many Bernstein functions $g$ from the list in [35, Chapter 15], so that the following derived function, $f = g'(|\cdot|)$ (up to multiplication by a positive normalising constant), is singular at the origin, and $h := f|_{(0,\infty)} \in \mathrm{CM}$. Next, we construct the corresponding family of functions $h_\epsilon$ defined in (59) (replacing $f$ with $h$). For any $\epsilon \geq 0$, $h_\epsilon \in \mathrm{CM}$ and thus, $\tilde{f}_\epsilon := h_\epsilon(|\cdot|) \in \mathrm{P}(\mathrm{L}^2_0(\mathbb{R}))$ by Corollary 2.7.6. Moreover, since $\tilde{f}_\epsilon$ is continuous for any $\epsilon > 0$, it follows from Proposition 2.4.6 that $\tilde{f}_\epsilon$ is classically positive definite for any $\epsilon > 0$. Using the results of Section 4.4, we then construct the vector of discrete minimising weights $\mathbf{w}^*_{\epsilon,N}$ for the energy $\Phi_{\tilde{f}_\epsilon}$ $(\epsilon > 0)$, see (67), (75), and, subsequently, derive continuous approximations to the density of the minimising probability measure (on $[0, 1]$) for $\Phi_f$, see (76). By [17, Th. 5.3], the numerically computed probability measures are also minimisers over the set of signed measures on $[0,1]$ with total mass 1.

Note, although the functions $\tilde{f}_\epsilon$ are classically positive definite, they need not be strictly positive definite. However, in the examples we consider below, the corresponding functions $\tilde{f}_\epsilon$ are sufficiently strictly positive definite for any $\epsilon > 0$. By "sufficiently" we mean that $\tilde{\mathbf{f}}_{\epsilon,N} := \left[\tilde{f}_\epsilon(x_i - x_j)\right]_{i,j=1}^N$ is invertible for any $N \in [0, \tilde{N}]$, where $\tilde{N}$ is reasonably large (usually $\tilde{N} = 250$), so that for any $\epsilon > 0$ and $N \in [0, \tilde{N}]$, all the eigenvalues of the matrix $\tilde{\mathbf{f}}_{\epsilon,N}$ are positive.

The corresponding energies $\Phi_{\tilde{f}_\epsilon}$, $\epsilon > 0$, are convex on the set of probability measures on $[0,1]$, see e.g. [47, Lemma 1]. Taking $\epsilon \to 0$ and $N \to \infty$ in the approximating density provides an estimate of the density of the minimising (not necessarily unique/optimal) measure for $\Phi_f$.

We assess the accuracy of the approximating measure by computing its potential $\mathcal{P}$, see (66). The closer this value is to a constant, the more accurate the approximation, see Theorem 4.3.1. The following method for computing the potential of an approximating measure will be used in the examples below.

Let $\mu_N$ denote the probability measure with density $p_N$ defined in (76), where $w_k$ are the $k^{th}$ elements of the vector of minimising weights, see (67). Then, for any $N \in \mathbb{N}$ and $x \in I_j = [z_j, z_{j+1})$ $(j = 0, \ldots, N)$,

$$\begin{aligned}
\mathcal{P}_{\mu_N}(x) &= \int_0^1 f(x-y)\, p_N(y)\, dy = \sum_{i=0}^{N} \frac{w_i}{l_i} \int_{I_i} g'(|x-y|)\, dy \\
&= \sum_{i=0}^{j-1} \frac{w_i}{l_i} \int_{z_i}^{z_{i+1}} g'(x-y)\, dy + \frac{w_j}{l_j} \int_{z_j}^{z_{j+1}} g'(|x-y|)\, dy \\
&\qquad + \sum_{k=j+1}^{N} \frac{w_k}{l_k} \int_{z_k}^{z_{k+1}} g'(y-x)\, dy \\
&= \sum_{i=0}^{j-1} \frac{w_i}{l_i} \left[g(x-z_i) - g(x-z_{i+1})\right] + \frac{w_j}{l_j} \left[g(z_{j+1}-x) + g(x-z_j) - 2g(0)\right] \\
&\qquad + \sum_{k=j+1}^{N} \frac{w_k}{l_k} \left[g(z_{k+1}-x) - g(z_k - x)\right]. \qquad (77)
\end{aligned}$$

We provide three examples of numerical approximations to the minimising measures for $\Phi_f$ with $f$ unbounded at the origin and $f|_{(0,\infty)} \in \mathrm{CM}$. Note, in each case we take $\mathcal{M}$ to be the set of probability measures on $[0,1]$.

1. Let $g_\alpha(t) = (1-\alpha)^{-1}\, t^{1-\alpha}/(1+t)^{1-\alpha}$ ($t > 0$, $\alpha \in (0,1)$), so that $f_\alpha(t) = g'_\alpha(|t|) = |t|^{-\alpha}/(1+|t|)^{2-\alpha}$ ($t \in \mathbb{R}$, $\alpha \in (0,1)$). Since $g_\alpha \in \mathrm{BF}$, then $h_\alpha := f_\alpha|_{(0,\infty)} \in \mathrm{CM}$, for any $\alpha \in (0,1)$: in fact, $h_\alpha$ is proportional to the function $g_8$ of Section 2.7. Moreover, for any $\alpha \in (0,1)$, $f_\alpha$ shares the same singularity as the function associated with the Riesz kernel, $|\cdot|^{-\alpha}$. Figure 6 below demonstrates this observation for $\alpha = 0.25$ and $\alpha = 0.75$.

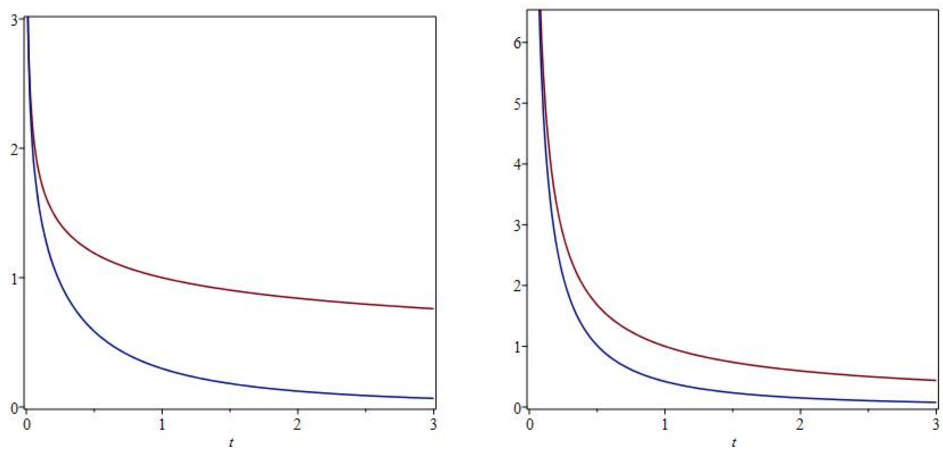



Figure 6: $t^{-\alpha}$ *(red) and* $h_\alpha(t)$ *(blue). Left:* $\alpha = 0.25$*. Right:* $\alpha = 0.75$*.*

For any $\alpha \in (0,1)$ and $\epsilon \geq 0$, the family of functions $h_{\alpha,\epsilon}$ are constructed by (59). Next, we define the classically continuous, positive definite functions $\tilde{f}_{\alpha,\epsilon} := h_{\alpha,\,\epsilon}(|\cdot|)$, and by (75), compute the vector of discrete minimising weights $\mathbf{w}^*_{\alpha,\epsilon,N}$ for the energy $\Phi_{\tilde{f}_{\alpha,\epsilon}}$. Finally, based on the analysis of Section 4.5, and due to the similarity between $f_\alpha$ and the function associated with the Riesz kernel, we use $N = 250$ and $\epsilon = 0.001$ to construct the continuous probability measures $\mu_N$ with densities $p_N$ defined in (76). The approximating probability densities turn out to be very close to the optimal densities $\phi_\alpha$ for the Riesz energy, see (69). Figures 7 and 8 below show the relation between $p_N$ and $\phi_\alpha$ for various values of $\alpha \in (0,1)$.

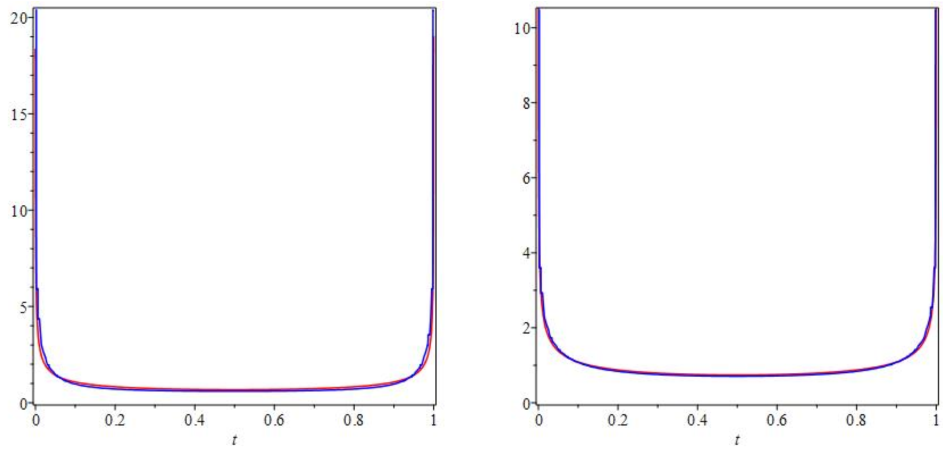



Figure 7: *Optimal density for the Riesz energy* $\phi_\alpha$ *(red), see* (69)*, and the numerically computed density* $p_N$ *(blue) on the uniform grid* $x_k = k/N$*,* $k = 0, 1, \ldots, N$*, for* $N = 250$*,* $\epsilon = 0.001$*. Left:* $\alpha = 0.1$*. Right:* $\alpha = 0.25$*.*

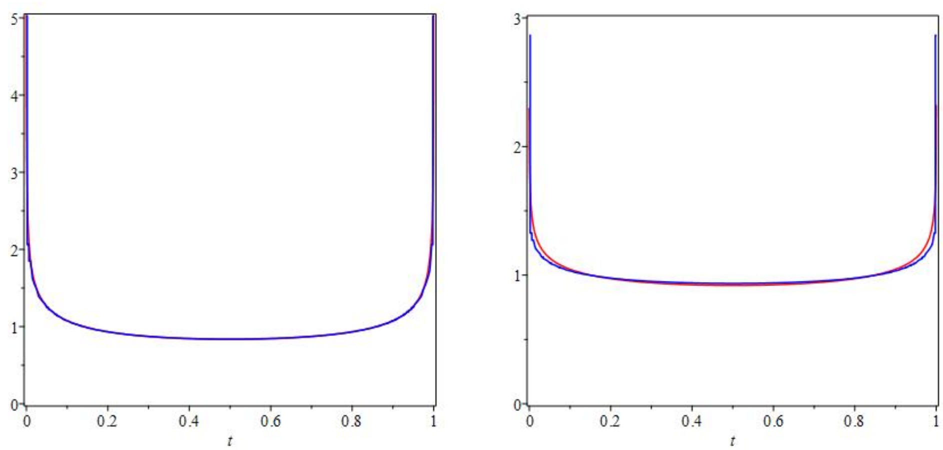



Figure 8: *Optimal density for the Riesz energy* $\phi_\alpha$ *(red), see* (69)*, and the numerically computed density* $p_N$ *(blue) on the uniform grid* $x_k = k/N$*,* $k = 0, 1, \ldots, N$*, for* $N = 250$*,* $\epsilon = 0.001$*. Left:* $\alpha = 0.5$*. Right:* $\alpha = 0.75$*.*

Indeed, it is true that for any $\alpha \in (0, 1)$, there is a strong resemblance between the approximating densities (for $N = 250$ and $\epsilon = 0.001$) and the optimal densities for the Riesz energy; which we believe is due to the similarity between the singularities of $f_\alpha$ and $|\cdot|^{-\alpha}$.

The potential of $\mu_N$ at $x \in [0,1]$,

$$\mathcal{P}_{\mu_N}(x) = \int_0^1 f_\alpha(x-y)\, p_N \, dy = \int_0^1 \frac{|x-y|^{-\alpha}}{(1+|x-y|)^{2-\alpha}}\, p_N(y)\, dy,$$

is approximately constant for any $\alpha \in (0,1)$, see e.g. Figures 45, 46 and 47 of Section 5.1. Thus, by Theorem 4.3.1, we conclude that the approximating probability measures $\mu_N$ are very close to the minimising measures for $\Phi_{f_\alpha}$.

2. Take $g_\lambda(t) = 2\sqrt{t}\left(1+e^{-\lambda\sqrt{t}}\right)$ $(t, \lambda > 0)$, so that $f_\lambda(t) = g'_\lambda(|t|) = \left(1-(\lambda\sqrt{|t|}-1)\, e^{-\lambda\sqrt{|t|}}\right)/\sqrt{|t|}$ $(t \in \mathbb{R},\, \lambda > 0)$. $h_\lambda := f_\lambda|_{(0,\infty)}$ is exactly the function $g_{19}$ of Section 2.7, and since $g_\lambda \in \mathrm{BF}$, then $h_\lambda \in \mathrm{CM}$ for any $\lambda > 0$. Computing the series expansion of $h_\lambda$ about $t = 0$ gives, for any $\lambda > 0$,

$$h_\lambda(t) = \frac{2}{\sqrt{t}} - 2\lambda + \frac{3}{2}\lambda^2\sqrt{t} - \frac{2}{3}\lambda^3 t + \mathcal{O}(t^{\frac{3}{2}}), \quad t \to 0,$$

and hence,

$$h_\lambda(t) \simeq \frac{2}{\sqrt{t}} - 2\lambda \quad (t \simeq 0,\, \lambda > 0).$$

It is clear that for $\lambda \simeq 0$, $h_\lambda(t) \simeq 2t^{-1/2}$ for small $t$. Thus, $f_\lambda(t) \to 2\,|t|^{-1/2}$, which is twice the function associated with the Riesz kernel (68) for $\alpha = 1/2$, as $t, \lambda \to 0$. Fig. 9 (left) below demonstrates that, in fact, for any $t > 0$, $h_\lambda(t) \to 2t^{-1/2}$ as $\lambda \to 0$ and $h_\lambda(t) \to t^{-1/2}$ as $\lambda \to \infty$.

For any $\lambda > 0$ and $\epsilon \geq 0$, the family of functions $h_{\lambda,\epsilon}$ are constructed by (59) (replacing $f$ with $h$). Again, we define the classically continuous, positive definite functions $\tilde{f}_{\lambda,\epsilon} := h_{\lambda,\epsilon}(|\cdot|)$, compute the vector of discrete minimising weights $\mathbf{w}^*_{\lambda,\epsilon,N}$ for the energy $\Phi_{\tilde{f}_{\lambda,\epsilon}}$ (using (75)) and construct the continuous probability measures $\mu_N$ with densities $p_N$ defined in (76). Since, for any $t \in \mathbb{R}$, $f_\lambda(t) \simeq 2\,|t|^{-1/2}$ for small $\lambda$, i.e. $\lambda \leq 0.01$ and $f_\lambda(t) \simeq |t|^{-1/2}$ for reasonably large $\lambda$, i.e. $\lambda \geq 25$, it follows that the approximating probability densities $p_N$ are almost identical to the optimal density $\phi_{1/2}$ for the Riesz energy, see (69), when $\lambda \leq 0.01$ and $\lambda \geq 25$. This is demonstrated in Figures 9 (right) and 10 below. Note, throughout the whole example, we use $N = 250$ and $\epsilon = 0.001$ in our approximations, since for any $\lambda > 0$, $f_\lambda \simeq |\cdot|^{-1/2}$ (up to multiplication by a positive constant) near the origin.

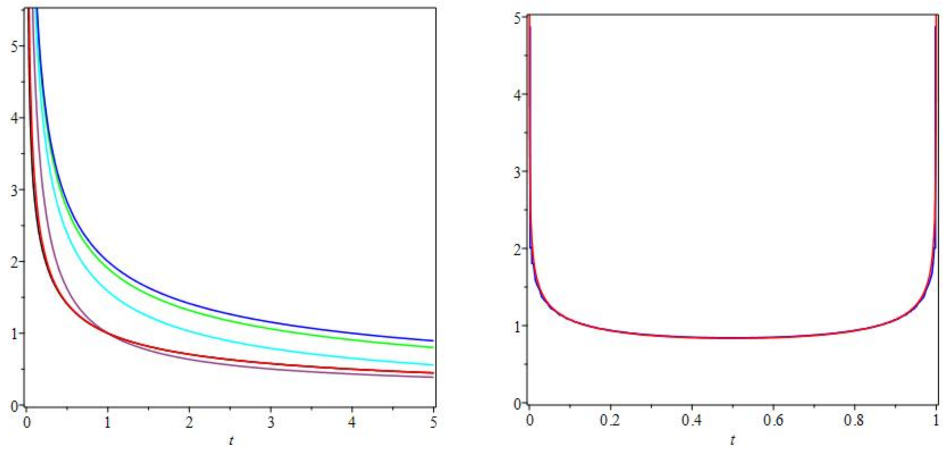



Figure 9: *Left:* $t^{-1/2}$ *(red),* $2\,t^{-1/2}$ *(blue),* $h_{0.05}(t)$ *(green),* $h_{0.25}(t)$ *(cyan),* $h_1(t)$ *(purple),* $h_{10}(t)$ *(black). Right: Optimal density for the Riesz energy* $\phi_{1/2}$ *(red), see* (69)*, and the numerically computed density* $p_N$ *(blue) on the uniform grid* $x_k = k/N$*,* $k = 0, 1, \ldots, N$*, for* $N = 250$*,* $\epsilon = 0.001$*,* $\lambda = 0.001$*.*

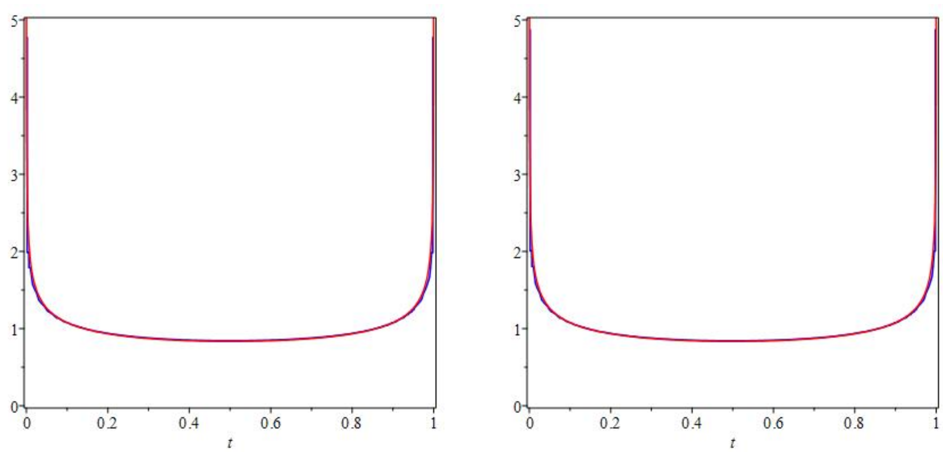



Figure 10: *Optimal density for the Riesz energy* $\phi_{1/2}$ *(red), see* (69)*, and the numerically computed density* $p_N$ *(blue) on the uniform grid* $x_k = k/N$*,* $k = 0, 1, \ldots, N$*, for* $N = 250$*,* $\epsilon = 0.001$*. Left:* $\lambda = 100$*. Right:* $\lambda = 1000$*.*

For $\lambda \in (0.01, 25)$, there exists $\beta \in (1, 2)$ such that $\beta\,|\cdot|^{-1/2}$ has the same singularity as $f_\lambda$ at zero, see e.g. Fig. 11 below, where $\lambda = 1$, $\beta = 1.75$ and $\lambda = 10$, $\beta = 1.1$ are shown.

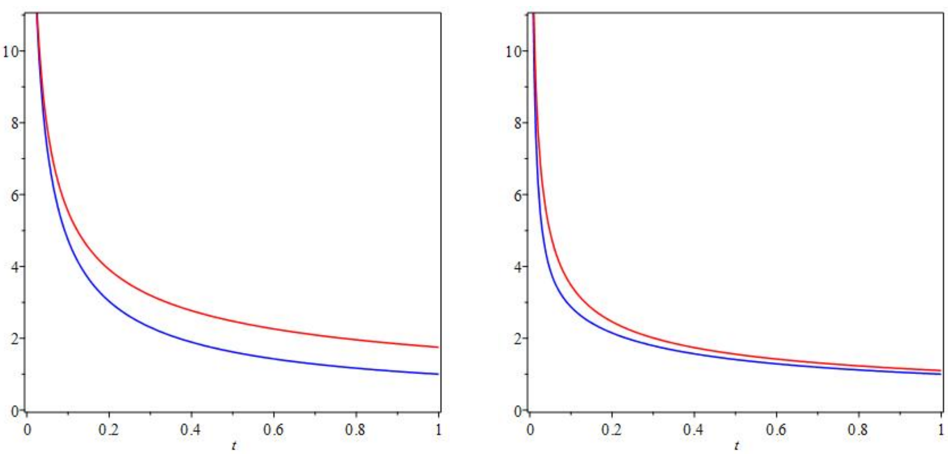



Figure 11: *Left:* $1.75t^{-1/2}$ *(red) and* $h_1(t)$ *(blue). Right:* $1.1\,t^{-1/2}$ *(red) and* $h_{10}(t)$ *(blue).*

For any $\lambda > 0$, the approximate minimising densities for $\Phi_{f_\lambda}$ are closely related to the optimal density $\phi_{1/2}$ for the Riesz energy. Indeed, for $\lambda \leq 0.01$ and $\lambda \geq 25$, they are almost identical, see the discussion above; for $\lambda \in (0.01, 25)$, we believe the similarity between the singularities of $f_\lambda$ and $\beta|\cdot|^{-1/2}$, for some $\beta \in (1,2)$, gives rise to a strong connection between $p_N$ and $\phi_{1/2}$. However, we now notice that the strength of the singularity of $f_\lambda$ plays an important role in how close $p_N$ is to $\phi_{1/2}$. In particular, the stronger the singularity, i.e. the smaller the $\lambda \in (0.01, 25)$, or the closer $f_\lambda$ is to $2\,|t|^{-1/2}$, the stronger the likeness between the approximate minimising densities for $\Phi_{f_\lambda}$ and the optimal density $\phi_{1/2}$ for the Riesz energy, see Fig. 12 below.

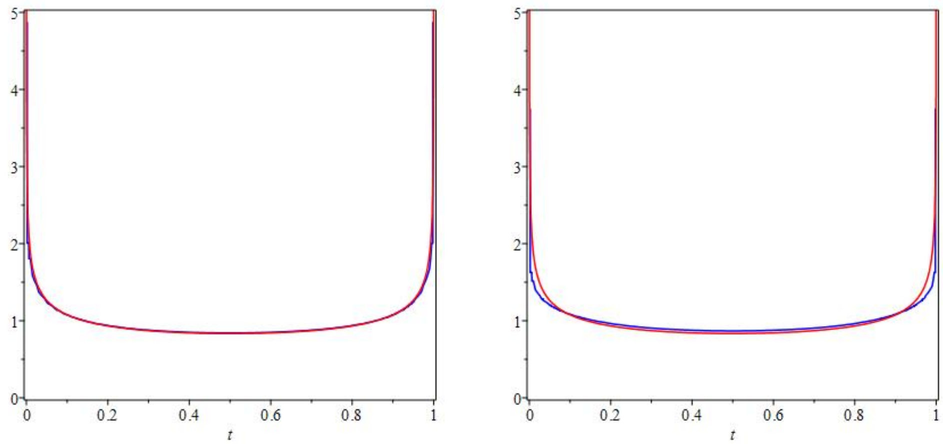



Figure 12: *Optimal density for the Riesz energy* $\phi_{1/2}$ *(red), see* (69), *and the numerically computed density* $p_N$ *(blue) on the uniform grid* $x_k = k/N$, $k = 0, 1, \ldots, N$, *for* $N = 250$, $\epsilon = 0.001$. *Left:* $\lambda = 0.1$. *Right:* $\lambda = 10$.

The potential of $\mu_N$ at $x \in [0, 1]$,

$$\mathcal{P}_{\mu_N}(x) = \int_0^1 f_\alpha(x - y)\, p_N\, dy = \int_0^1 \frac{1-(\lambda\sqrt{|x-y|}-1)\, e^{-\lambda\sqrt{|x-y|}}}{\sqrt{|x-y|}}\, p_N(y)\, dy,$$

is approximately constant for any $\lambda > 0$, see e.g. Figures 48, 49 and 50 of Section 5.1. Hence, by Theorem 4.3.1, we conclude that the probability measures $\mu_N$ accurately approximate the minimising measures for $\Phi_{f_\lambda}$.

3. Let $g(t) = t(1 + 1/t)^{1+t} - 1$ $(t > 0)$, so that $f(t) = g'(|t|) = |t|(1 + 1/|t|)^{1+|t|} \log(1 + 1/|t|)$ $(t \in \mathbb{R})$. $h := f|_{(0,\infty)}$ is equal to the function $g_{23}$ of Section 2.7, and since $g \in \mathrm{BF}$, then $h \in \mathrm{CM}$. On computing the series expansion of $h$ about $t = 0$ we obtain

$$h(t) = -\log(t) + \big(1 - \log(t) - \log(t)^2\big)\, t + \mathcal{O}(t), \quad t \to 0,$$

and hence,

$$h(t) \simeq -\log(t) \quad (t \simeq 0).$$

The density of the optimal measure for the energy $\Phi_f$ with $f = -\log|\cdot|$ is the arcsine density $\phi_{\mathrm{arc}}(t) = t^{-1/2}(1-t)^{-1/2}/\Gamma\big(1/2\big)^2$ $(t \in [0, 1])$, see e.g. [47, Corollary 1]. Note, $\phi_{\mathrm{arc}} = \lim_{\alpha \to 0} \phi_\alpha$; the limiting case of the optimal density for the Riesz energy.

We construct the family of functions $h_\epsilon$ using (59) and define the classically continuous, positive definite functions $\tilde{f}_\epsilon := h_\epsilon(|\cdot|)$. Next, by (75), we compute the vector of discrete minimising weights $\mathbf{w}^*_{\epsilon,N}$ for the energy $\Phi_{\tilde{f}_\epsilon}$. Note that, formally, the matrix $\tilde{\mathbf{f}}_{\epsilon,N}$ blows up on the diagonal. However, by construction, we know that $\tilde{f}_\epsilon$ is bounded at the origin and thus, we replace the diagonal elements $\tilde{f}_\epsilon(0)$ with the numerically computed series expansion of $\tilde{f}_\epsilon(t)$ (or $h_\epsilon(t)$) at $t = 0$. Finally, we construct the continuous probability measures $\mu_N$ with densities $p_N$ using (76).

Since $f$ shares the same singularity as $-\log|\cdot|$, then, based on the analysis of examples 1 and 2, one would expect there to be a similarity between the approximate minimising densities $p_N$ and the optimal probability density for $\Phi_{-\log|\cdot|}$, $\phi_{\mathrm{arc}}$. However, the $p_N$s do not resemble the optimal density as much as we have witnessed in the previous two examples, see Figures 13 and 14 below. Intuitively, we believe this is due (at least partly) to the fact that $f$ and $-\log|\cdot|$ have a relatively weak singularity; for instance, $f_\lambda$ of example 2 is much sharper than $f$ at the origin. We do not investigate this claim further, however, since it is not the focus of our research.

We have for the potential of $\mu_N$ at $x \in [0,1]$,

$$\mathcal{P}_{\mu_N}(x) = \int_0^1 |x-y|\left(1+|x-y|^{-1}\right)^{1+|x-y|}\log\left(1+|x-y|^{-1}\right)p_N(y)\,dy.$$

For $N = 250$ and $\epsilon = 0.001$, $\mathcal{P}_{\mu_N}(x) \simeq 3.139$ for all $x \in [0,1]$, see Fig. 52 (right) of Section 5.1. Hence, by Theorem 4.3.1, we conclude that the corresponding probability measure $\mu_N$ is very close to the minimising measure for $\Phi_f$.

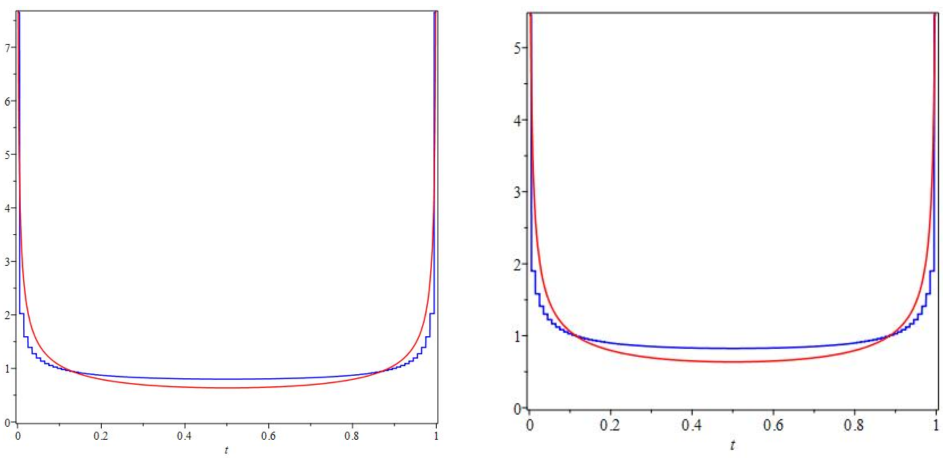



Figure 13: *Arcsine density (red) and the numerically computed density* $p_N$ *(blue) on the uniform grid* $x_k = k/N$*,* $k = 0, 1, \ldots, N$*, for* $N = 100$*. Left:* $\epsilon = 0.01$*. Right:* $\epsilon = 0.001$*.*

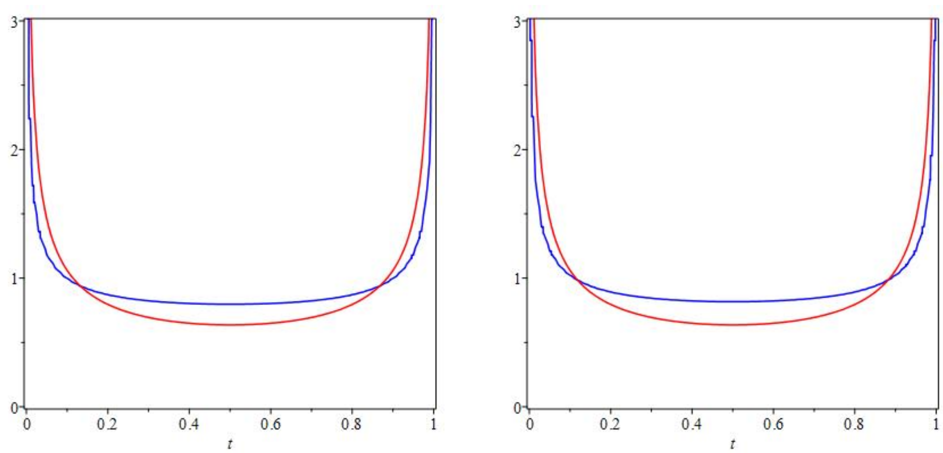



Figure 14: *Arcsine density (red) and the numerically computed density* $p_N$ *(blue) on the uniform grid* $x_k = k/N$*,* $k = 0, 1, \ldots, N$*, for* $N = 250$*. Left:* $\epsilon = 0.01$*. Right:* $\epsilon = 0.001$*.*

# 5 Appendix

## 5.1 Miscellaneous tables and figures

The following tables highlight the efficiency of the measures $\mu_N$ ($\mathrm{eff}(\mu_N) = \Phi_\alpha^*/\Phi_\alpha(\mu_N)$) with densities $p_N$, see (76), for various values of $\alpha \in (0,1)$, $N \in \mathbb{N}$ and $\epsilon > 0$, see Section 4.5 for details.

Table 1: $\mathrm{eff}(\mu_N)$, $\alpha = 0.1$.

| $\alpha$ | $N$ | $\epsilon$ | $\mathrm{eff}(\mu_N)$ |
|---|---|---|---|
| 0.1 | 100 | 0.1 | 0.994455 |
| | | 0.01 | 0.999508 |
| | | 0.001 | 0.999888 |
| | | 0.0001 | 0.999594 |
| | 200 | 0.1 | 0.993406 |
| | | 0.01 | 0.999388 |
| | | 0.001 | 0.999957 |
| | | 0.0001 | 0.999836 |
| | 250 | 0.1 | 0.993066 |
| | | 0.01 | 0.999344 |
| | | 0.001 | 0.999957 |
| | | 0.0001 | 0.999888 |

Table 2: $\mathrm{eff}(\mu_N)$, $\alpha = 0.25$.

| $\alpha$ | $N$ | $\epsilon$ | $\mathrm{eff}(\mu_N)$ |
|---|---|---|---|
| 0.25 | 100 | 0.1 | 0.982312 |
| | | 0.01 | 0.998394 |
| | | 0.001 | 0.999620 |
| | | 0.0001 | 0.998074 |
| | 200 | 0.1 | 0.978736 |
| | | 0.01 | 0.997881 |
| | | 0.001 | 0.999846 |
| | | 0.0001 | 0.999133 |
| | 250 | 0.1 | 0.977586 |
| | | 0.01 | 0.997708 |
| | | 0.001 | 0.999866 |
| | | 0.0001 | 0.999359 |

Table 3: $\mathrm{eff}(\mu_N)$, $\alpha = 0.5$.

| $\alpha$ | $N$ | $\epsilon$ | $\mathrm{eff}(\mu_N)$ |
|---|---|---|---|
| 0.5 | 100 | 0.1 | 0.953078 |
| | | 0.01 | 0.995425 |
| | | 0.001 | 0.998697 |
| | | 0.0001 | 0.993893 |
| | 200 | 0.1 | 0.942722 |
| | | 0.01 | 0.993577 |
| | | 0.001 | 0.999546 |
| | | 0.0001 | 0.995921 |
| | 250 | 0.1 | 0.939382 |
| | | 0.01 | 0.992941 |
| | | 0.001 | 0.999653 |
| | | 0.0001 | 0.996648 |

Table 4: $\mathrm{eff}(\mu_N)$, $\alpha = 0.75$.

| $\alpha$ | $N$ | $\epsilon$ | $\mathrm{eff}(\mu_N)$ |
|---|---|---|---|
| 0.75 | 100 | 0.1 | 0.927120 |
| | | 0.01 | 0.991419 |
| | | 0.001 | 0.997564 |
| | | 0.0001 | 0.995842 |
| | 200 | 0.1 | 0.908509 |
| | | 0.01 | 0.988184 |
| | | 0.001 | 0.998694 |
| | | 0.0001 | 0.996052 |
| | 250 | 0.1 | 0.902328 |
| | | 0.01 | 0.986822 |
| | | 0.001 | 0.998955 |
| | | 0.0001 | 0.996242 |

Table 5: $\mathrm{eff}(\mu_N)$, $\alpha = 0.9$.

| $\alpha$ | $N$ | $\epsilon$ | $\mathrm{eff}(\mu_N)$ |
|---|---|---|---|
| 0.9 | 100 | 0.1 | 0.940572 |
| | | 0.01 | 0.989972 |
| | | 0.001 | 0.995564 |
| | | 0.0001 | 0.995886 |
| | 200 | 0.1 | 0.921485 |
| | | 0.01 | 0.989034 |
| | | 0.001 | 0.997549 |
| | | 0.0001 | 0.997626 |
| | 250 | 0.1 | 0.914352 |
| | | 0.01 | 0.988011 |
| | | 0.001 | 0.997950 |
| | | 0.0001 | 0.997950 |

Next, we present several figures comparing the optimal densities $\phi_\alpha$ for the Riesz energy, see (69), which appear in red throughout, and the numerically computed densities $p_N$, see (76), represented in blue. Note, for simplicity, we use the uniform grid $x_k = k/N$, $k = 0, 1, \ldots, N$, in all approximations.

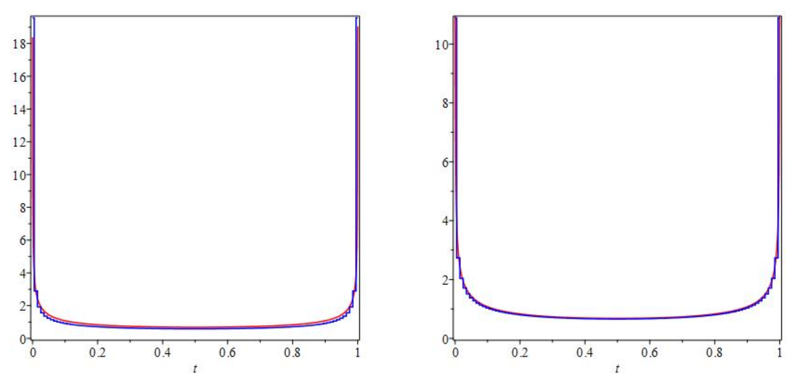


Figure 15: $N = 100$, $\alpha = 0.1$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

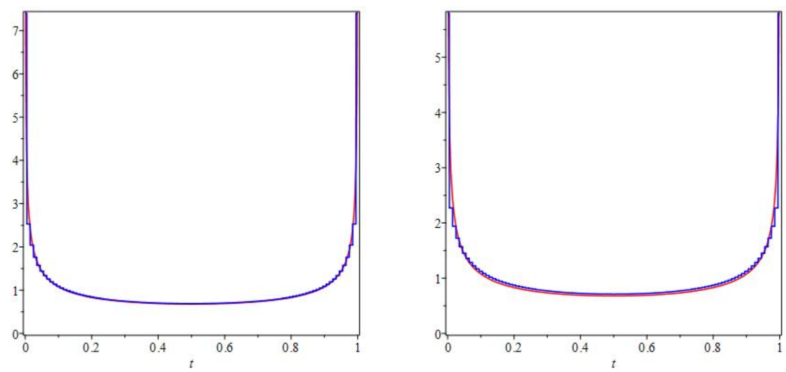



Figure 16: $N = 100$, $\alpha = 0.1$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

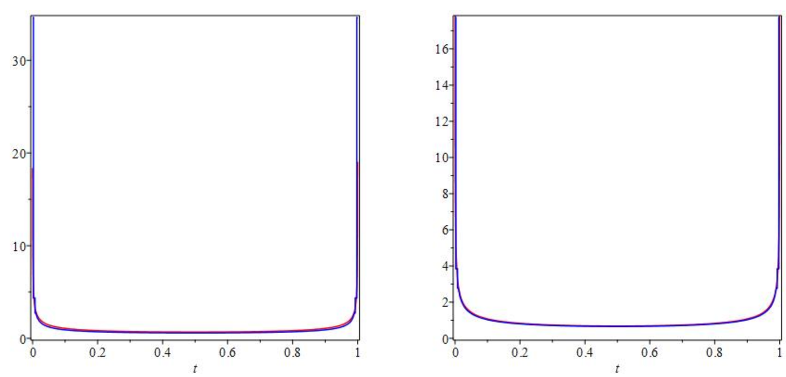



Figure 17: $N = 200$, $\alpha = 0.1$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

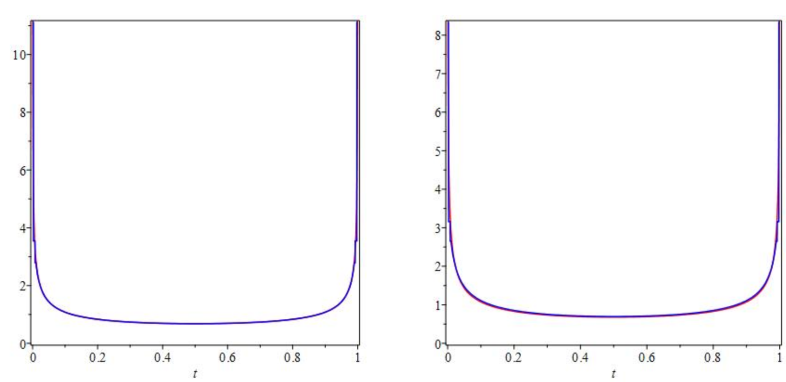



Figure 18: $N = 200$, $\alpha = 0.1$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

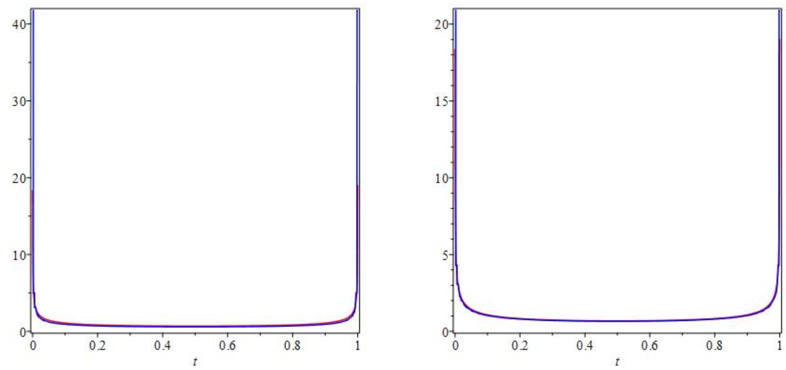



Figure 19: $N = 250$, $\alpha = 0.1$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

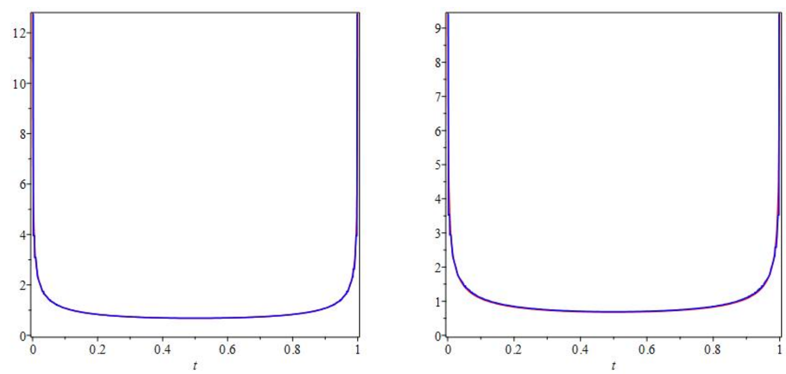



Figure 20: $N = 250$, $\alpha = 0.1$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

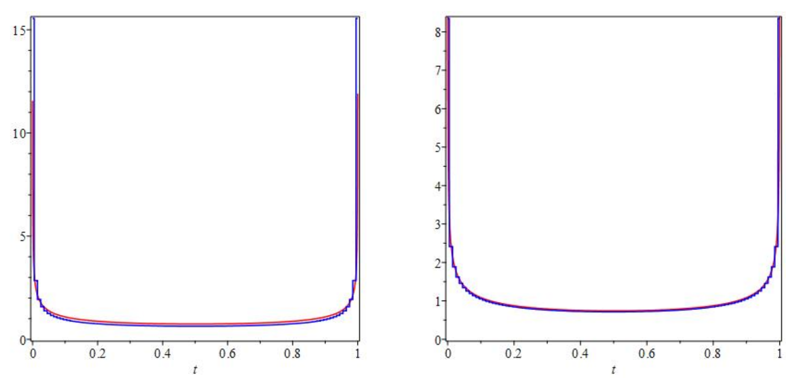



Figure 21: $N = 100$, $\alpha = 0.25$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

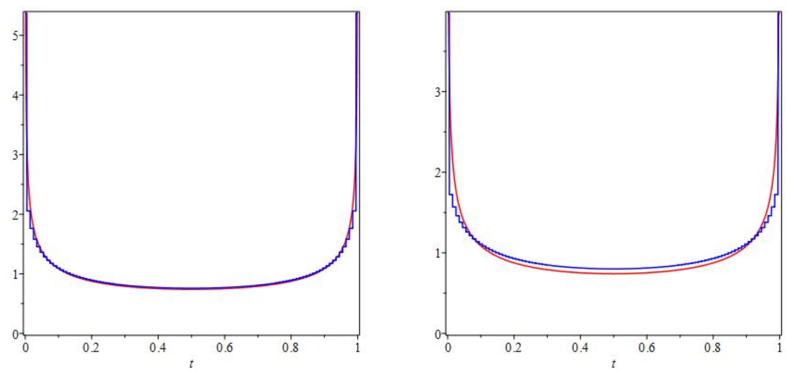



Figure 22: $N = 100$, $\alpha = 0.25$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

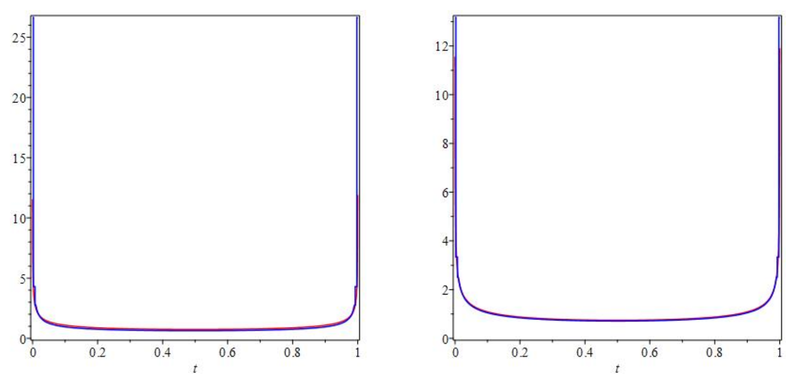



Figure 23: $N = 200$, $\alpha = 0.25$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

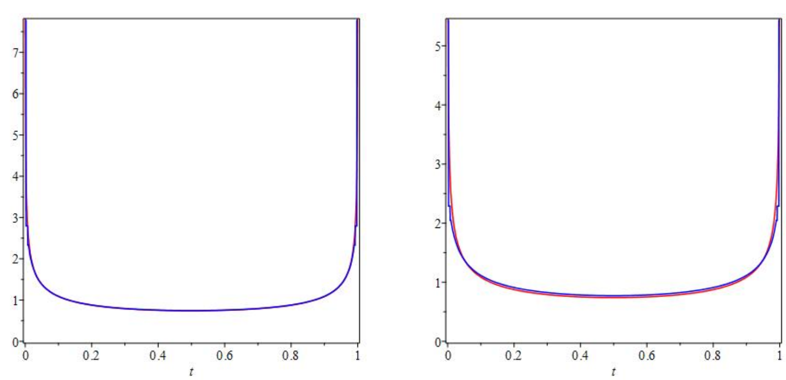



Figure 24: $N = 200$, $\alpha = 0.25$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

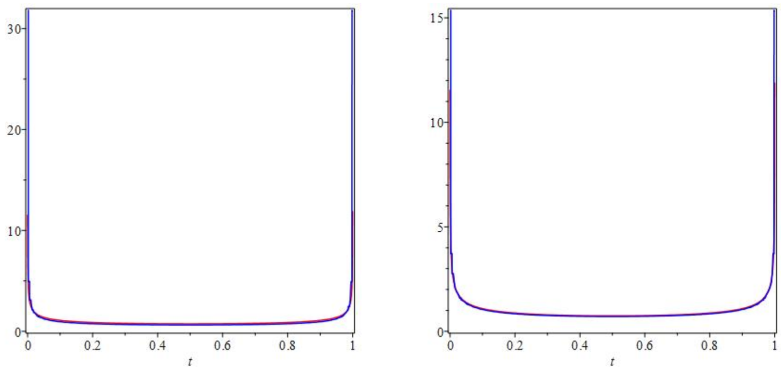



Figure 25: $N = 250$, $\alpha = 0.25$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

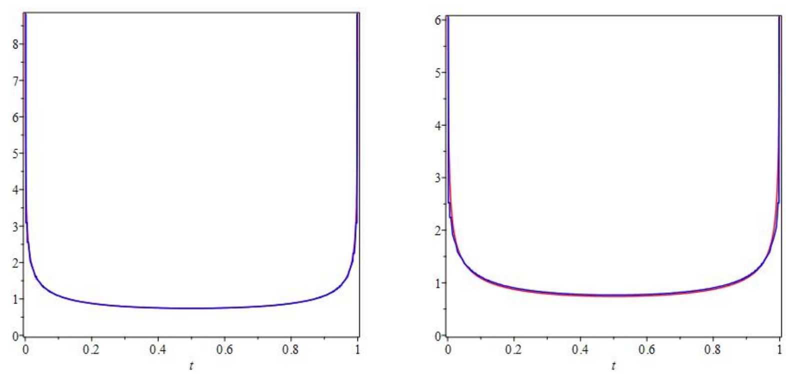



Figure 26: $N = 250$, $\alpha = 0.25$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

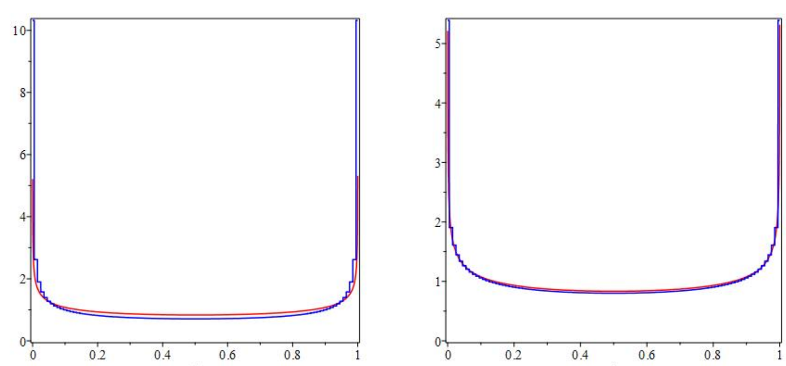



Figure 27: $N = 100$, $\alpha = 0.5$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

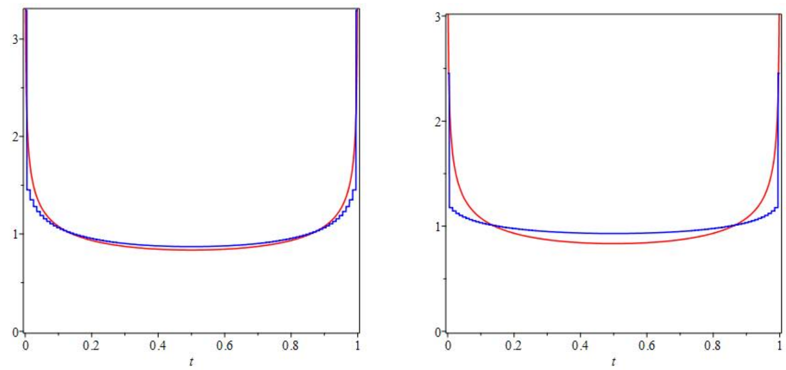


Figure 28: $N = 100$, $\alpha = 0.5$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

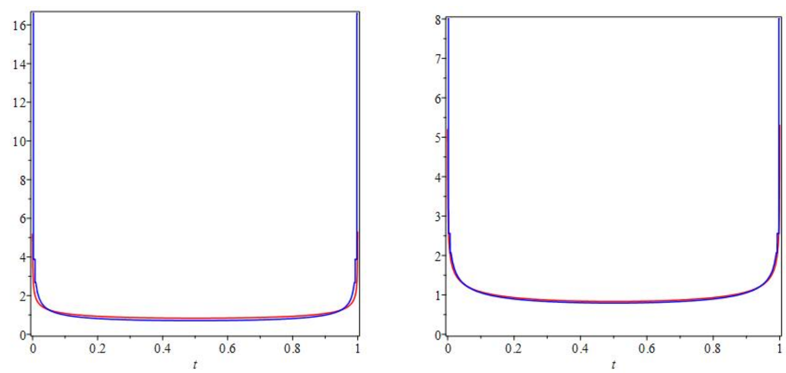


Figure 29: $N = 200$, $\alpha = 0.5$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

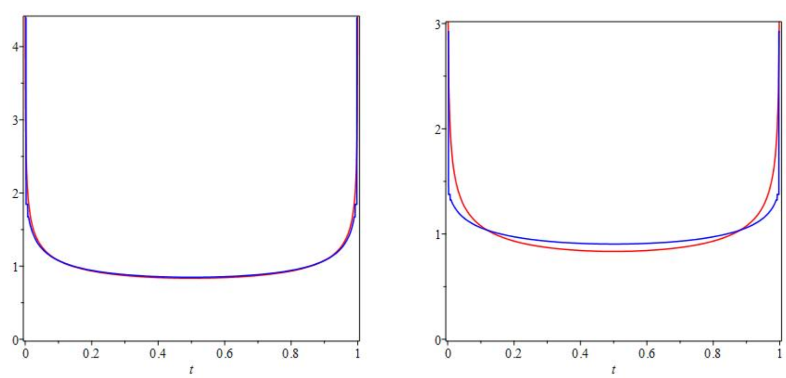


Figure 30: $N = 200$, $\alpha = 0.5$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

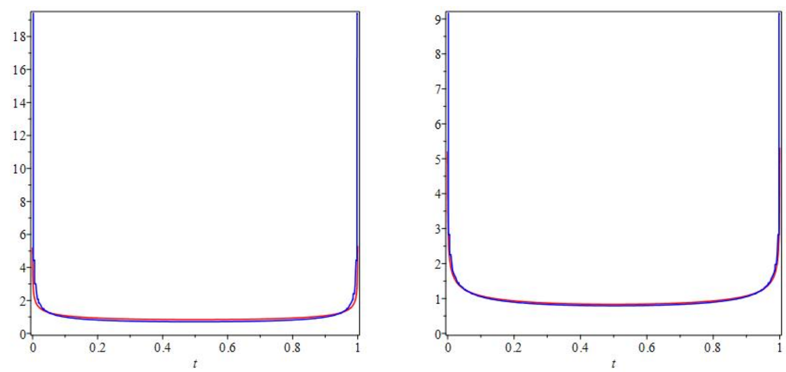


Figure 31: $N = 250$, $\alpha = 0.5$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

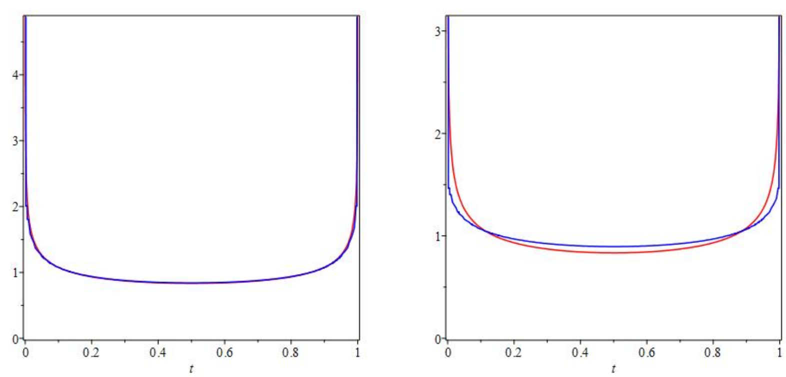


Figure 32: $N = 250$, $\alpha = 0.5$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

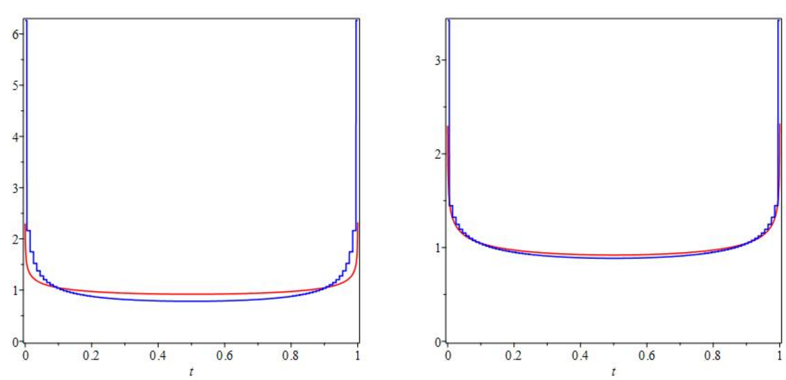


Figure 33: $N = 100$, $\alpha = 0.75$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

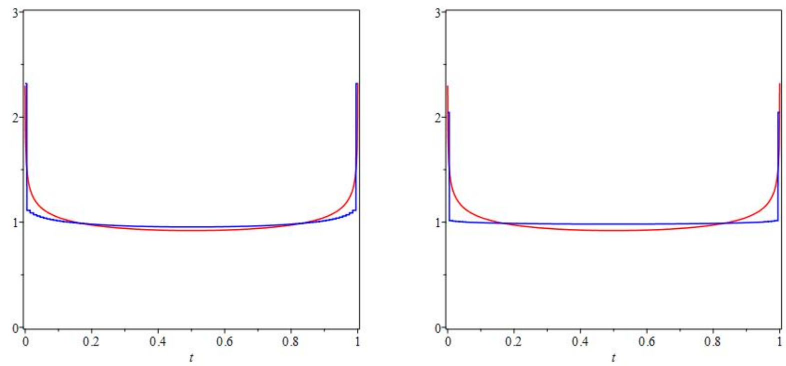



Figure 34: $N = 100$, $\alpha = 0.75$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

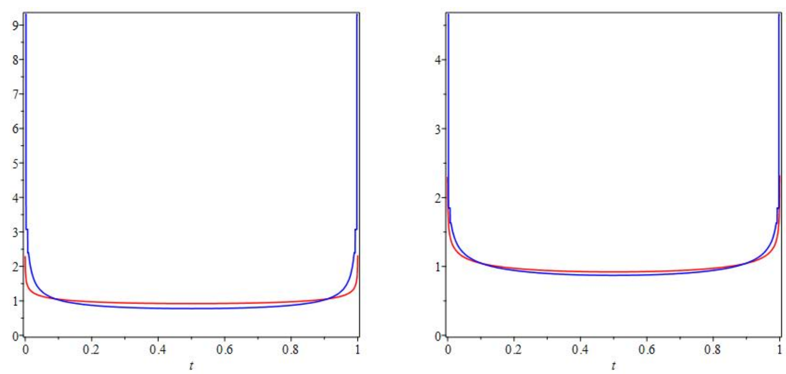



Figure 35: $N = 200$, $\alpha = 0.75$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

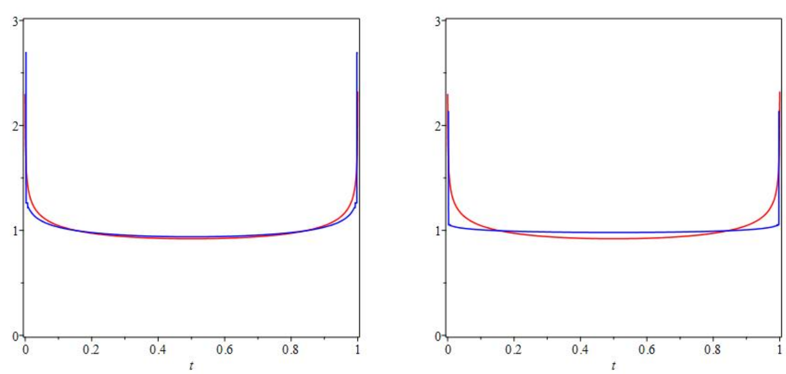



Figure 36: $N = 200$, $\alpha = 0.75$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

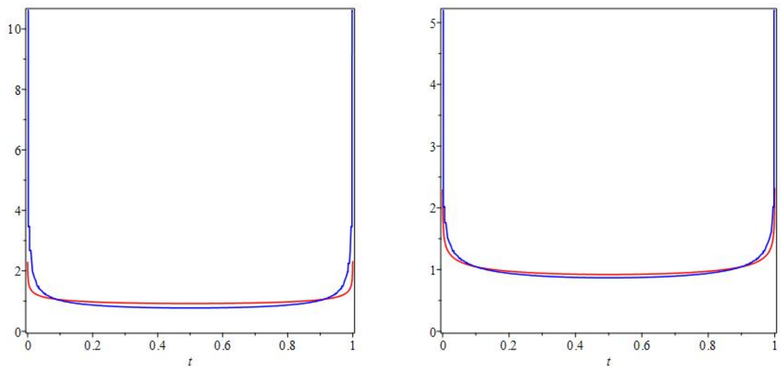


Figure 37: $N = 250$, $\alpha = 0.75$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

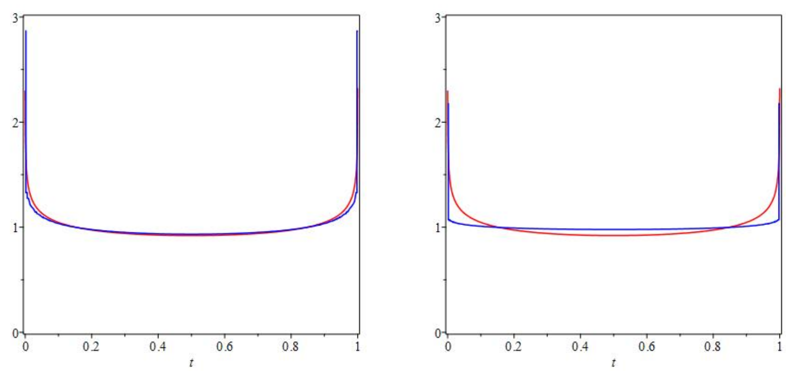


Figure 38: $N = 250$, $\alpha = 0.75$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

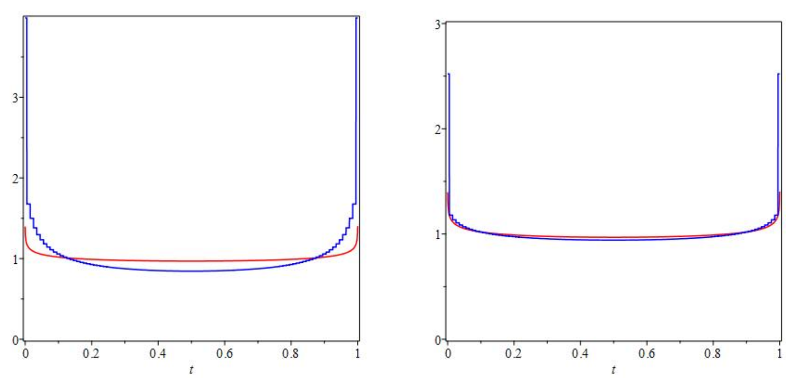


Figure 39: $N = 100$, $\alpha = 0.9$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

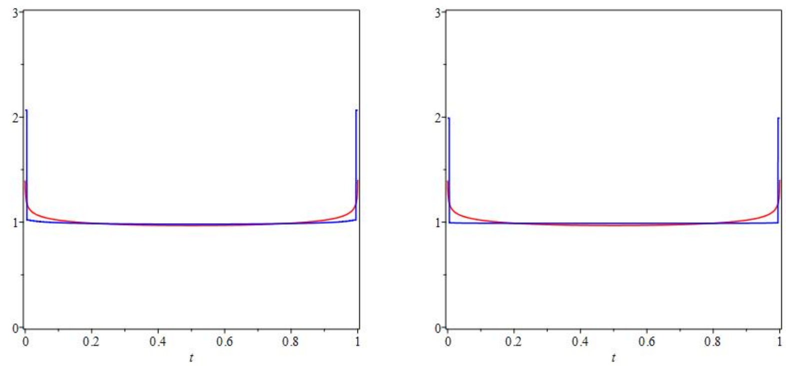



Figure 40: $N = 100$, $\alpha = 0.9$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

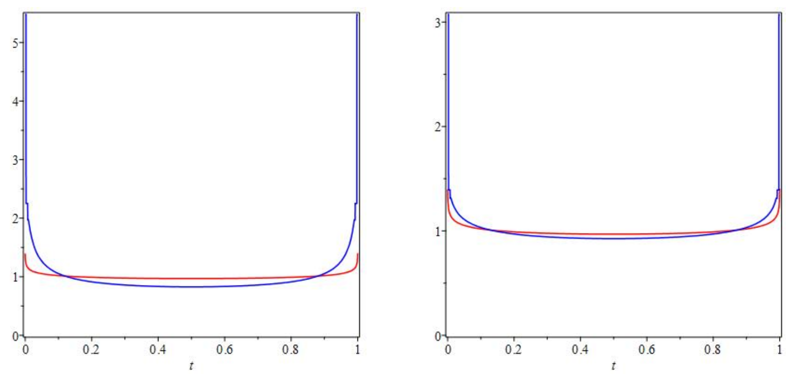



Figure 41: $N = 200$, $\alpha = 0.9$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

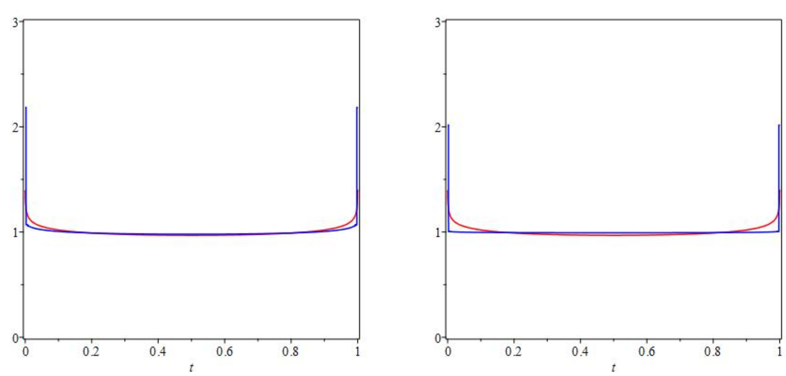



Figure 42: $N = 200$, $\alpha = 0.9$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

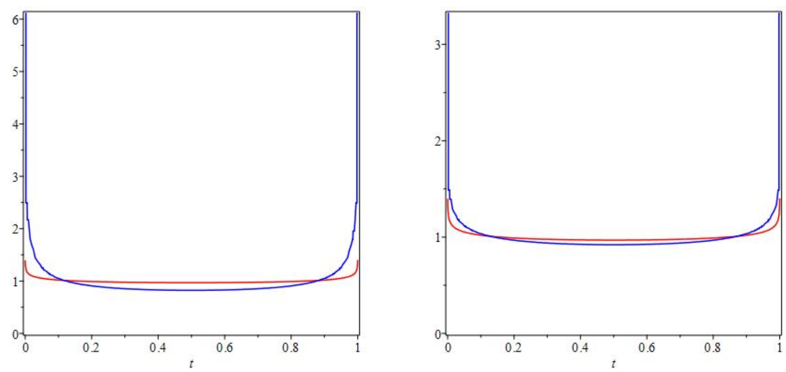



Figure 43: $N = 250$, $\alpha = 0.9$. *Left:* $\epsilon = 0.1$. *Right:* $\epsilon = 0.01$.

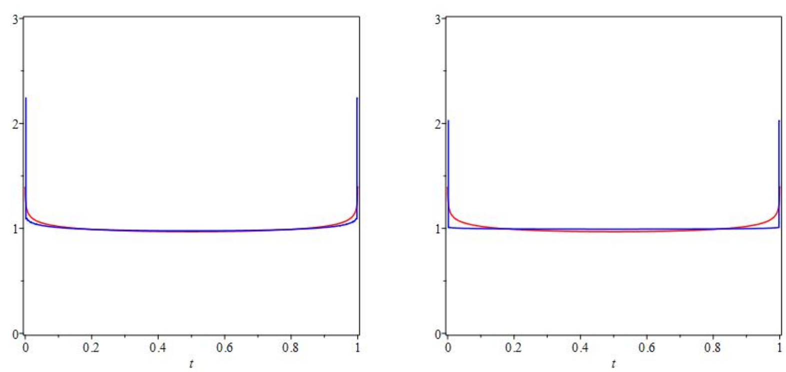



Figure 44: $N = 250$, $\alpha = 0.9$. *Left:* $\epsilon = 0.001$. *Right:* $\epsilon = 0.0001$.

The following figures display the potentials $\mathcal{P}_\mu(x) := \int_0^1 f(x-y)\, p_N \, dy$ ($x \in [0,1]$) of the approximate minimising measures constructed in examples 1, 2 and 3 of Section 4.6. In each case, we plot the expression in (77) at 200 equally spaced points in $[0,1]$. Note, $N = 250$ and $\epsilon = 0.001$, unless stated otherwise.

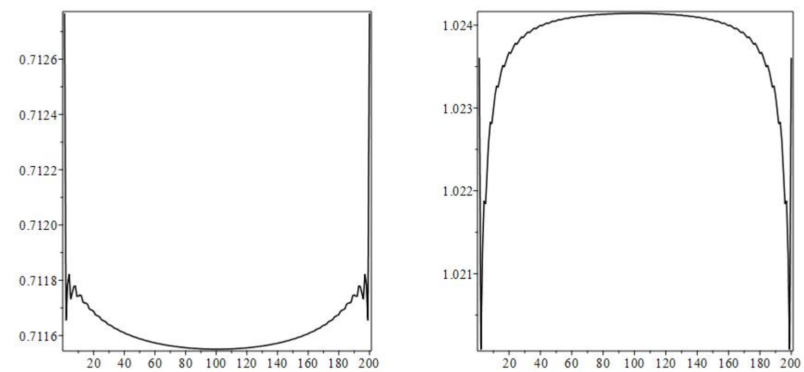



Figure 45: $f_\alpha = |\cdot|^{-\alpha}/(1+|\cdot|)^{2-\alpha}$. *Left:* $\alpha = 0.1$. *Right:* $\alpha = 0.25$.

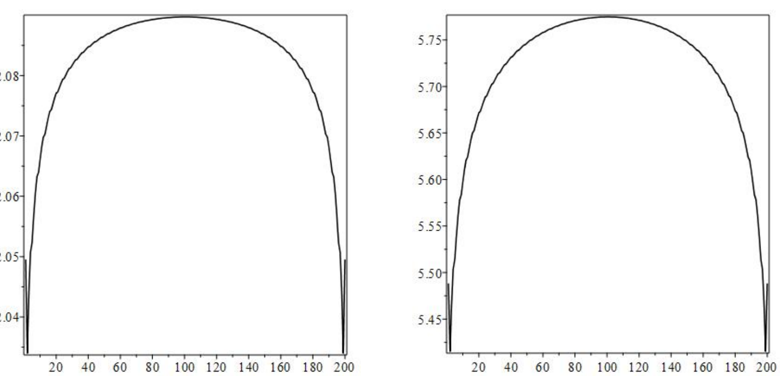



Figure 46: $f_\alpha = |\cdot|^{-\alpha}/(1+|\cdot|)^{2-\alpha}$. *Left:* $\alpha = 0.5$. *Right:* $\alpha = 0.75$.

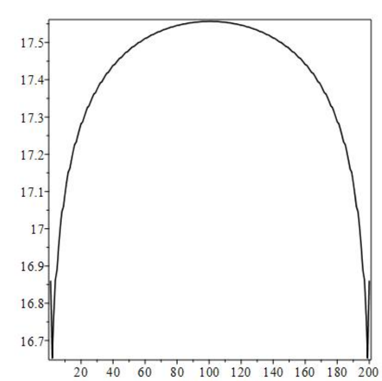



Figure 47: $f_\alpha = |\cdot|^{-\alpha}/(1+|\cdot|)^{2-\alpha}$, $\alpha = 0.9$.

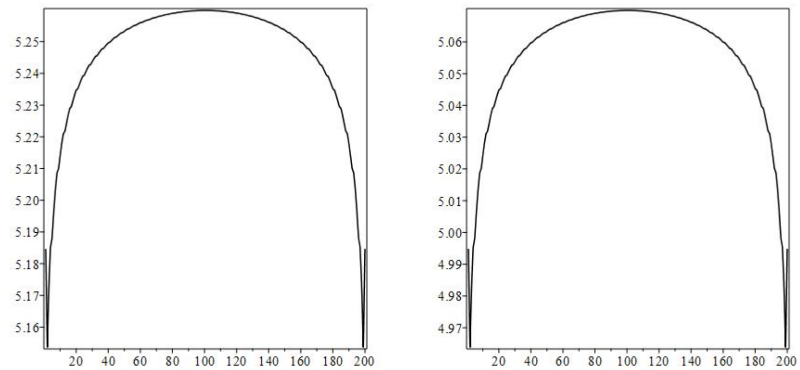



Figure 48: $f_\lambda = \left(1 - (\lambda\sqrt{|\cdot|} - 1)\, e^{-\lambda\sqrt{|\cdot|}}\right)/\sqrt{|\cdot|}$. *Left:* $\lambda = 0.001$. *Right:* $\lambda = 0.1$

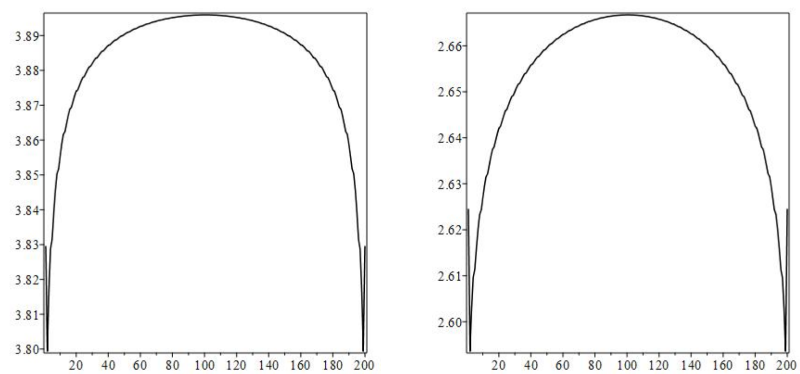



Figure 49: $f_\lambda = \left(1 - (\lambda\sqrt{|\cdot|} - 1)\, e^{-\lambda\sqrt{|\cdot|}}\right)/\sqrt{|\cdot|}$. *Left:* $\lambda = 1$. *Right:* $\lambda = 10$

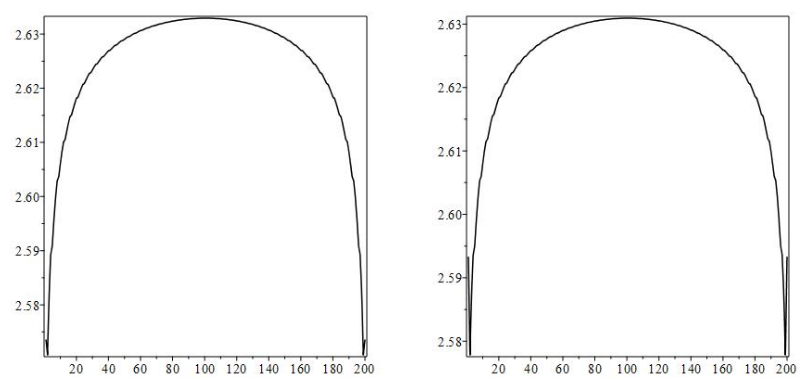



Figure 50: $f_\lambda = \left(1 - (\lambda\sqrt{|\cdot|} - 1)\, e^{-\lambda\sqrt{|\cdot|}}\right)/\sqrt{|\cdot|}$. *Left:* $\lambda = 100$. *Right:* $\lambda = 1000$

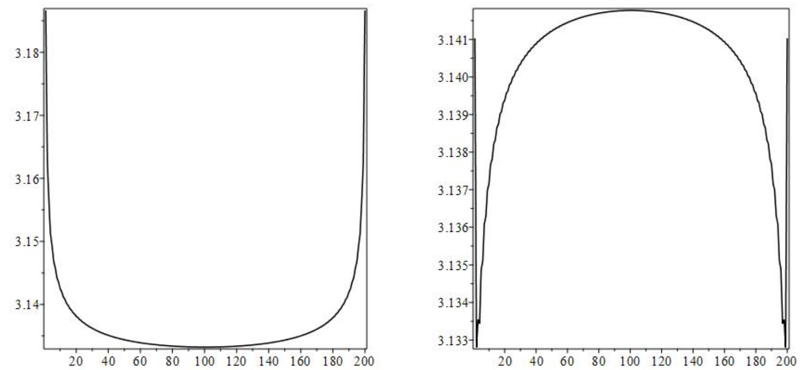



Figure 51: $f = |\cdot|(1+|\cdot|^{-1})^{1+|\cdot|}\log(1+|\cdot|^{-1})$, $N = 100$. *Left:* $\epsilon = 0.01$. *Right:* $\epsilon = 0.001$.

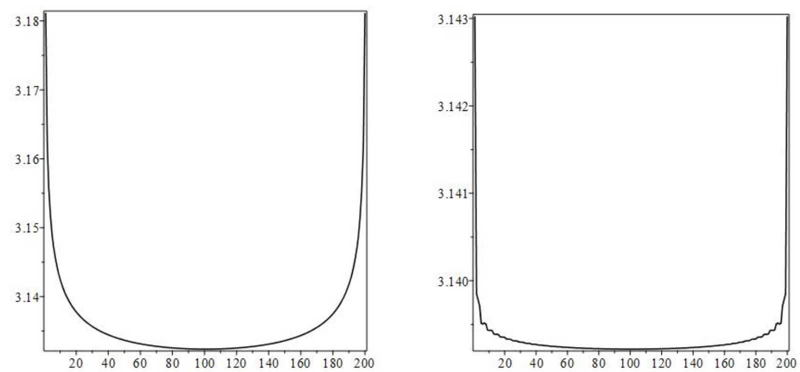



Figure 52: $f = |\cdot|(1+|\cdot|^{-1})^{1+|\cdot|}\log(1+|\cdot|^{-1})$, $N = 250$. *Left:* $\epsilon = 0.01$. *Right:* $\epsilon = 0.001$.

## References


[1] Atkinson K. & Han W. (2009). *Theoretical Numerical Analysis: A Functional Analysis Framework, Vol. 39* (Springer Science & Business Media).

[2] Bekka B., de La Harpe P. & Valette A. (2008). *Kazhdan's property (T), Vol. 11* (Cambridge university press).

[3] Berg C., Christensen J. P. R. & Ressel P. (1984). *Harmonic analysis on semigroups, Vol. 100* (Springer Graduate Texts in Mathematics, New York).

[4] Berg C. & Forst G. (2012). *Potential theory on locally compact Abelian groups, Vol. 87* (Springer Science & Business Media).

[5] Berg C. (2008). Stieltjes-Pick-Bernstein-Schoenberg and their connection to complete monotonicity, *Positive Definite Functions: From Schoenberg to Space-Time Challenges*, 15–45.

[6] Bernstein S. (1914). Sur la définition et les propriétés des fonctions analytiques d'une variable réelle, *Mathematische Annalen*, **75(4)**, 449–468.

[7] Bernstein S. (1929). Sur les fonctions absolument monotones, *Acta Mathematica*, **52(1)**, 1–66.

[8] Björck G. (1956). Distributions of positive mass, which maximize a certain generalized energy integral, *Arkiv för matematik*, **3(3)**, 255–269.

[9] Bochner S. (1932). Vorlesungen über Fouriersche Integerale, *Akademische Verlagsgesellschaft Leipzig*.

[10] Bochner S. (1955). Harmonic analysis and probability theory, *Berkeley and Los Angeles*.

[11] Bochner S. & Chandrasekharan K. (1950). *Fourier Transforms (AM-19), Vol. 19* (Princeton University Press).

[12] Buescu J., Paixao A. C., Garcia F. & Lourtie I. (2004). Positive-definiteness, integral equations and Fourier transforms, *Journal of Integral Equations and Applications*, 33–52.

[13] Carathéodory C. (1907). Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen, *Mathematische Annalen*, **64(1)**, 95–115.

[14] Cooper J. L. B. (1960). Positive definite functions of a real variable, *Proceedings of the London Mathematical Society* **3(1)**, 53–66.

[15] Cramér H. (1939). On the representation of a function by certain Fourier integrals, *Transactions of the American Mathematical Society*, **46(2)**, 191–201.

[16] Frostman O. (1935). Potentiel d'équilibre et capacité des ensembles, *Doctoral dissertation, Gleerup.*

[17] Hajek J. (1956). Linear estimation of the mean value of a stationary random process with convex correlation function, *Czechoslovak Mathematical Journal*, **6(81)**, 94–117.

[18] Herglotz G. (1911). Über Potenzreihen mit positivem, reellen Teil im Einheitskreis, *Leipziger Berichte, math.-phys.*, **63**, 501–511.

[19] Kato T. (2013). *Perturbation theory for linear operators, Vol. 132* (Springer Science & Business Media).

[20] Katznelson Y. (2002). *An introduction to harmonic analysis, third corrected edition* (Cambridge University Press).

[21] Landkof N. S. (1972). *Foundations of modern potential theory* (Springer, Berlin).

[22] Lukacs E. (1960). *Characteristic functions* (Griffinâs statistical monographs and courses, no. 5).

[23] Mathias M. (1923). Über positive Fourier-Integrale, *Mathematische Zeitschrift*, **16(1)**, 103–125.

[24] Mercer J. (1909). Functions of positive and negative type, and their connection with the theory of integral equations, *Philosophical transactions of the royal society of London, Ser. A*, **209**, 415-446.

[25] Niculescu C. & Persson L. E. (2006). *Convex functions and their applications: a contemporary approach* (Springer Science & Business Media).

[26] Phillips T. R. L. & Schmidt K. M. (In Press). *On Unbounded Positive Definite Functions* (Mathematica Pannonica).

[27] Phillips T. R. L., Schmidt K. M. & Zhigljavsky A. (2018). *Extension of the Schoenberg Theorem to Integrally Positive Definite Functions* (Journal of Mathematical Analysis and Applications).

[28] Pólya G. & Szegö G. (1931). Über den transfiniten Durchmesser (Kapazitätskonstante) von ebenen und räumlichen Punktmengen, *Journal für die reine und angewandte Mathematik*, **165**, 4–49.

[29] Rao C. Radhakrishna (1984). Convexity properties of entropy functions and analysis of diversity, *Inequalities in Statistics and Probability*, **5**, 68–77.

[30] Reed M. & Simon B. (1972). *Functional Analysis, Methods of Modern Mathematical Physics, Vol. I* (Academic Press).

[31] Reed M. & Simon B. (1975). *Fourier Analysis, Self-Adjointness, Methods of Modern Mathematical Physics, Vol. II* (Academic Press).

[32] Riesz M. (1938). Intégrales de Riemann-Liouville et potentiels, *Acta Sci. Math. Szeged*, **9(1)**, 1–42.

[33] Saff E. B. (2010). Logarithmic potential theory with applications to approximation theory, *Surv. Approx. Theory*, **5(14)**, 165–200.

[34] Saff E. B. & Totik V. (2013). *Logarithmic potentials with external fields, Vol. 316* (Springer Science & Business Media).

[35] Schilling R., Song R. & Vondracek Z. (2012). *Bernstein functions, theory and applications, 2nd edition* (De Gruyter).

[36] Schmidt K. M. & Zhigljavsky A. (2009). A characterization of the arcsine distribution. *Statistics & Probability Letters* **79(24)**, 2451–2455.

[37] Schmidt K. M. & Zhigljavsky A. (2013). An extremal property of the generalized arcsine distribution. *Metrika* **76(3)**, 347–355.

[38] Schoenberg I. J. (1938). Metric spaces and completely monotone functions, *Annals of Mathematics*, 811–841.

[39] Schoenberg I. J. (1938). Metric spaces and positive definite functions, *Transactions of the American Mathematical Society*, **44(3)**, 522–536.

[40] Schur J. (1911). Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen, *Journal für die reine und angewandte Mathematik*, **140**, 1–28.

[41] Schwartz L. (1951). *Théorie des distributions, tome II* (Hermann & Cie, Paris).

[42] Stein E. M. & Weiss G. L. (1971). *Introduction to Fourier analysis on Euclidean spaces, Vol. 1* (Princeton university press).

[43] Stewart J. (1976). Positive definite functions and generalizations, an historical survey, *Rocky mountain journal of mathematics*, **6(3)**, 409–434.

[44] Toeplitz O. (1911). Über die fourierâsche entwickelung positiver funktionen, *Rendiconti del Circolo Matematico di Palermo*, **32(1)**, 191-192.

[45] Tuck E. O. (2006). On positivity of Fourier transforms, *Bulletin of the Australian Mathematical Society*, **74(01)**, 133–138.

[46] Weidmann J. (2012). *Linear operators in Hilbert spaces, Vol. 68* (Springer Science & Business Media).

[47] Zhigljavsky A., Dette H. & Pepelyshev A. (2010). A new approach to optimal design for linear models with correlated observations. *Journal of the American Statistical Association* **105(491)**, 1093–1103.