TECHNISCHE UNIVERSITÄT MÜNCHEN
FAKULTÄT FÜR MATHEMATIK
Angewandte Numerische Analysis

LOCALIZED KERNELS AND SUPER-RESOLUTION ON SPECIAL MANIFOLDS

Kristof Schröder

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:	Prof. Dr. Michael M. Wolf
Prüfer der Dissertation:	1. Prof. Dr. Massimo Fornasier
	2. Prof. Dr. Stefan Kunis (Universität Osnabrück)
	3. Prof. Dr. Armin Iske (Universität Hamburg)

Die Dissertation wurde am 10.11.2017 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 28.02.2018 angenommen.

Abstract

In this thesis, we consider the problem of super-resolution on certain manifolds. Broadly speaking, super-resolution aims to recover higher resolution information from low resolution measurements. In this context, the given information has to be understood in the regime of analyzing signals with respect to the eigenfunctions of the Laplace-Beltrami operator on the manifold. Especially, the two manifolds $SO(3)$ of rotation matrices in three dimensions and the two-dimensional Euclidean sphere $\mathbb{S}^2$ are of interest to us. On the one hand, these naturally appear in applications and on the other hand, these considerations generalize the previous work on super-resolution, which is mainly situated on the one-dimensional torus $\mathbb{T}$, to more complex geometries. More concrete, the problem consists of recovering a weighted sum of Dirac measures from its low-frequency information only. On the rotation group, the low frequency information are moments of the sought measure with respect to Wigner D-functions, whereas on the sphere the given moments are with respect to spherical harmonics. We investigate the recovery of the sought measure using a total variation minimization approach. Regarding the theoretical aspects, we provide recovery guarantees of a discrete signed measure both on the rotation group and the sphere in terms of the separation distance of the support of the sought measure. In addition, we give error estimates for the recovery on the rotation group in the presence of noise in the given data. The main ingredients for the theoretical aspects are localization estimates for interpolation kernels and their derivatives. We make numerical considerations regarding the recovery on the rotation group and investigate two different recovery algorithms, one that utilizes a semi-definite relaxation of an infinite-dimensional optimization problem and one that builds on an a priori discretization.

Zusammenfassung

In dieser Arbeit untersuchen wir die Problemstellung der Superauflösung auf speziellen Mannigfaltigkeiten. Grob gesprochen zielt die Superauflösung darauf ab höher auflösende Informationen aus niedrig auflösenden zu rekonstruieren. In diesem Kontext verstehen wir Informationen als Analyse von Signalen mittels Eigenfunktionen des Laplace-Operators auf der Mannigfaltigkeit. Insbesondere die Mannigfaltigkeit $SO(3)$ der Rotationsmatrizen in drei Dimensionen und die zweidimensionale Euklidische Sphäre $\mathbb{S}^2$ stehen im Fokus dieser Arbeit, da diese Manigfaltigkeiten zum einen in natürlicher Weise in den Anwendungen von Bedeutung sind und zum anderen die Betrachtung dieser die bisherigen Arbeiten zur Superauflösung, die sich zum größten Teil mit dem eindimensionalen Torus befassen, auf kompliziertere Geometrien verallgemeinert. Genauer betrachtet besteht das Problem der Superauflösung in der Rekonstruktion einer gewichteten Summe von Dirac-Maßen nur mithilfe der niedrigfrequenten Informationen. Auf der Rotationsgruppe bestehen diese niedrigfrequenten Informationen aus Momenten des gesuchten Maßes bezüglich der sogenannten Wigner D-Funktionen eines bestimmten Grades. Auf der Sphäre hingegen sind die Momente des gesuchten Maßes gegeben bezüglich der Kugelflächenfunktionen. Wir untersuchen die Möglichkeit das gesuchte Maß über die Minimierung der Totalvariation zu rekonstruieren. Bezüglich der theoretischen Aspekte betrachten wir Rekonstruktionsgarantien abhängig von der Separationsdistanz des Trägers des gesuchten Maßes sowohl auf der Rotationsgruppe als auch auf der Sphäre. Für den Fall der Rotationsgruppe geben wir in der Situation von gestörten Daten Fehlerabschätzungen an. Das Fundament für diese theoretischen Überlegungen bilden Lokalisationsabschätzungen für Interpolationskerne und deren Ableitungen. Wir diskutieren zwei unterschiedliche Algorithmen für die Rekonstruktion eines diskreten Maßes auf der Rotationsgruppe, wobei der eine auf der semidefiniten Relaxation eines unendlichdimensionalen Optimierungsproblems beruht und der andere eine a priori Diskretisierung verwendet.

Acknowledgement

Writing this thesis, would not have been possible without the help and support of many. I would like to take the opportunity to express my deepest gratitude to them.

First of all, my sincere thanks go to my doctoral advisor Massimo Fornasier. His confidence in me was a constant support and I am grateful, that he gave me the opportunity to do my doctoral studies at his chair.

I wish to express my unbounded gratitude to my advisor at the Helmholtz Center Munich, Frank Filbir. It is not just that he accompanied me through all the time and helped me to develop my mathematical interests, but was also a great source of motivation, which is every bit as important. I would like to thank him for his trust and support and all that I learned from him.

Moreover, I would like to thank Stefan Kunis and Armin Iske for fruitful discussions and serving as referees for this thesis. My special thanks got to Hrushikesh Mhaskar, who agreed to be part of the thesis committee at the Helena graduate school. I appreciate the long and insightful discussions during my stay in California and his visits in Munich.

Working on this thesis would have been only half as much fun without all my colleagues at the Helmholtz Center Munich and the chair for Applied Numerical Analysis at the Technical University Munich. In particular, I want to thank Felix Krahmer, Christian Kümmerle, Juliane Sigl and Nada Sissouno for organizing the workshop WDI^2 and Wolfgang zu Castell, Sam Maurus and Hannah Schrenk for the nice and funny discussions.

Last but not least, I am greatly indebted to my family, my parents Kristina and Karsten, my brothers Kai and Konstantin, for all the support. Furthermore, I thank all my friends for having fun and a great time. Especially but not exclusively, I want to thank Tamara and Ralf for their advice, support and all the great and hilarious moments we share.

Financial support by the Helmholtz Center Munich and by Deutsche Forschungsgemeinschaft Grant FI 883/3-2 is gratefully acknowledged.

Contents

Introduction

The meaning of the expression *super-resolution* is manifold. Depending on the field of studies, quite different phenomena and techniques are subsumed under this term. In optics, the phrase super-resolution denotes instrumental techniques to overcome the *diffraction limit* of an optical system, [Lindberg, 2012]. Especially, the development of *super-resolution fluorescence microscopy* makes it possible to observe biological structure beyond the diffraction limit, see e.g. [Huang et al., 2009], [Schermelleh et al., 2010] and [Cremer and Masters, 2013].

In contrast to this instrumental super-resolution, in the field of imaging the process of recovering a high-resolution image from several low-resolution images is also called super-resolution, see e.g. [Park et al., 2003] and [Nasrollahi and Moeslund, 2014] for a good overview. To distinguish these algorithmic techniques from the instrumental techniques, sometimes the term *computational super-resolution* is used, see [Bertero and Boccacci, 2003].

In this thesis, we will use the term super-resolution to describe the recovery of a *spatially highly resolved signal* from its *coarse scale information only*. To make this statement more precise, the highly resolved signal is modeled by a weighted sum of Dirac measures, i.e.

$$\mu^\star = \sum_j c_j \delta_{x_j}.$$

The coarse scale information is a measurement, modeled as a linear mapping A^*, such that $A^*\mu^\star$ is an approximation to the measure $\mu^\star$, but this approximation may not point out the locations x_i of the support of $\mu^\star$. One can think of the information $A^*\mu^\star$ as a smooth function, which approximates the measure $\mu^\star$ so badly, that the unknown locations x_i of the support of the measure are not identifiable anymore. The process of super-resolution aims to recover the unknown parameters x_i and c_i from this coarse scale information.

To make this abstract problem more concrete, consider a 2π-periodic signal of the form

$$f(t) = \sum_{j=1}^{M} c_j \mathrm{e}^{-ix_j t}, \quad t \in \mathbb{R},$$

with unknown parameters $c_j \in \mathbb{C}, x_j \in \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ and $M \in \mathbb{N}$, i.e. a non-harmonic Fourier expansion, see e.g. [Young, 2001]. What we can access are the samples

$$f(k) = \sum_{j=1}^{M} c_j \mathrm{e}^{-ix_j k}, \quad k = -N, \dots, N, \tag{1}$$

which we want to use to recover the unknown parameters. From a statistical point of view, this problem is also known as *spectral line estimation*, see [Tang et al., 2015]. This problem of detecting *hidden frequencies* was first considered by G.R de Prony in 1795. He proposed a method to extract the unknown frequencies as the simple roots of a certain polynomial, the Prony polynomial, whose coefficients can be computed from the samples $f(k)$ by solving a system of equations. Nevertheless, the locations of the roots are sensitive to perturbation on the samples, which makes *Prony's method* unstable in the presence of noise. An additional restriction of Prony's method is given by the fact, that one has to know the number M of hidden frequencies, or at least a good estimate of it.

Recently, Prony's method is undergoing a revival in order to stabilize and generalize the method in various directions. To name only a few, stabilization approaches are ESPRIT [Roy and Kailath, 1989], MUSIC [Schmidt, 1986], Approximate Prony methods [Potts and Tasche, 2010], [Potts and Tasche, 2011] and [Potts and Tasche, 2013], the Matrix pencil method [Hua and Sarkar, 1990] and the use of orthogonal polynomials on the unit circle (OPUC) [Filbir et al., 2012]. We point out, that this is only a small selection and we refer to the book [Stoica et al., 2005] and the thesis [Peter, 2013] for a detailed overview of existing work in this direction. Generalization to higher dimensions using projections can be found in [Plonka and Wischerhoff, 2013], [Diederichs and Iske, 2015], and [Wischerhoff and Plonka, 2016]. A generalization to higher dimensions, revealing the algebraic structure of Prony's method can be found in [Kunis et al., 2016b].

To relate the problem of detecting hidden frequencies to a measure recovery problem, observe that the samples in equation (1) can also be interpreted in the following way. Consider the measure

$$\mu^\star = \sum_{j=1}^{M} c_j \delta_{x_j}. \tag{2}$$

Observe, that for the expressions in equation (1) we have

$$f(k) = \widehat{\mu^\star}(k), \quad k = -N, \dots, N, \tag{3}$$

where $\widehat{\mu^\star}(k) = \sum_{j=1}^{M} c_j \mathrm{e}^{-ix_j k}$ denotes the k-th Fourier moment of the measure. Thus, we can observe the first $2N+1$ Fourier moments of $\mu^\star$. In other words, we can only access a band-limited approximation of $\mu^\star$ of the form

$$\sum_{k=-N}^{N} \widehat{\mu^\star}(k)\mathrm{e}^{ikt}.$$

In this approximation the 'peaks' of the Dirac measures are 'smoothed out' and their locations are not clear any more.

Very recently, several authors proposed a variational recovery method, that does not need to know the number of support points of the measure beforehand, see [de Castro and Gamboa, 2012], [Candés and Fernandez-Granda, 2014]. They proposed to minimize the *total variation* over the space of regular Borel measures, given the low-frequency information in (3). It turns out, that the total variation norm induces *discreteness* of the measure, which solves the restricted minimization problem. Thus, minimizing the total variation norm over the space of measure can be seen as having a *sparsifying property* on the support of the measure, analog to the minimization of the ℓ^1 norm on a finite space. In the following, we will shortly describe, what we mean by sparsifying property. *Beurling* considered the following extrapolation problem, see [Beurling, 1989a], [Beurling, 1989b]. Given a complex measure on the real line $\mathbb{R}$, we know its *Fourier-Stieltjes transform* only on a subset $[-\lambda, \lambda] \subset \mathbb{R}$. He considered to extrapolate from this subset by taking the Fourier-Stieltjes transform of a measure, whose transform agrees with the given data on $[-\lambda, \lambda]$ and that minimizes the *total variation norm*. His observation is, that under certain conditions such a measure always exists, is unique and is a *discrete* measure, regardless of whether the data generating measure itself is discrete or not. Hence, minimizing the total variation with respect to the given data results in a discrete measure. He called this measure *minimal extrapolation*. The authors of [Benedetto and Li, 2016] considered this extrapolation problem for measures on the torus $\mathbb{T}$ and observed the same behavior, i.e. under certain conditions the minimal extrapolation is always a discrete measure.

The natural question, that rises is under which conditions is the minimal extrapolation of a discrete measure the measure itself? In other words, when is the measure $\mu^\star$, given in (2), the unique solution of the minimization problem

$$\min_{\mu \in \mathcal{M}(\mathbb{T},\mathbb{C})} \|\mu\|_{TV}, \quad \text{subject to} \quad \hat{\mu}(k) = \widehat{\mu^\star}(k), \quad \text{for } |k| \leq N. \tag{TP}$$

The authors of [de Castro and Gamboa, 2012] and [Candés and Fernandez-Granda, 2014] investigate this question using measure theoretic techniques. A *sufficient* criteria for $\mu^\star$ being the unique solution is the existence of so called *dual certificates*. In more detail, consider the support points $\{x_j\}_{j=1}^M$ given in the definition (2) of the measure $\mu^\star$. If for each sign sequence $\{u_j\}_{j=1}^M \subset \mathbb{C}$, i.e. $|u_j| = 1$, there is a trigonometric polynomial of degree N, the so-called *dual certificate*, that interpolates the sign u_j at the location x_j and is strictly smaller than one in absolute value everywhere else, then $\mu^\star$ is the unique solution of the minimization problem (TP). This somehow technical condition becomes more clear in view of the convexity of the problem (TP), which we will see later on. Understandably, using this technique, the points $\{x_j\}_{j=1}^M$ cannot become arbitrarily close, since the existence of dual certificates would conflict with Bernstein's inequality for trigonometric polynomials. In [Candés and Fernandez-Granda, 2014], the authors show, that a minimal separation of $\frac{4\pi}{N}$ is sufficient for the existence of dual certificates and thus $\mu^\star$ being the unique solution of the total variation minimization. The method of proof relies on an explicit construction of a dual certificate by solving a Hermite interpolation problem.

Apart from the theoretical considerations, the optimization problem (TP) is an *infinite-dimensional* problem and therefore not feasible directly. The authors of [Candés and Fernandez-Granda, 2014] propose to solve the convex *dual problem*. In the dual problem, the minimization takes place on the space of trigonometric polynomials of a fixed degree and is therefore finite-dimensional, but the constraints of the minimization are infinite-dimensional in the form of a supremum norm bound on the trigonometric polynomial. Nevertheless, this supremum norm bound on the trigonometric polynomial can be equally cast as a *semi-definite* constraint by introducing an auxiliary matrix variable, based on the representation of non-negative trigonometric polynomials as *sum of squares*. Using this, the dual problem is equivalent to a finite-dimensional semi-definite program and can be solved by standard solvers for convex problems. The solution of the convex problem is a trigonometric polynomial that acts as a dual certificate, i.e. it interpolates the sign of the sought measure at its support points and is strictly smaller than one in absolute value everywhere else. Thus, the support of the sought measure can be identified as those points, where the solution of the dual problem approaches one in absolute value.

Next to the fundamental papers [Candés and Fernandez-Granda, 2014], [de Castro and Gamboa, 2012], the trigonometric super-resolution problem gained a lot attention very recently. Considerations regarding recovery from trigonometric moments corrupted by noise is considered in [Candés and Fernandez-Granda, 2013], [Tang et al., 2015], [Duval and Peyré, 2015], [Li and Tang, 2016], [Fernandez-Granda, 2016] and [Boyer et al., 2017]. Also, the restriction to the case of positive measures gained attention, see e.g. [Morgenshtern and Candés, 2014], [Denoyelle et al., 2015b] and [Denoyelle et al., 2015a]. The generalization to higher dimensions is considered in [Xu et al., 2014]. Numerical treatment of the problem is the content of [Duval and Peyré, 2015] and [Duval and Peyré, 2016].

Beside the fact that this variational recovery approach does not need the number of unknown points as prior information, an advantage is the adaptability to more general settings than recovery on the torus. Indeed, one could consider the problem of recovery of a weighted sum of Dirac measures from its moments with respect to a systems of functions in a quite general setting. Concentrating on the problem of super-resolution, i.e. recovering of highly resolved signals from coarse scale information, a straightforward generalization would be to consider signals on a compact smooth Riemannian manifold and moments with respect to the eigenfunctions of the first few eigenvalues of the Laplace-Beltrami operator on the manifold. Nevertheless, considering the problem in this generality has the drawback, that the eigenfunctions are not known explicitly in most cases, which makes it more difficult to construct the dual certificates. In this thesis, we concentrate on two concrete examples, which are also interesting from the viewpoint of applications, the rotation group $SO(3)$ and the two-dimensional Euclidean sphere $\mathbb{S}^2$.

The group $SO(3)$ of all rotation matrices in dimension three plays a crucial role in various applications ranging from crystallographic texture analysis, see [Bunge, 1982], [Hielscher et al., 2008], [v.d. Boogart et al., 2007], [Schaeben and v.d. Boogart, 2003], over the calculation of magnetic resonance spectra [Stevensson and Edén, 2011] to applications in biology such as protein-protein docking, see [Castrillon-Candas et al., 2005], [Bajaj et al., 2013], [Kovacs et al., 2003]. For a good overview regarding applications see also [Chirikjian and Kyatkin, 2000]. Signals or functions on the rotation group $SO(3)$ are often analyzed with respect to a harmonic basis arising from representation theory of the group, the so called *Wigner-D functions*. Since these functions are also eigenfunctions of the Laplace operator on the manifold $SO(3)$, they can be regarded as a natural analog to Fourier series in the case of the torus group. Thus, the super-resolution problem on the rotation group corresponds to the recovery of a weighted sum of Dirac measures from its low degree approximation with respect to those Wigner D-functions. In this thesis, we investigate the recovery from Wigner D-moments using a total variation minimization approach. We analyze the problem using the concept of dual certificates as described before. We adapt the construction of a dual certificate by solving a Hermite interpolation problem to signals on the rotation group $SO(3)$ and provide sufficient recovery guarantees in terms of the separation distance of the support points of the sought measure. The crucial ingredient for this are *localization estimates* for interpolation kernels and their derivatives on the rotation group. We also consider the numerical solution of the involved optimization problems using two different approaches.

The second example, we consider in this thesis, corresponds to the two-dimensional Euclidean sphere $\mathbb{S}^2$. In this geometry, the involved harmonic basis functions are known as *spherical harmonics*. Applications are ranging from acoustic source detection [Teutsch and Kellermann, 2006] over astrophysics [Vielva et al., 2003] to magnetic resonance imaging [Deslauriers-Gauthier and Marziliano, 2012]. The recovery of weighted sums of Dirac measures on the sphere from moments with respect to spherical harmonics has been considered using different approaches such as *sampling at finite rate of innovation* [Deslauriers-Gauthier and Marziliano, 2013] and Prony like methods [Kunis et al., 2016a]. Recovery using a total variation minimization approach was first considered in [Bendory et al., 2015a]. Although the authors analyze the recovery using dual certificates, there are gaps in the construction of those. In this thesis, we close these gaps and provide explicit reconstruction guarantees in terms of the separation distance of the support points of the sought measure.

Contribution

In this thesis, we consider the super-resolution problem on the rotation group and the sphere in the context of recovery using a total variation minimization. We contribute in the following aspects. On the rotation group, we provide an explicit recovery guarantee of a discrete signed measure in terms of the separation distance of its support. The guarantee builds on the construction of dual certificates, which we approach using a Hermite interpolation. Fundamental for the involved bounds are localization estimates for certain interpolation kernels and their derivatives. We construct interpolation kernels from weights, that are generated by sampling certain B-splines. These kernels allow for bounds with explicit constants, which is necessary in the construction of the dual certificates. Next to the recovery guarantee, we provide error estimates in the presence of noise. Beside the theoretical aspects, we investigate two different recovery algorithms. Whereas the first approach is build on a semi-definite relaxation of the dual problem and does not need an a priori discretization, the second corresponds to recovery on a predefined grid.

On the sphere, we follow the meta-scheme of constructing dual certificates as the solution of Hermite interpolation problems. We provide new bounds that involve derivatives of the Jackson kernel, which are necessary for the construction. Building on this, similar to that on the rotation group, we provide a recovery guarantee of a discrete signed measure in terms of the separation distance of its support.

Outline of the Thesis

Chapter 1: In the first part of this chapter, we introduce the recovery of a discrete measure from linear measurements in an abstract setting. We formulate the recovery from these measurements using a minimization of the *total variation* over the space of regular Borel measures. The key ingredients for the analysis of this convex optimization problem is the notion of the *null-space property* of an operator and the existence of so called *dual certificates*. The bottom line of this part is, that the existence of dual certificates guarantees the recovery of the sought measure by minimizing the total variation.

The second part of this chapter is meant to introduce the problem of super-resolution on the *rotation group* $SO(3)$ in more detail. We give a short overview on the analysis on the rotation group and introduce the recovery problem from moments with respect to *Wigner D-functions*. In this setting, a dual certificate corresponds to a finite linear combination of Wigner D-functions up to a given degree, which interpolates a given sign on the support of the sought measure and is less than one in absolute value elsewhere. We close the part with a glimpse on the construction of dual certificates in this setting using a *Hermite interpolation*.

Chapter 2: To show the existence of a solution of the proposed Hermite interpolation problem, in this chapter we construct *polynomial interpolation kernels*, i.e. kernels that have a finite expansion with respect to Wigner D-functions. We show pointwise bounds for these kernels and their derivatives, which we call *localization estimates*. These sort of estimates are used to show the invertibility of the matrix arising from the Hermite interpolation problem. To derive suitable bounds on the coefficients of the interpolation problem, we need *explicit constants* in all estimates. Beside the application to the problem of super-resolution on the rotation group, these localization estimates might be of interest on their own.

Chapter 3: Building on the derived estimates from Chapter 2, we show that under a suitable *separation condition* on the support points of the sought measure, it is the unique solution with minimal total variation given the available data. We show, that, if the minimal separation of a point set scales proportional to $1/N$, where N is the degree of the given Wigner D-moments, there is always a dual certificate, which interpolates any given real sign at these points and is strictly smaller than one in absolute value elsewhere. This guarantees the recovery of a signed measure, whose support obeys the prescribed separation condition, from exact data.

In addition to the theoretical recovery guarantee for the noise-free data case, we analyze the case of noisy data. We derive L^∞-error estimates for the super-resolution, seen as a spectral extrapolation problem.

Chapter 4: Whereas the previous chapters correspond to theoretical considerations, we investigate the numerical aspects of the proposed total variation minimization. Due to the infinite-dimensional nature of the optimization problem, it is not tractable directly.

In this chapter, we propose two different strategies to search for a solution of the minimization problem. The first considers the dual problem. This problem can be relaxed to a finite-dimensional *semi-definite program*, building on the *Bounded Real Lemma*. From the solution of the dual problem, we can compute the support of the sought measure. In the second approach, we discretize the problem on a *predefined grid* and solve the resulting finite-dimensional problem. We analyze the convergence of the solutions as the predefined discretization gets finer. For both approaches we provide numerical experiments, which should be understood as proof of principle.

Chapter 5: This chapter covers the super-resolution problem on the *two-dimensional Euclidean sphere* $\mathbb{S}^2$. In contrast to the super-resolution problem on the rotation group, the problem on the sphere has been considered before by different authors. Nevertheless, we provide the first valid proof for recovery guarantees on the sphere.

Chapter 1

Super-Resolution and Exact Recovery

At the beginning of this chapter, we give a short introduction to the problem of super-resolution. Afterwards, in Section 1.1 we formulate the problem in an abstract setting as an exact recovery problem. The second part of this chapter introduces the super-resolution problem on the rotation group $SO(3)$. We briefly state the necessary analysis results connected to the rotation group and discuss dual certificates in this setting.

Broadly speaking, the problem of super-resolution aims to resolve a spatial highly resolved signal, modeled by a sum of Dirac measures, using only its low frequency information. In the spectral domain, this corresponds to an extrapolation of the given spectrum.

The setting we will use for this introduction is a signal on the torus $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, i.e. a 2π-periodic signal, that is analyzed using its Fourier coefficients. More concrete, consider a discrete measure having M support points, i.e.

$$\mu^\star = \sum_{j=1}^{M} c_j \delta_{x_j}, \tag{1.1}$$

where x_j are unknown locations in $[-\pi, \pi]$ and $c_j \in \mathbb{C}$ are unknown amplitudes. What we can observe are the first $2N + 1$ Fourier moments, given by

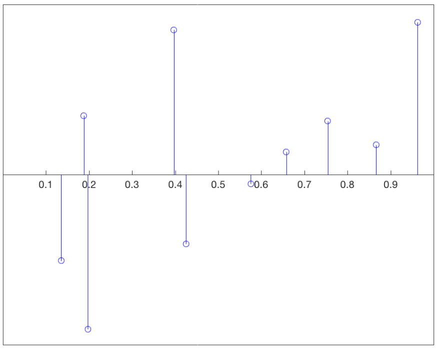
$$y_k = \widehat{\mu}^\star(k) = \int_{-\pi}^{\pi} \mathrm{e}^{-ikx} \, \mathrm{d}\mu^\star(x), \tag{1.2}$$

for $-N \leq k \leq N$. On the spatial side, this means that we observe a convolved version of the signal, given by

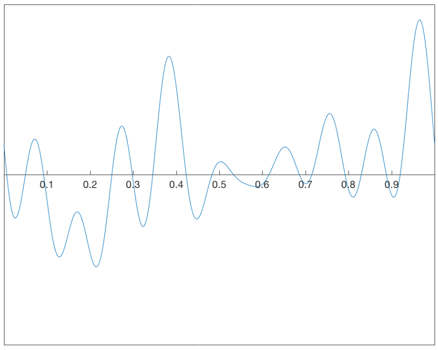
$$(\mu^\star * D_N)(x) = \int_{-\pi}^{\pi} D_N(x - t) \, \mathrm{d}\overline{\mu^\star}(t), \quad x \in [-\pi, \pi],$$

where $D_N(t) = \sum_{k=-N}^{N} \mathrm{e}^{ikt}$ is the Dirichlet kernel. The process of super-resolution aims to deconvolve this signal to recover the measure $\mu^\star$, i.e. the unknown locations x_j and amplitudes c_j. The difficult part of this is the recovery of the support locations x_j, since after finding the support, the amplitudes can be computed by solving a linear system.

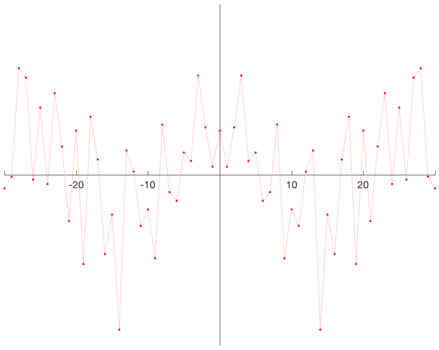
At a first sight, the recovery seems not to be possible, since one theoretically needs all frequency information to recover the support exactly. The problem becomes even more difficult, if the low frequency information is not exact but corrupted by noise, due to the measurement process. In order to still be able to recover the sought measure, one has to incorporate the knowledge of the special structure of the measure $\mu^\star$ into the recovery process. It turns out, that the discreteness of the measure in combination with a separation condition on its support enables the recovery by minimizing the total variation over the measure space with respect to the given Fourier information.



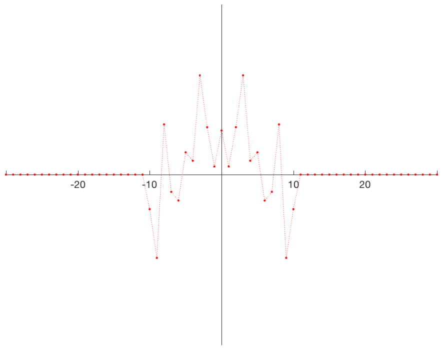
(a) The discrete measure $\mu^\star$.



(b) Its convolved version $(\mu^\star * D_{10})$.



(c) Real part of the Fourier coefficients $\widehat{\mu}^*(k)$.



(d) Available frequency information up to $N = 10$.

Figure 1.1: Illustration of the problem of super-resolution. On the spatial side, it can be seen as a deconvolution problem, whereas on the spectral domain, it is an extrapolation problem.

To clarify the capability of recovering via total variation minimization, we look at the problem of super-resolution as an extension problem in the frequency space. We give a short excursion to the problem of *Beurlings minimal extrapolation*, which was stated by A. Beurling in [Beurling, 1938], [Beurling, 1989a], [Beurling, 1989b]. Denote the space of bounded complex Borel measures on the real line by $\mathcal{M}(\mathbb{R}, \mathbb{C})$, see Appendix B. The Fourier-Stieltjes transform of $\mu^\star \in \mathcal{M}(\mathbb{R}, \mathbb{C})$ is given by

$$\widehat{\mu^\star}(\xi) = \int_{\mathbb{R}} \mathrm{e}^{-i\xi t} \mathrm{d}\mu^\star(t).$$

Suppose, we know $\widehat{\mu^\star}$ on the interval $\Lambda = [-\lambda, \lambda]$. A measure $\mu \in \mathcal{M}(\mathbb{R}, \mathbb{C})$ is called *minimal extrapolation* from Λ, if

$$\begin{aligned} \widehat{\mu}(\xi) &= \widehat{\mu^\star}(\xi), \quad \xi \in \Lambda, \\ \|\mu\|_{TV} &= \inf\{\|\nu\|_{TV} : \nu \in \mathcal{M}(\mathbb{R}, \mathbb{C}), \widehat{\nu}(\xi) = \widehat{\mu^\star}(\xi), \xi \in \Lambda\}. \end{aligned}$$

Beurling asked the question about the existence and uniqueness of a minimal extrapolation. He showed, that if we denote

$$m := \inf\{\|\nu\|_{TV} : \nu \in \mathcal{M}(\mathbb{R}, \mathbb{C}), \widehat{\nu}(\xi) = \widehat{\mu^\star}(\xi), \xi \in \Lambda\},$$

the measure $\mu^\star$ admits a *unique* minimal extrapolation provided that

$$|\widehat{\mu^\star}(\xi)| \neq m, \quad \text{for all } \xi \in \Lambda.$$

Moreover, he showed that this extrapolation is a *discrete measure*, i.e it is of the form

$$\mu = \sum_{j=1}^{\infty} c_j \delta_{x_j},$$

for $(x_j)_{j\in\mathbb{N}} \subset \mathbb{R}$, $(c_j)_{j\in\mathbb{N}} \subset \mathbb{C}$. This is true for all $\mu^\star \in \mathcal{M}(\mathbb{R}, \mathbb{C})$, regardless of whether $\mu^\star$ is a discrete measure itself or not. If one likes, one can call this a *sparsifying* property of the total variation norm. Building on the work of Beurling, Donoho considered the recovery of a discrete measure on $\mathbb{R}$ supported on a grid, see [Donoho, 1992].

Very recently, the authors of [Benedetto and Li, 2016] showed, that this behavior holds true for measures on the Torus $\mathbb{T}$. To be more precise, they showed that for $\mu^\star \in \mathcal{M}(\mathbb{T}, \mathbb{C})$ the minimal extrapolation μ, such that

$$\widehat{\mu}(k) = \widehat{\mu^\star}(k), \quad k \in \Lambda = \{-N, \dots N\},$$
$$\|\mu\|_{TV} = \inf\{\|\nu\|_{TV} : \nu \in \mathcal{M}(\mathbb{T}, \mathbb{C}), \widehat{\nu}(k) = \widehat{\mu^\star}(k), k \in \Lambda\}.$$

exists, is unique and a discrete measure of the form (1.1), if

$$\widehat{\mu^\star}(k) \neq \inf\{\|\nu\|_{TV} : \nu \in \mathcal{M}(\mathbb{T}, \mathbb{C}), \widehat{\nu}(k) = \widehat{\mu^\star}(k), k \in \Lambda\}, \quad \text{for all } k \in \Lambda. \tag{1.3}$$

For the cases, such that (1.3) is not satisfied, $\Lambda \subset \mathbb{Z}$ is a more general subset or the extension problem is on the d-dimensional Torus $\mathbb{T}^d$, we refer to [Benedetto and Li, 2016].

In short, under certain conditions the minimal extrapolation always gives a discrete measure, regardless of whether $\mu^\star$ is itself discrete or not. In the case $\mu^\star$ is known to be discrete, this immediately raises the question of whether the minimal extrapolation is $\mu^\star$ itself? In other words, when is the measure $\mu^\star$ the unique solution of the minimization problem

$$\min_{\mu \in \mathcal{M}(\mathbb{T})} \|\mu\|_{TV}, \quad \text{subject to} \quad \widehat{\mu}(k) = \widehat{\mu^\star}(k), \quad \text{for } |k| \leq N. \tag{TP}$$

In the articles [de Castro and Gamboa, 2012] and [Candés and Fernandez-Granda, 2014] it was proposed to use the minimization (TP) to recover a discrete measure on the torus $\mathbb{T}$ from its low frequency information. Both articles show that the sought measure $\mu^\star$ is a solution of the minimization problem under certain separation conditions on the points in the support of the discrete measure $\mu^\star$. If we denote $\mathcal{X} = \text{supp}(\mu^\star) = \{x_i\}_{i=1}^M$, the separation distance is given by

$$\rho(\mathcal{X}) = \min_{x_i \neq x_j} |x_i - x_j|, \quad x_i, x_j \in \mathcal{X}.$$

Whereas in [de Castro and Gamboa, 2012] it was shown, that under the assumption

$$N \geq \frac{2}{\sqrt{\pi}} \left(\frac{\sqrt{\mathrm{e}}}{\rho(\mathcal{X})} \right)^{5/2+1+\rho(\mathcal{X})}$$

the measure $\mu^\star$ is *a* solution of the minimization problem (TP), the authors of [Candés and Fernandez-Granda, 2014] showed the following.

Theorem 1.1 ([Candés and Fernandez-Granda, 2014], Thm. 1.2)**.** *Suppose the measure* $\mu^\star$ *is given by* (1.1) *and one can observe its Fourier coefficients* $\widehat{\mu^\star}(k)$ *up to order* $N \geq 128$*, given by* (1.2)*. If the support of the measure* $\mathcal{X} = \{x_i\}_{i=1}^M$ *obeys*

$$\rho(\mathcal{X}) \geq \frac{4\pi}{N},$$

then μ^* *is the unique solution of the optimization problem*

$$\min_{\mu \in \mathcal{M}(\mathbb{T})} \|\mu\|_{TV}, \quad \textit{subject to} \quad \widehat{\mu}(k) = \widehat{\mu^\star}(k), \quad \textit{for } |k| \leq N. \tag{TP}$$

The backbone of the proof is the construction of a so called *dual certificate*. This is a trigonometric polynomial of maximal degree N, that interpolates a given sign pattern on the support of the measure $\mu^\star$ and is strictly less than one in absolute value elsewhere. We will provide more details in Section 1.1. Interestingly, a very similar concept already appears in the work of Beurling, where he calls it *extremal function*, see [Beurling, 1989b, pp. 360 − 362].

We give several examples, that the recovery of discrete measures is a general theme. We start with the recovery from low frequency information, which is the analog to the previous example of trigonometric moments. We state this problem for two specific settings, which are the main topic of this thesis, the rotation group $SO(3)$ and the two-dimensional sphere $\mathbb{S}^2$. The finite trigonometric moment problem and both these examples can be understood as the realization of an abstract super-resolution problem.

Example 1.2 (Super-resolution)**.** *Let* X *be a compact smooth Riemannian manifold. In this setting, it is common to analyze functions with respect to the eigenspaces of differential operators, especially the Laplace-Beltrami operator on* X*. Since this operator is compact in the* L^2*-topology, the eigenspaces are dense in* $L^2(X) = L^2(X, \nu)$*, where* ν *is the Riemannian volume measure. This means,*

$$L^2(X) = \mathrm{cl}_{\|\cdot\|_2} \bigcup_{l=1}^{\infty} H_l,$$

where H_l *is the eigenspace to the* $l-$*th eigenvalue. The frequency information is now carried in the ascending spaces*

$$\Pi_N(X) = \mathrm{span} \bigcup_{l=1}^{N} H_l, \quad N = 1, 2 \ldots.$$

The L^2*-projection operator onto the space* $\Pi_N(X)$ *for a fixed* N *can be written in the following way. Choose for each* $l = 1, 2, \ldots$ *an orthonormal basis* $\{\varphi_{l,k}\}_{k=1}^{\dim(H_l)}$ *of* H_l *and set*

$$K_N(x, y) = \sum_{l=1}^{N} \sum_{k=1}^{\dim(H_l)} \varphi_{l,k}(x)\overline{\varphi_{l,k}(y)}.$$

Then the projection operator $\mathcal{S}_N : L^2(X) \to C(X)$ *onto the space* $\Pi_N(X)$ *can be written as*

$$\mathcal{S}_N f(x) = \int_X f(y) K_N(x, y) \mathrm{d}\nu(y).$$

This is the setting of most interest to us, since this corresponds to the reconstruction of point measures from low frequency information in the meaning of analysis of functions with respect to eigenspaces of a differential operator. The super-resolution problem now reads as follows. Given a discrete measure

$$\mu^\star = \sum_{j=1}^{M} c_j \delta_{x_j},$$

with $x_j \in X$ *and* $c_j \in \mathbb{C}$*, recover* $\mu^\star$ *from the low frequency information* $\mathcal{S}_N^* \mu^\star$*.*

In the case $X = \mathbb{T}$ *this resembles the trigonometric setting with*

$$\varphi_{l,l}(x) = \mathrm{e}^{ilx}, \quad l = -N, \dots, N$$

and $K_N(x, y) = D_N(x - y)$ *is the classical Dirichlet kernel.*

In this thesis, we concentrate on two concrete realizations of this abstract super-resolution problem. The first regards to measures on the group $SO(3)$ *of rotations in three dimensions. Here, the eigenspaces of the Laplace-Beltrami operator are spanned by the so called* Wigner D-functions *and the kernel of the projection operator has the form*

$$K_N(x, y) = \sum_{l=0}^{N} (2l + 1) U_{2l} \left(\cos \left(\frac{\omega(x^{-1}y)}{2} \right) \right),$$

where U_n *denotes the* n*-th Chebychev polynomial of the second kind and* $\omega(x^{-1}y)$ *is the rotation angle of the rotation matrix* $x^{-1}y$*. We give more details in Section 1.2. To the best of our knowledge, the super-resolution problem on the rotation group is first considered in this thesis.*

Secondly, we consider measures on the $2-$*Sphere, i.e.* $\mathbb{S}^2 = \{x \in \mathbb{R}^3, \|x\| = 1\}$*. The involved basis functions of the eigenspaces are called* spherical harmonics*. The kernel of the projection operator has the form*

$$K_N(x, y) = \sum_{l=1}^{N} \frac{2l + 1}{4\pi} P_l(x \cdot y),$$

where P_l *is the* l*-th Legendre polynomial. This setting was first considered in [Bendory et al., 2015a], [Bendory et al., 2015b] and [Bendory and Eldar, 2015]. Nevertheless, there are severe gaps in the proofs, which we discuss and close in Chapter 5.*

Example 1.3 (Generalized moment problem)**.** *The recovery of measures from a given set of moments was also considered under other assumptions than discreteness of the sought measure like positivity or absolute continuity with respect to a given prior measure, see e.g. [Gamboa and Gassiat, 1994], [Gamboa and Gassiat, 1996], [Lewis, 1996] and references therein.*

To introduce a moment problem in the context of sparsity assumptions, let $X = [-1, 1]$ *and consider a discrete measure* $\mu^\star = \sum_{j=1}^{M} c_j \delta_{x_j}$*, with* $c_j \in \mathbb{C}, x_j \in [-1, 1], M \in \mathbb{N}$ *and its polynomial moments.*

$$y_k = \int_{-1}^{1} x^k \, \mathrm{d}\mu^\star(x), \quad \textit{for } k = 0, \dots, N.$$

In [Bendory et al., 2014] and [Castro and Mijoule, 2015] these polynomial moment information were used to recover unknown knots of non-uniform spline approximations using total variation minimization. More general, one can consider any system of continuous functions φ_k *and the moments*

$$y_k = \int_{-1}^{1} \varphi_k(x) \, \mathrm{d}\mu^\star(x), \quad \textit{for } k = 0, \dots, N.$$

The generalized moment problem *asks to recover the measure* $\mu^\star$ *from these moments. In the case, that* $\{\varphi_k\}_{k=0}^{\infty}$ *is a Markov system, the recovery of a discrete measure using the total variation norm minimization was analyzed in [de Castro and Gamboa, 2012]. By Markov system, we mean that for each* N *the system* $\{\varphi_k\}_{k=0}^{N}$ *is a Chebychev system, i.e. each non-trivial function in* $\operatorname{span}\{\varphi_k\}_{k=0}^{N}$ *has at most* N *zeros. Nevertheless, the presented theory depends on* φ_k *being a Markov system and thus is restricted to univariate settings by the* Mairhuber-Curtis *Theorem, see [Mairhuber, 1956], which states that there are no Chebyshev systems on higher-dimensional domains.*

For numerical considerations regarding polynomial moments on semi-algebraic sets, also in higher dimensions, see [De Castro et al., 2017].

Example 1.4 (Point-spread deconvolution)**.** *Let $X \subset \mathbb{R}^d$ be a compact set. Suppose, we measure a signal of the form*

$$f(x) = \sum_{i=i}^{M} c_i K(x, x_i)$$

for a given continuous kernel $K : X \times X \to \mathbb{R}$ and would like to recover the unknown parameters $c_i \in C, x_i \in X, M \in \mathbb{N}$. Typically the kernel K is a radial kernel, i.e. it is of the form

$$K(x, y) = \tilde{K}(\varepsilon \|x - y\|_2),$$

where ε is called scaling parameter and $\tilde{K}$ is called point spread function*. Again, this can be understood to recover the discrete measure*

$$\mu^\star = \sum_{j=1}^{M} c_j \delta_{x_j}$$

from the linear measurement.

$$\int_X K(x, y) \mathrm{d}\mu^\star(x).$$

Approaches, using a total variation minimization for recovery from measurements of this kind, can be found in [Duval and Peyré, 2015], [Bendory et al., 2016a], [Bendory et al., 2016b], [Bendory, 2017] and [Bernstein and Fernandez-Granda, 2017].

Building on these examples, in the next section of this chapter we discuss the possibility of recovering a discrete measure from linear measurements in an abstract setting using a total variation minimization approach. It turns out, that the crucial ingredient to analyze this convex minimization problem is the existence of so-called *dual certificates*.

1.1 Exact Recovery in an Abstract Setting

We state the problem of exact recovery in a more abstract way, which was also considered in [Bredies and Pikkarainen, 2013], to incorporate several different measurement situations. Let X be a compact Hausdorff space and $\mathcal{X} \subset X$ a finite subset. The measure is given by

$$\mu^\star = \sum_{x \in \mathcal{X}} c(x) \delta_x, \tag{1.4}$$

with $c(x) \in \mathbb{K}$, where $\mathbb{K}$ is $\mathbb{C}$ or $\mathbb{R}$, for all $x \in \mathcal{X}$. Let H be a Hilbert space and

$$A : H \to C(X, \mathbb{K})$$

be a bounded linear operator. Then the adjoint $A^* : \mathcal{M}(X, \mathbb{K}) \to H$ is weak* to weak continuous and will serve as a model for the measurements of the sought measure, i.e. given the measurements

$$y = A^* \mu^\star,$$

one has to reconstruct the measure $\mu^\star$. For example, in the recovery problem from Fourier coefficients, discussed in the previous section, A is the convolution with the Dirichlet kernel.

To find the measure $\mu^\star$ given the information $A^*\mu^\star$ one can consider the following minimization problem

$$\inf\{|\mathcal{X}| : \mu = \sum_{x \in \mathcal{X}} c(x) \delta_x, \mathcal{X} \text{ finite set}, c(x) \in \mathbb{K}\}, \quad \text{subject to} \quad A^*\mu = y, \tag{GP$_0$}$$

with $y = A^*\mu^\star$, i.e. minimizing the sparsity with respect to the given information provided by the mapping A^*. Even in a complete finite-dimensional setting, i.e. X is finite and $\mathrm{ran}(A^*)$ is finite-dimensional, this problem is in general NP-hard, see e.g. [Rauhut and Foucart, 2013].

We observe that for some $t > 0$ the measure $\mu^\star$ is an element of

$$t \cdot \mathrm{conv}(\mathcal{E}),$$

where $\mathcal{E} = \{s \cdot \delta_x, x \in X, s \in \{-1, 1, i, -i\}\}$ and $\mathrm{conv}(\mathcal{E})$ denotes the convex hull of $\mathcal{E}$. In the finite-dimensional case, i.e. X is finite and $\mathrm{ran}(A^*)$ is finite-dimensional, the authors of [Chandrasekaran et al., 2012] propose to use the convex surrogate

$$\inf\{t : \mu \in t \cdot \mathrm{conv}(\mathcal{E}), t \geq 0\},$$

which leads to the $\|\cdot\|_{\ell_1}$-norm, since $\mathrm{conv}(\mathcal{E})$ is the norm one ball with respect to $\|\cdot\|_{\ell_1}$. This is the classical *Basis pursuit problem*, see [Chen et al., 1998].

In the case X is not finite, one replaces $\mathrm{conv}(\mathcal{E})$ with

$$B_1 = \mathrm{cl}_{\mathrm{w}^\star} \mathrm{conv}(\mathcal{E}).$$

The theorem of Krein-Milman, see e.g. [Simon, 2011], shows that B_1 is indeed the norm one ball with respect to $\|\cdot\|_{TV}$ and

$$\|\mu\|_{TV} = \inf\{t : \mu \in t \cdot B_1, t \geq 0\},$$

being the corresponding convex surrogate. Accordingly, one replaces the optimization problem (GP_0) with the convex relaxation

$$\min_{\mu \in \mathcal{M}(X,\mathbb{K})} \|\mu\|_{TV}, \quad \text{subject to} \quad A^*\mu = y. \tag{GP}$$

Now the question is, under which conditions on the operator A^* and the set $\mathcal{X}$ does the minimization of the convex objective functional leads to the sought measure.

1.1.1 Null Space Property and Dual Certificates

Similarly to the finite-dimensional case, the null space of the mapping A^*, i.e. the set of all measures with $A^*\mu = 0$, plays a crucial role in determining the uniqueness of the minimizers. In this section, we see that the sought measure $\mu^\star$ is the unique minimizer of the problem (GP), if and only if the mapping A^* has the *null space property*. In the following, we state a sufficient criteria for the null space property to hold true, i.e. the existence of so called *dual certificates*, which is a major tool to study the solutions of the convex minimization problem (GP).

Definition 1.5. *[Null-space property] Let $\mathcal{X} \subset X$ be a discrete set. The operator $A^* : \mathcal{M}(X,\mathbb{K}) \to H$ has the* null space property *with respect to $\mathcal{X}$, if for all $\mu \in \ker(A^*) \setminus \{0\}$*

$$\|\mu_{\mathcal{X}}\|_{TV} < \|\mu_{\mathcal{X}^c}\|_{TV}, \tag{NSP}$$

where $\mu = \mu_{\mathcal{X}} + \mu_{\mathcal{X}^c}$ is the Lebesgue decomposition of μ with respect to the measure $\nu = \sum_{b\in\mathcal{X}} \delta_b$.

Theorem 1.6. *Let $\mathcal{X} \subset X$ be a discrete set. Then each $\mu^\star \in \mathcal{M}(X,\mathbb{K})$ with $\operatorname{supp}(\mu^\star) \subseteq \mathcal{X}$ is the unique minimizer of*

$$\min_{\mu\in\mathcal{M}(X,\mathbb{K})} \|\mu\|_{TV}, \quad \textit{subject to} \quad A^*\mu = y,$$

with $y = A^*\mu^\star$, if and only if A^* has the null space property (NSP) *with respect to $\mathcal{X}$.*

Proof. First, we argue that the minimization problem has a solution. For this, set

$$\chi_y(\mu) := \begin{cases} 0, & A^*\mu = y, \\ \infty, & \text{else,} \end{cases}$$

i.e. χ_y is the indicator function of the feasible set of the optimization problem (GP). Hence, the minimization is equivalent to the unconstrained minimization

$$\min_{\mu\in\mathcal{M}(X,\mathbb{K})} J(\mu),$$

with

$$J(\mu) = \|\mu\|_{TV} + \chi_y(\mu). \tag{1.5}$$

Due to the dual representation of the norm $\|\cdot\|_{TV}$, see appendix B, one veryfies that the norm is sequentially lower semicontinuous with respect to the weak*-topology (w*-s.l.s.c.) on $\mathcal{M}(X,\mathbb{K})$. Since the effective domain of χ_y is w*-sequentially closed, χ_y is also w*-s.l.s.c. Therefore, J is w*-s.l.s.c and the existence of a minimizer follows by a compactness argument, see Lemma D.1.

Let $\mu \in \ker(A^*)$, with the Lebesgue decomposition

$$\mu = \mu_{\mathcal{X}} + \mu_{\mathcal{X}^c}.$$

By assumption, $\mu_{\mathcal{X}}$ is the unique minimizer of the optimization problem (GP) with the data given by $y = A^*\mu_{\mathcal{X}}$. Moreover, $A^*\mu_{\mathcal{X}} = -A^*\mu_{\mathcal{X}^c}$ and $\mu_{\mathcal{X}} \neq \mu_{\mathcal{X}^c}$, which yields

$$\|\mu_{\mathcal{X}}\|_{TV} < \|\mu_{\mathcal{X}^c}\|_{TV},$$

since otherwise $\mu_{\mathcal{X}^c}$ would be a minimizer of (GP).

For the opposite direction, assume that $\mu_0 = \mu^\star + \lambda$ is a minimizer and the difference measure λ is non-zero. By assumption we have $\lambda \in \ker(A^*) \setminus \{0\}$ and

$$\|\lambda_{\mathcal{X}}\|_{TV} < \|\lambda_{\mathcal{X}^c}\|_{TV},$$

where $\lambda = \lambda_{\mathcal{X}} + \lambda_{\mathcal{X}^c}$, with $\lambda_{\mathcal{X}^c} \neq 0$, is the Lebesgue decomposition with respect to the measure $\nu = \sum_{x\in\mathcal{X}} \delta_x$. Since $\operatorname{supp}(\mu^\star) \subseteq \mathcal{X}$, the measures $\mu^\star + \lambda_{\mathcal{X}}$ and $\lambda_{\mathcal{X}^c}$ are mutually singular. Consequently

$$\begin{aligned} \|\mu^\star\|_{TV} \geq \|\mu^\star + \lambda\|_{TV} &= \|\mu^\star + \lambda_{\mathcal{X}}\|_{TV} + \|\lambda_{\mathcal{X}^c}\|_{TV}, \\ &\geq \|\mu^\star\|_{TV} - \|\lambda_{\mathcal{X}}\|_{TV} + \|\lambda_{\mathcal{X}^c}\|_{TV} > \|\mu^\star\|_{TV}, \end{aligned}$$

which is a contradiction, meaning $\lambda = 0$ and $\mu^\star$ is the unique minimizer. □

Although the null-space property is an equivalent characterization, it is still hard to check this property. However, it is well known, how to derive a sufficient condition for the null-space property. This condition involves the existence of so-called *dual certificates* or dual interpolating polynomials. This connection has been explicitly exploited in various settings, see e.g. [de Castro and Gamboa, 2012], [Candés and Fernandez-Granda, 2014], [Bendory et al., 2014] and [Bendory et al., 2015a]. For completeness, we will state this condition in our abstract setting, as the proof is purely measure theoretic.

Theorem 1.7. *Suppose for all sign combinations $u(x) \in \mathbb{K}$, where $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$, i.e. $|u(x)| = 1$, there is a function $q \in \operatorname{ran}(A)$, such that*

$$\begin{aligned} q(x) &= u(x), \quad \textit{for } x \in \mathcal{X}, \\ |q(x)| &< 1, \quad \textit{for } x \in \mathcal{X}^c = X \setminus \mathcal{X}, \end{aligned} \tag{DC}$$

then A^* has the null space property (NSP) *with respect to $\mathcal{X}$.*

Proof. Let $\lambda \in \ker(A^\star) \setminus \{0\}$ and

$$\lambda = \lambda_{\mathcal{X}} + \lambda_{\mathcal{X}^c}$$

be the Lebesgue decomposition with respect to $\nu = \sum_{x \in \mathcal{X}} \delta_x$. If $\lambda_{\mathcal{X}} = 0$, then the inequality

$$0 = \|\lambda_{\mathcal{X}}\|_{TV} < \|\lambda_{\mathcal{X}^c}\|_{TV}$$

holds trivially since $\lambda \neq 0$.

Otherwise, using the *polar decomposition* of $\lambda_{\mathcal{X}}$, see Appendix B, we find a function u with $|u(x)| = 1$ for all $x \in X$, such that

$$\lambda_{\mathcal{X}} = u \cdot |\lambda_{\mathcal{X}}|.$$

By assumption, we can find $q \in \operatorname{ran}(A)$, i.e. $q = Ac$ for some $c \in H$, such that

$$\begin{aligned} q(x) &= \overline{u(x)}, \quad \text{for } x \in \mathcal{X}, \\ |q(x)| &< 1, \quad \text{for } x \in \mathcal{X}^c = X \setminus \mathcal{X}, \end{aligned}$$

which yields together with $A^*\lambda = 0$,

$$\|\lambda_{\mathcal{X}}\|_{TV} + \langle \lambda_{\mathcal{X}^c}, q\rangle = \langle \lambda, q\rangle = \langle A^*\lambda, c\rangle_H = 0.$$

If $\lambda_{\mathcal{X}^c} = 0$, then $\lambda_{\mathcal{X}} = 0$, otherwise

$$|\langle \lambda_{\mathcal{X}^c}, q\rangle| < \|\lambda_{\mathcal{X}^c}\|_{TV},$$

and thus

$$\|\lambda_{\mathcal{X}}\|_{TV} < \|\lambda_{\mathcal{X}^c}\|_{TV}.$$

□

Remark 1.8. *In the case the measure is real-valued and one restricts the minimization to the space $\mathcal{M}(X, \mathbb{R})$ of signed measures, the proof shows that it is sufficient to fulfill the condition* (DC) *only for real-valued signs, i.e. $u(x) \in \{-1, 1\}$. Later on, in the case of the rotation group and the sphere, we will restrict ourselves to signed measures. Furthermore, if the measure is known to be real-valued and positive, one can even restrict to $u \equiv 1$.*

The connection of the existence of dual certificates to the minimization property of the measure becomes more clear in the context of the convexity of the problem. We state this connection in order to make clear the relation of the condition (DC) and a slightly different condition, that can be found in [de Castro and Gamboa, 2012, Lemma 1.1]. In view of Fermat's rule, $\mu \in \mathcal{M}(X, \mathbb{K})$ is a minimizer of (GP), if and only if

$$0 \in \partial J(\mu),$$

where J is given in (1.5) and $\partial J(\mu)$ is the subdifferential of J at μ, see Appendix D. By the Moreau-Rockafellar Theorem, this is equivalent to

$$0 \in \partial\|\cdot\|_{TV}(\mu) + \partial\chi_y(\mu),$$

particularly there is $Q \in \partial\|\cdot\|_{TV}(\mu)$ such that $-Q \in \partial\chi_y(\mu)$. Since

$$\partial\chi_y(\mu) = \{Q \in \mathcal{M}(X,\mathbb{K})' : \langle Q, \nu - \mu\rangle \leq 0, \text{ for all } \nu \in \mathcal{M}(X,\mathbb{K}), A^*\nu = y\},$$

we have

$$\partial\chi_y(\mu) = \ker(A^*)^\perp = \{Q \in \mathcal{M}(X,\mathbb{K})' : \langle Q, \nu\rangle = 0, \text{ for all } \nu \in \ker(A^*)\},$$

i.e. the annihilator of $\ker(A^*)$, meaning μ is a minimum if and only if there is a subgradient at μ that is 'perpendicular' to $\ker(A^*)$. By duality, one always has $\operatorname{ran}(A^{**}) \subseteq \ker(A^*)^\perp$ and the stronger assumption $\operatorname{ran}(A^{**}) \cap \partial\|\cdot\|_{TV}(\mu) \neq \emptyset$, which means there exists $h \in H$, such that

$$A^{**}h \in \partial\|\cdot\|_{TV}(\mu), \tag{SC}$$

is an additional regularity assumption for the measure μ. In the context of inverse problems, the condition (SC) is known as *source condition*, see e.g. [Bredies and Pikkarainen, 2013], [Burger and Osher, 2004], [Hofmann et al., 2007] and [Scherzer and Walch, 2009].

In the case $\operatorname{ran}(A^*)$ is closed, one can show $\ker(A^*)^\perp = \operatorname{ran}(A^{**})$, which means μ is a minimizer of (GP), if and only if the condition (SC) holds.

In the situation A^* is the adjoint of some operator, one can further simplify this condition. Following [Bredies and Pikkarainen, 2013], we have

$$\partial\|\cdot\|_{TV}(\mu) = \{Q \in \mathcal{M}(X,\mathbb{K})' : \langle Q, \mu\rangle = \|\mu\|_{TV}, \|Q\|_{\mathcal{M}(X,\mathbb{K})'} = 1\},$$

and therefore (SC) is equivalent to

$$\exists\, h \in H, \text{ such that } \langle A^{**}q, \mu\rangle = \|\mu\|_{TV}, \text{ and } \|A^{**}q\|_{\mathcal{M}(X,\mathbb{K})'} = 1,$$

or

$$\exists\, h \in H, \text{ such that } \langle \mu, Ah\rangle = \|\mu\|_{TV}, \text{ and } \|Ah\|_\infty = 1.$$

Using the polar decomposition $\mu = \operatorname{sign}_\mu \cdot|\mu|$, we have $\langle \mu, Ah\rangle = \|\mu\|_{TV}$ is equivalent to

$$\int_B (1 - Ah(x) \cdot \overline{\operatorname{sign}_\mu(x)})\,\mathrm{d}|\mu|(x) = 0,$$

which means $Ah(x) \cdot \operatorname{sign}_\mu(x) = 1$ for $|\mu|-$almost all $x \in X$. In the case the measure is discrete with $\operatorname{supp}(\mu) \subset \mathcal{X}$, this is equivalent to to the existence of $h \in H$ such that

$$Ah(x) = \operatorname{sign}_\mu(x), \quad x \in \mathcal{X}, \tag{1.6}$$

$$|Ah(x)| \leq 1, \quad x \in \mathcal{X}^c. \tag{1.7}$$

This means, in the case that $\operatorname{ran}(A^*)$ is closed, μ is a minimizer of (GP), if and only if there exists $q = Ah \in \operatorname{ran}(A)$ such that the interpolation condition (1.6), i.e. interpolation of the sign of the measure μ is fulfilled and the supremum norm of Ah is bounded by one. Such an element $q = Ah$ certifies the optimality of the measure μ, hence the name *dual certificate*.

If one tightens the condition (1.7) to

$$|Ah(x)| < 1, \quad x \in \mathcal{X}^c,$$

then one can show in addition, that each minimizer is a discrete measure with support in $\mathcal{X}$, see [de Castro and Gamboa, 2012, Lemma A.1]. Finally, the assumption that A^* is injective on all measures with support in $\mathcal{X}$ is sufficient for recovery of the measure μ. Summarizing, the condition

$$\begin{aligned} Ah(x) &= \operatorname{sign}_\mu(x), \quad x \in \mathcal{X}, \\ |Ah(x)| &< 1, \quad x \in \mathcal{X}^c, \\ A^*\nu = 0,\ \operatorname{supp}(\nu) \subset \mathcal{X} &\implies \nu = 0, \end{aligned} \tag{DCa}$$

guarantees recovery of the specific measure μ by minimizing (GP), see [de Castro and Gamboa, 2012, Lemma 1.1]. The proof is very similar to that of Theorem 1.7, despite the fact that one only gets recovery of the specific measure μ. Indeed, the condition (DC) implies the condition (DCa).

1.1.2 Exact Recovery as an Inverse Problem

In general, one cannot assume to have access to the pure information $y = A^*\mu^\star$ but rather to a noisy version y^δ of it. It is therefore necessary to replace the minimization problem (GP) including the exact side condition $A^*\mu = y$. One popular approach is to solve the Thikonov-type minimization problem

$$\min_{\mu\in\mathcal{M}(X,\mathbb{K})} \frac{1}{2}\|y^\delta - A^*\mu\|_H^2 + \lambda\|\mu\|_{TV}. \tag{GP$_\lambda$}$$

The aim of this section is to summarize the convergence properties of the minimizers of (GP$_\lambda$) to the minimizers of (GP) as y^δ converges to y and $\lambda \to 0$. In the most general setting, the convergence results are stated in terms of *Bregman distances* and we refer mainly to results especially from [Burger and Osher, 2004], [Hofmann et al., 2007], [Scherzer and Walch, 2009] and [Bredies and Pikkarainen, 2013].

As utilized in these references, we will choose a deterministic noise model. This means, for $\delta > 0$ we will assume that y^δ obeys

$$\|y - y^\delta\|_H \leq \delta,$$

where $y = A^*\mu^\star$ denotes the exact data. The first theorem states, that under the additional regularity condition on the measure $\mu^\star$, given by the specialized source condition (DCa) discussed in the previous section, there is a a priori parameter choice for λ depending on the noise level δ, such that the solution of the Thikonov-type problems converge to the exact solution $\mu^\star$.

Theorem 1.9. *Suppose $\mu^\star$ is a discrete measure as given in* (1.4)*, that fulfils the source condition* (DCa)*. Further assume sequences $\delta_n \to 0, \lambda_n \to 0$ monotonically decreasing, such that*

$$\frac{\delta_n^2}{\lambda_n} \to 0.$$

Then each sequence of solution μ_{λ_n,δ_n} of

$$\min_{\mu\in\mathcal{M}(X,\mathbb{K})} \frac{1}{2}\|y^{\delta_n} - A^*\mu\|_H^2 + \lambda_n\|\mu\|_{TV},$$

converges to $\mu^\star$ in the weak* topology.

Proof. The specialized source condition assures that $\mu^\star$ is the unique solution to the exact problem (GP). Then the statement is a consequence of [Hofmann et al., 2007, Thm. 3.5], respectively [Bredies and Pikkarainen, 2013, Prop. 5]. □

To show quantitative estimates in this very general setting, one has to make use of the notion of *generalized Bregman distances* introduced in the context of inverse problems in [Burger and Osher, 2004]. In this case the distances between two measures is a set of distance functions depending on the subgradient, i.e. for $q \in \partial\|\cdot\|_{TV}(\nu)$ a distance is given by

$$D_q(\mu,\nu) = \|\mu\|_{TV} - \mathrm{Re}(\langle q,\mu\rangle).$$

For a good overview of Bregman distances in the context of inverse problems and further references, see [Burger, 2016]. Assuming the source condition (DCa), one can derive the following bound

$$D_{A^{**}h}(\mu_{\lambda,\delta},\mu^\star) \leq \frac{\|h\|_H^2}{2\lambda} + \frac{\lambda\delta^2}{2},$$

see e.g. [Burger and Osher, 2004, Thm. 2]. Instead of further exploring this very general setting, we will concentrate on the special case of harmonic information on the rotation group in the next section.

1.2 Super-Resolution on the Rotation Group

In this section, we further investigate the problem of super-resolution on the rotation group. We start this section with a short reminder on analysis on the rotation group, including the smooth structure as well as the analysis with respect to harmonic basis functions.

In the second subsection, we state the problem of super-resolution on the rotation group in more detail. As seen in the previous section, the key ingredient to analyze the recovery via total variation minimization is the existence of dual certificates. We describe a possible construction of a dual certificate using a Hermite interpolation.

1.2.1 Analysis on the Rotation Group

The rotation group $SO(3)$ is defined as the space of matrices

$$SO(3) := \{x \in \mathbb{R}^{3\times 3} : x^T x = I, \det x = 1\},$$

which is a group under the action of matrix multiplication. We will use two different parametrizations of the rotation group.

By *Euler's Rotation Theorem*, there is for each $x \in SO(3)$ a unit vector $e \in \mathbb{R}^3$ and an angle $\omega \in [0,\pi]$, such that x is a rotation with rotation axis e and rotation angle ω. Using *Rodrigues rotation formula* yields

$$x = I\cos(\omega) + (1-\cos(\omega))ee^T + [e]\sin(\omega),$$

where

$$[e] = \begin{pmatrix} 0 & -e_3 & e_2 \\ e_3 & 0 & -e_1 \\ -e_2 & e_1 & 0 \end{pmatrix}.$$

This identification shows that $SO(3)$ is diffeomorphic to the real three-dimensional projective space and is therefore a connected compact Lie group. The corresponding Lie algebra of the Lie group $SO(3)$ is given by the skew symmetric matrices,

$$\mathfrak{so}(3) = \{v \in \mathbb{R}^{3\times 3} : v^T = -v\}.$$

The tangent space at $x \in SO(3)$ can be written as

$$T_x SO(3) = \{v \in \mathbb{R}^{3\times 3} : vx^T = -xv^T\} = \{v \in \mathbb{R}^{3\times 3} : v = xw, w \in \mathfrak{so}(3)\}.$$

The generators of the Lie-Algebra $\mathfrak{so}(3)$ are given by

$$\mathcal{L}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathcal{L}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \mathcal{L}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{1.8}$$

with the commutator relations

$$[\mathcal{L}_1, \mathcal{L}_2] = \mathcal{L}_1\mathcal{L}_2 - \mathcal{L}_2\mathcal{L}_1 = \mathcal{L}_3, \quad [\mathcal{L}_3, \mathcal{L}_1] = \mathcal{L}_2, \quad [\mathcal{L}_2, \mathcal{L}_3] = \mathcal{L}_1.$$

A basis of the tangent space at $x \in SO(3)$ is thus given by $x\mathcal{L}_1, x\mathcal{L}_2, x\mathcal{L}_3$.

The exponential map $\exp_x : T_x SO(3) \to SO(3)$ at $x \in SO(3)$ is defined by

$$\exp_x(v) = x\mathrm{e}^{x^T v},$$

where $\mathrm{e}^A = \sum_k \frac{A^k}{k!}$ denotes the matrix exponential. The unique geodesic originating from $x \in SO(3)$ in direction $v \in T_x SO(3)$ has the form

$$\gamma_{x,v}(t) = \exp_x(tv).$$

Set

$$B_\varepsilon(0) = \{v \in \mathbb{R}^{3\times 3} : \|v\|_F < \varepsilon\},$$

where $\|v\|_F = \sqrt{\operatorname{tr}(v^T v)}$ denotes the Frobenius norm. Restricted to the set $T_x SO(3) \cap B_\varepsilon(0)$ the exponential map is invertible for $\varepsilon < \log(2)$ and its inverse $\log_x : SO(3) \cap B_\delta(x) \to T_x SO(3) \cap B_\varepsilon(0)$ is given by

$$\log_x(w) = x \log(x^T w),$$

where $\log(A) = \sum_{k=1}^\infty (-1)^{k+1} \frac{(A-I)^k}{k}$ denotes the matrix logarithm and

$$B_\delta(x) = \{w \in \mathbb{R}^{3\times 3} : \|\log(x^T w)\|_F < \delta\},$$

with $\delta = \sqrt{2}\varepsilon$. This parametrization is called *normal coordinates*. The Riemannian metric is defined for $x \in SO(3)$ as

$$g_x(v, w) = \frac{1}{2}\operatorname{tr}(v^T w), \quad v, w \in T_x SO(3),$$

and $x\mathcal{L}_1, x\mathcal{L}_2, x\mathcal{L}_3$ is an orthogonal basis with respect to this inner product. Hence, the gradient of f in normal coordinates centered at x is represented by

$$\nabla f(x) = \sum_{i=1}^3 X_i f(x) \cdot x\mathcal{L}_i,$$

where

$$X_i f(x) = \lim_{t\to 0} t^{-1}(f(x e^{t\mathcal{L}_i}) - f(x)),$$

whenever f is differentiable. For $t \in \mathbb{R}$, the corresponding elements in $SO(3)$ are given by

$$\mathrm{e}^{t\mathcal{L}_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(t) & -\sin(t) \\ 0 & \sin(t) & \cos(t) \end{pmatrix}, \mathrm{e}^{t\mathcal{L}_2} = \begin{pmatrix} \cos(t) & 0 & \sin(t) \\ 0 & 1 & 0 \\ -\sin(t) & 0 & \cos(t) \end{pmatrix},$$
$$\mathrm{e}^{t\mathcal{L}_3} = \begin{pmatrix} \cos(t) & -\sin(t) & 0 \\ \sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Similarly, for a two times differentiable function f the *Hessian* can be represented in normal coordinates by the matrix

$$Hf = \begin{pmatrix} X_1X_1f & X_1X_2f & X_1X_3f \\ X_2X_1f & X_2X_2f & X_2X_3f \\ X_3X_1f & X_3X_2f & X_3X_3f \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 0 & X_3f & -X_2f \\ -X_3f & 0 & X_1f \\ X_2f & -X_1f & 0 \end{pmatrix}. \tag{1.9}$$

Using these differential operators, we have the following Taylor formula, see e.g. [Chirikjian, 2012]. Let $f : SO(3) \to \mathbb{R}$ be a two times continuous differentiable function and $Y \in \mathfrak{so}(3)$ with $\|Y\|_F = 1$, then for $x \in SO(3)$ and $t \in \mathbb{R}$

$$f(x\exp(tY)) = f(x) + t\nabla f(x)^T e(Y) + \frac{t^2}{2} e(Y)^T Hf(x\exp(\xi Y))e(Y), \tag{1.10}$$

for some ξ with $|\xi| \leq |t|$, where $e(Y) = \begin{pmatrix} e_1(Y) & e_2(Y) & e_3(Y) \end{pmatrix}^T$ and $Y = \sum_{i=1}^3 e_i(Y)\mathcal{L}_i$.

The second parametrization we will use is given by *Euler angles*. Each element $x \in SO(3)$ can be represented by

$$x = R_Z(\alpha)R_Y(\beta)R_Z(\gamma),$$

with $(\alpha, \beta, \gamma) \in [0, 2\pi) \times [0, \pi] \times [0, 2\pi)$ and

$$R_Z(t) = \begin{pmatrix} \cos(t) & -\sin(t) & 0 \\ \sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad R_Y(t) = \begin{pmatrix} \cos(t) & 0 & \sin(t) \\ 0 & 1 & 0 \\ -\sin(t) & 0 & \cos(t) \end{pmatrix}.$$

The triplet (α, β, γ) is called *Euler angles* in the ZYZ-convention.

A distance on $SO(3)$, that is compatible with its topology and invariant with respect to the group action, is given by

$$\omega(y^{-1}x) := \arccos\left(\frac{\operatorname{tr}(y^{-1}x) - 1}{2}\right) = \frac{1}{\sqrt{2}}\|\log(y^T x)\|_F,$$

which is equal to the rotation angle of the matrix $y^{-1}x$.

Since $SO(3)$ is a compact group, there is a regular Borel measure λ, that is invariant under the group action, i.e. $\lambda(xB) = \lambda(B) = \lambda(Bx)$ for all Borel sets B. This measure can be normalized such that

$$\int_{SO(3)} \mathrm{d}\lambda(x) = 1.$$

Using an Euler angle parametrization, we can write down the integral for each measurable function $f : SO(3) \to \mathbb{C}$ explicitly as

$$\int_{SO(3)} f(x)\mathrm{d}\lambda(x) = \frac{1}{8\pi^2}\int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} f(x(\alpha,\beta,\gamma))\sin(\beta)\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma.$$

For functions that only depend on the rotation angle, i.e. $f(x) = \tilde{f}(\omega(x))$ the integral reduces to

$$\int_{SO(3)} f(x)\mathrm{d}\lambda(x) = \frac{2}{\pi}\int_0^{\pi} \tilde{f}(t)\sin^2\left(\frac{t}{2}\right)\mathrm{d}t.$$

The space $L^2(SO(3))$ of square-integrable functions with respect to λ is defined in the usual way. The Peter-Weyl Theorem now states that the right regular representation of $SO(3)$ splits up into an orthogonal direct sum of irreducible finite-dimensional representations and the matrix coefficients of these irreducible representation form an orthogonal basis for $L^2(SO(3))$. The dimensions of the irreducible representations are given by $2l + 1$, $l \in \mathbb{N}$ and the matrix coefficients $D^l_{k,m}$, $-l \leq k, m \leq l$ are called *Wigner-D functions*. We have that

$$\{\sqrt{2l+1}D^l_{k,m}, -l \leq k, m \leq l, l \in \mathbb{N}\}$$

form an orthonormal basis of $L^2(SO(3))$. The value $l \in \mathbb{N}$ is called *degree*. In the Euler angle parametrization the Wigner D-functions are given for $l \in \mathbb{N}$ and $-l \leq k, m \leq l$ by

$$D^l_{k,m}(\alpha,\beta,\gamma) = \mathrm{e}^{-ik\alpha}d^l_{k,m}(\cos(\beta))\mathrm{e}^{-im\gamma},$$

where $d^l_{k,m}$ is defined as

$$d^l_{k,m}(t) = C_{l,k,m}(1-t)^{-(m-k)/2}(1+t)^{-(m+k)/2}\frac{d^{l-m}}{dt^{l-m}}\left((1-t)^{l-k}(1+t)^{l+k}\right),$$

with $C_{l,k,m} = \frac{(-1)^{l-k}}{2^l(l-k)!}\sqrt{\frac{(l-k)!(l+m)!}{(l+k)!(l-m)!}}$.

Another very useful representation of the functions $d^l_{k,m}$ is given by

$$d^l_{k,m}(\cos(\beta)) = i^{m+k}\sum_{j=-l}^{l}(-1)^j d^l_{j,k}(0)d^l_{m,j}(0)\mathrm{e}^{ij\beta}, \tag{1.11}$$

see e.g. [Edmonds, 1957, p. 62] or [Risbo, 1996]. The addition theorem for Wigner D-functions states

$$\sum_{-l\le k,m\le l} D^l_{k,m}(x)\overline{D^l_{k,m}(y)} = U_{2l}\left(\cos\left(\frac{\omega(y^{-1}x)}{2}\right)\right), \tag{1.12}$$

where U_n denotes the n-th Chebychev polynomial of the second kind.

The space of all finite linear combinations of Wigner-D functions with degree less or equal to N is be denoted as

$$\Pi_N(SO(3)) := \operatorname{span}\{D^l_{k,m} :\ -l \le k,m \le l,\ l \le N\}$$

and is also called *generalized polynomials* of degree N. The projection onto the space of generalized polynomials of degree N can be written as

$$\mathcal{S}_N : L^2(SO(3)) \to C(SO(3)),$$
$$\mathcal{S}_N f(x) = \int_{SO(3)} f(y)D_N(x,y)\mathrm{d}\lambda(y),$$

where

$$D_N(x,y) = \sum_{l=0}^{N}(2l+1)\sum_{k,m=-l}^{l} D^l_{k,m}(x)\overline{D^l_{k,m}(y)},$$

is the Dirichlet kernel on the rotation group. Using the addition theorem, one also has

$$D_N(x,y) = \sum_{l=0}^{N}(2l+1)U_{2l}\left(\frac{\cos(\omega(y^{-1}x)))}{2}\right).$$

The closed form expression of D_N is given by $D_N(x,y) = \tilde{D}_N(\omega(y^{-1}x))$, with

$$\tilde{D}_N(\omega) = \begin{cases} \frac{(2N+3)\sin(N+1/2)\omega)-(2N+1)\sin((N+3/2)\omega)}{4\sin^3(\omega/2)}, & \omega \ne 0, \\ \frac{1}{3}(N+1)(2N+1)(2N+3), & \omega = 0, \end{cases}$$

see [Schmid, 2008]. The differential operators defined above map $\Pi_N(SO(3))$ to itself. More concrete, we have

$$X_1 D^l_{k,m}(x) = \frac{1}{2}ic^l_{-m}D^l_{k,m-1}(x) + \frac{1}{2}ic^l_m D^l_{k,m}(x),$$
$$X_2 D^l_{k,m}(x) = -\frac{1}{2}c^l_{-m}D^l_{k,m-1}(x) + \frac{1}{2}c^l_m D^l_{k,m}(x),$$
$$X_3 D^l_{k,m}(x) = -imD^l_{k,m}(x),$$

with $c^l_m = \sqrt{(l-m)(l+m+1)}$, see e.g. [Chirikjian and Kyatkin, 2000].

Having introduced the Wigner D-functions, we show a possible construction of a dual certificate in the next subsection, which is the important part to analyze the recovery with respect to these Wigner D-functions.

1.2.2 A Dual Certificate on the Rotation Group

In the following, we will introduce the problem of super-resolution on the rotation group, i.e. exact recovery of Dirac measures from low frequency information. To be more specific, consider a Dirac measure of the form

$$\mu^\star = \sum_{i=1}^{M} c_i \delta_{x_i}, \tag{1.13}$$

where $M \in \mathbb{N}$, $c_i \in \mathbb{R}$ are real valued coefficients and δ_{x_i} are the point measures centered at pairwise distinct $x_i \in SO(3)$. All parameters M, c_i, x_i are unknown and we can only access

$$\mathcal{S}_N^* \mu^\star(x) = \int_{SO(3)} D_N(x,y) \mathrm{d}\mu^\star(y),$$

for a possible low degree N. On the spectral side, this means we can access the moments of $\mu^\star$ with respect to the functions $D^l_{k,m}$ for $-l \le k, m \le l$ only for $l \le N$, i.e.

$$\langle \mu^*, D^l_{k,m} \rangle := \int_{SO(3)} \overline{D^l_{k,m}(x)} \mathrm{d}\mu^\star(x) = \sum_{i=1}^{M} c_i \overline{D^l_{k,m}(x_i)},$$

for $-l \le k, m \le l, l \le N$. The question is, under which conditions on the support of the measure $\mu^\star$ it is the solution of the total variation minimizaton

$$\min_{\mu \in \mathcal{M}(SO(3),\mathbb{R})} \|\mu\|_{TV}, \quad \text{subject to} \quad \mathcal{S}_N^* \mu = \mathcal{S}_N^* \mu^\star. \tag{RP}$$

As seen in the previous section, this relies on the existence of a dual certificate, i.e. a $q \in \Pi_N(SO(3))$, such that

$$\begin{aligned} q(x_i) &= u_i, \quad \text{for } x_i \in \mathcal{X}, \\ |q(x)| &< 1, \quad \text{for } x \in SO(3) \setminus \mathcal{X}. \end{aligned}$$

Clearly, the existence of such a function should be coupled to the minimal separation of the support points, i.e. the value

$$\rho(\mathcal{X}) := \min_{x_i \neq x_j} \omega(x_j^{-1} x_i).$$

Otherwise, two collapsing interpolation points would result in a growing value of the derivatives, which is not possible due to the Bernstein inequality

$$\|X_i q\|_\infty \le N \|q\|_\infty, \quad \text{for } q \in \Pi_N(SO(3)), \tag{1.14}$$

see [Schmid, 2008].

Indeed, a proportional coupling of the minimal separation to the degree of the given moments is sufficient for the existence of a dual certificate. More precisely, if

$$\rho(\mathcal{X}) \ge \frac{36}{N+1}$$

for $N \ge 20$, then for each sign combination $u_i \in \{-1, 1\}$, there is a $q \in \Pi_N(SO(3))$ such that

$$\begin{aligned} q(x_i) &= u_i, \quad \text{for } x_i \in \mathcal{X}, \\ |q(x)| &< 1, \quad \text{for } x \in SO(3) \setminus \mathcal{X}. \end{aligned} \tag{1.15}$$

The proof is based on an explicit construction of the dual certificate. We will follow ideas from [Candés and Fernandez-Granda, 2014], where the construction was done for *trigonometric polynomials*. Here, we shortly describe the construction process to point out the importance of localization estimates for kernels and their derivatives, which are established in Chapter 2. The actual proof, i.e. showing the properties (1.15), is postponed to the proof of Theorem 3.6 in Chapter 3.

In order to satisfy the conditions (1.15), one formulates the *Hermite-type* interpolation problem

$$\begin{aligned} q(x_i) &= u_i, \\ X_1 q(x_i) = X_2 q(x_i) &= X_3 q(x_i) = 0, \end{aligned} \tag{1.16}$$

for $x_i \in \mathcal{X}$, where X_k are the differential operators defined in Section 1.2.1. This means, beside the interpolation itself, we ask for local extrema at the interpolation points. One then seeks a solution q to this interpolation problem in the space $\Pi_N(SO(3))$, that satisfies, due to the local extrema conditions, $|q(x)| < 1$ for $x \in SO(3) \setminus \mathcal{X}$. The constructed interpolant is of the form

$$q(x) = \sum_{i=1}^{M} \alpha_{i,0}\sigma_N(x, x_i) + \alpha_{i,1} X_1^y \sigma_N(x, x_i) + \alpha_{j,2} X_2^y \sigma_N(x, x_i) + \alpha_{j,3} X_3^y \sigma_N(x, x_i),$$

where σ_N is an *interpolation kernel* of the form

$$\sigma_N(x, y) = \sum_{l=0}^{N} h_N(l) \sum_{-l \le k,m \le l} D_{k,m}^l(x) \overline{D_{k,m}^l(y)},$$

with positive weights $h_N(l) > 0$. Observe, that the expressions $\sigma_N(x, x_i)$ and $X_j^y \sigma_N(x, x_i)$, where the superscript indicates the action of the differential operator on the second variable, are by construction generalized polynomials of degree N in the first variable, which means $q \in \Pi_N(SO(3))$. Applying the interpolation conditions (1.16) leads to the linear system of equations

$$K\alpha := \begin{pmatrix} \sigma_N & X_1^x \sigma_N & X_2^x \sigma_N & X_3^x \sigma_N \\ X_1^y \sigma_N & X_1^x X_1^y \sigma_N & X_2^x X_1^y \sigma_N & X_3^x X_1^y \sigma_N \\ X_2^y \sigma_N & X_1^x X_2^y \sigma_N & X_2^x X_2^y \sigma_N & X_3^x X_2^y \sigma_N \\ X_3^y \sigma_N & X_1^x X_3^y \sigma_N & X_2^x X_3^y \sigma_N & X_3^x X_3^y \sigma_N \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} u \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{1.17}$$

where the entries in the matrix corresponds to blocks of the form $\sigma_N = (\sigma_N(x_i, x_j))_{i,j=1}^M$ and in the same way for the derivatives. The blocks in the vectors are given by $\alpha_k = (\alpha_{k,j})_{j=1}^M$ for $k = 0, 1, 2, 3$ and $u = (u_j)_{j=1}^M$. To find the coefficients, we have to show the invertibility of the matrix K. Due to the block structure of K this is done using an *iterative block inversion*, explained in Section 3, and the fact that a matrix A is invertible if

$$\|I - A\|_\infty < 1,$$

where $\|A\|_\infty = \max_i \sum_j |a_{i,j}|$. In this case the norm of the inverse is bounded by

$$\|A^{-1}\|_\infty \le \frac{1}{1 - \|I - A\|_\infty},$$

see Appendix C. Thus, to show the invertibility of the matrix K, we have to employ *localization estimates* for the entries of the matrix K, which means we have to bound the expressions $|\sigma_N(x_i, x_j)|$, $|X_k^y \sigma_N(x_i, x_j)|$ and $|X_n^x X_k^y \sigma_N(x_i, x_j)|$. The values of these expressions should decrease, if the distance of $\omega(x_j^{-1} x_i)$ becomes bigger. We are locking for estimates of the form

$$|\sigma_N(x_i, x_j)| \le \frac{c}{((N+1)\omega(y^{-1}x))^s},$$

for some constants s and c only depending on the weights h_N and similar estimates for the derivatives. Using these estimates we find explicit bounds on the supremum norm of the coefficients. Once we have bound the coefficients, we have to show the condition $|q(x)| < 1$, where x is not an interpolation point. This includes convexity arguments for the interpolant q, which means we have to deal with the entries of the *Hessian matrix* of q, where third mixed derivatives appear. For this reason, we also need localization estimates for third derivatives.

The key ingredients for the construction of the interpolant q are *localization estimates* for the interpolation kernel σ_N and its various derivatives. Moreover, we need *explicit constants* in these estimates to show the claimed properties of the interpolant. This is the topic of the next chapter.

Notes and References. *The assertions regarding the abstract recovery problem can be found in one form or another in [Burger and Osher, 2004], [Hofmann et al., 2007], [Scherzer and Walch, 2009] and [Bredies and Pikkarainen, 2013]. Especially the existence of minimizers and the notion of* source conditions *are concepts, that are valid for general Banach spaces.*

The abstract setting we chose in Section 1.1 is very close to that chosen in [Bredies and Pikkarainen, 2013], including the connection of a dual certificate to a source condition. Rather than originality, the intention of Section 1.1 is to give a concise summary of the 'standard' framework for the recovery via the convex minimization problem (GP) *and to point out the central importance of a* dual certificate.

To the best of our knowledge, the problem of super-resolution on the rotation group $SO(3)$ *is first considered in this thesis. The construction of the candidate of a dual certificate, i.e. solving the Hermite interpolation problem* (1.16)*, is inspired by the article [Candés and Fernandez-Granda, 2014], where this procedure was proposed for trigonometric moments. Nevertheless, the realization of this so to say meta-principle is the crucial point and the content of the Chapters 2 and 3.*

Chapter 2

Localized Kernels

As seen in the previous chapter, the proposed construction of a dual certificate by solving the linear system (1.17) requires to control pointwise the quantities $|\sigma_N(x,y)|$, $|X_n^y \sigma_N(x,y)|$ and $|X_i^x X_n^y \sigma_N(x,y)|$ for an interpolation kernel σ_N. By pointwise control we mean estimates of the form

$$|\sigma_N(x,y)| \leq \frac{c_s}{(N+1)^s \omega(x^{-1}y)^s}, \tag{2.1}$$

where $s > 0$ and c_s is a constant depending only on s. In addition, since bounds of the form (2.1) are only meaningful in the case that x and y are sufficiently separated, we need different bounds for x and y being close to each other. Deriving bounds of the form (2.1) and bounds that control the behavior of the kernels near the diagonal $x = y$ is the content of this chapter.

The kernels we use are of the form

$$\sigma_N(x,y) = \sum_{l=0}^{N} h_N(l) \sum_{-l \leq k,m \leq l} D_{k,m}^l(x) \overline{D_{k,m}^l(y)},$$

with positive weights $h_N(l) > 0$. By the addition formula (1.12) for Wigner D-functions, this can also be written as

$$\sigma_N(x,y) = \sum_{l=0}^{N} h_N(l) U_{2l} \left(\cos \left(\frac{\omega(y^{-1}x)}{2} \right) \right) = \sum_{l=0}^{N} h_N(l) \sum_{k=-l}^{l} \mathrm{e}^{ik\omega(y^{-1}x)}, \tag{2.2}$$

where U_{2l} is the Chebyshev polynomial of the second kind of order $2l$. In other words, the interpolation kernels are *radial* kernels, i.e. they are of the form

$$\sigma_N(x,y) = \tilde{\sigma}_N(\omega(x^{-1}y)),$$

where $\tilde{\sigma}_N$ is a trigonometric polynomial. For this reason, the localization properties of the interpolation kernel can be derived from corresponding localization principles for *trigonometric polynomials*, see e.g. [Gräf and Kunis, 2008]. In contrast to this, the derivatives of these interpolation kernels are no longer radial, but should be controllable by the ordinary derivatives of the trigonometric polynomial $\tilde{\sigma}_N$, which we show in Section 2.2. Beforehand, we choose specific weights, that allow for good control of the trigonometric polynomial $\tilde{\sigma}_N$ and its derivatives, which is the content of Section 2.1.

2.1 Localized Trigonometric Polynomials

We start this section by fixing specific weights in the kernel expansion (2.2). The weights are generated by sampling a function. More specific, we define the weights $h_N(l)$ by

$$h_N(l) = \frac{1}{\|g\|_{1,N}} \begin{cases} g\left(\frac{l}{2(N+1)}\right) - g\left(\frac{l+1}{2(N+1)}\right), & 0 \leq l < N, \\ g\left(\frac{N}{2(N+1)}\right), & l = N, \end{cases}$$

where $g : \mathbb{R} \to \mathbb{R}_+$ is a symmetric positive function with $\operatorname{supp}(g) \subseteq [-\frac{1}{2}, \frac{1}{2}]$, which is decreasing for positive values. Its discrete coefficient norm is given by

$$\|g\|_{1,N} := \sum_{l=-N}^{N} g\left(\frac{l}{2(N+1)}\right).$$

This results in

$$\sigma_N(x,y) = \tilde{\sigma}_N(\omega(y^{-1}x)) = \frac{1}{\|g\|_{1,N}} \sum_{k=-N}^{N} g\left(\frac{k}{2(N+1)}\right) \mathrm{e}^{ik\omega(y^{-1}x)}.$$

In [Mhaskar and Prestin, 2000] and also in [Kunis and Potts, 2007], it was shown that a trigonometric polynomial of the form

$$\sum_{k=-N}^{N} g\left(\frac{k}{2(N+1)}\right) \mathrm{e}^{ikt}$$

obeys a localization property, as long as the function g is sufficiently smooth with derivatives of *bounded variation*. The variation of a function f defined on $\left[-\frac{1}{2}, \frac{1}{2}\right]$ is given by

$$|f|_V := \sup \left\{ \sum_{i=1}^{n-1} |f(t_{i+1}) - f(t_i)| \right\},$$

where the supremum is taken over all partitions $(t_i)_{i=1}^n$ of the interval $[-\frac{1}{2}, \frac{1}{2}]$. The space of $(s-1)$-times differentiable functions g with compact support in $\left[-\frac{1}{2}, \frac{1}{2}\right]$, such that $|g^{(s-1)}|_V < \infty$, will be denoted as $\mathcal{BV}_0^{s-1}([-\frac{1}{2}, \frac{1}{2}])$. Equipped with these definitions, we have for $g \in \mathcal{BV}_0^{s-1}([-\frac{1}{2}, \frac{1}{2}])$, see [Kunis and Potts, 2007, Lemma 3.2],

$$\left| \sum_{k=-N}^{N} g\left(\frac{k}{2(N+1)}\right) \mathrm{e}^{ikt} \right| \leq \frac{(2^s - 1)\zeta(s)|g^{(s-1)}|_V}{(4(N+1))^{s-1}|t|^s}, \tag{2.3}$$

for $t \in (0, \pi]$ with $N \geq s - 1 \geq 1$. Here, $\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{n^s}$ denotes the Riemannian zeta function. In addition, for positive $g \in \mathcal{BV}_0^{s-1}([-\frac{1}{2}, \frac{1}{2}])$ one has the bounds

$$\left(\|g\|_1 - \frac{2\zeta(s)}{(2N\pi)^s} |g^{(s-1)}|_V \right) \leq \frac{\|g\|_{1,N}}{2(N+1)} \leq \left(\|g\|_1 + \frac{2\zeta(s)}{(2N\pi)^s} |g^{(s-1)}|_V \right), \tag{2.4}$$

see [Kunis and Potts, 2007, Lemma 3.2]. This leads to explicit constants for localization results of the trigonometric polynomial

$$\tilde{\sigma}_N(t) = \frac{1}{\|g\|_{1,N}} \sum_{k=-N}^{N} g\left(\frac{k}{2(N+1)}\right) \mathrm{e}^{ikt},$$

as long as one can compute the L^1-norm of the filter function and the total variation of its $(s-1)$-th derivative. The $l-$th derivative of the trigonometric polynomial $\tilde{\sigma}_N$ is given by

$$\tilde{\sigma}_N^{(l)}(t) = \frac{(2(N+1)i)^l}{\|g\|_{1,N}} \sum_{k=-N}^{N} \left(\frac{k}{2(N+1)}\right)^l g\left(\frac{k}{2(N+1)}\right) \mathrm{e}^{ikt}.$$

Thus, to achieve an analog localization result for the derivatives we have to estimate the total variation of the function $(z^l g_s(z))^{(s-1)}$. To be able to bound the corresponding norms and variations, we will choose a specific function to be sampled. We seek for a $(s-1)$-smooth function, whose $(s-1)$-th derivative has small total variation. This leads to functions, whose $(s-1)$-th derivative is piecewise constant. One way to construct such a function is to use a *B-spline* function of order $s-1$, see e.g. [Gräf and Kunis, 2008]. We will use the so called *perfect B-spline* of order $s-1$ as filter function, since in this case $|g^{(s-1)}(x)| = 1$. These functions are given by

$$g_{s-1}(x) = \frac{(-1)^{s-1}}{(s-2)!} \int_{-1}^{x} \sum_{k=0}^{s-1} (-1)^k \chi_{(\cos(\frac{k+1}{s}\pi),\cos(\frac{k}{s}\pi)]}(t)(x-t)^{s-2}\mathrm{d}t,$$

for $s \in 2\mathbb{N}$. For reasons of clarity and comprehensibility, the details, properties and bounds for the corresponding variations and norms are stated in appendix A.

The aim of this section is to show, that for this specific choice of weights the trigonometric polynomial $\tilde{\sigma}_N$ and its derivatives obey

$$|\tilde{\sigma}_N^{(l)}(t)| \leq \frac{c_{l,s}}{(N+1)^{s-l}|t|^s}, \quad l = 0, \ldots, 3,$$

with explicit constants $c_{l,s}$ depending only on the order of the B-spline and the order of the derivative. The scaled function

$$\tilde{g}_{s-1}(x) = g_{s-1}(2x)$$

has its support in $[-\frac{1}{2}, \frac{1}{2}]$ and we have $\tilde{g}_{s-1} \in \mathcal{BV}_0^{s-1}([-\frac{1}{2}, \frac{1}{2}])$. We define the kernel by

$$\tilde{\sigma}_N(t) = \frac{1}{\|\tilde{g}_s\|_{1,N}} \sum_{k=-N}^{N} \tilde{g}_{s-1}\left(\frac{k}{2(N+1)}\right) \mathrm{e}^{ikt}.$$

Theorem 2.1. *Let $s \geq 6$, $s \in 2\mathbb{N}$ and $N \geq 2s$. Using the scaled perfect B-spline $\tilde{g}_{s-1}$ as filter function leads to localization estimates for the kernel $\tilde{\sigma}_N$ and its derivatives up to order 3, i.e. for $t \in [-\pi, \pi]\setminus\{0\}$ we have*

$$|\tilde{\sigma}_N^{(l)}(t)| \leq \frac{c_{l,s}}{(N+1)^{s-l}|t|^s}, \quad l = 0, \ldots, 3, \tag{2.5}$$

where the constants are given by

$$\begin{aligned}
c_{0,s} &= 1.02 \cdot (s-1)! \cdot 2^s \cdot s, \\
c_{1,s} &= 1.02 \cdot (s-1)! \cdot 2^s \cdot 2s, \\
c_{2,s} &= 1.02 \cdot (s-1)! \cdot 2^s \cdot (4s+1), \\
c_{3,s} &= 1.02 \cdot (s-1)! \cdot 2^s \cdot (9s-2).
\end{aligned}$$

Proof. The kernel and its derivatives up to order 3 are given by

$$\tilde{\sigma}_N^{(l)}(t) = \frac{(2(N+1)i)^l}{\|\tilde{g}_{s-1}\|_{1,N}} \sum_{k=-N}^{N} \left(\frac{k}{2(N+1)}\right)^l \tilde{g}_{s-1}\left(\frac{k}{2(N+1)}\right) \mathrm{e}^{ikt}, \quad l = 0, \ldots, 3.$$

In view of Proposition A.1 and Lemma A.3, we have $\|\tilde{g}_{s-1}\|_1 = \frac{1}{2}\|g_{s-1}\|_1 = \frac{1}{(s-1)!2^{s-1}}$ and $|\tilde{g}_{s-1}^{(s-1)}|_V = 2^s s$. Thus, using estimate (2.4) we can bound the discrete norm $\|\tilde{g}_{s-1}\|_{1,N}$ from below by

$$\|\tilde{g}_{s-1}\|_{1,N} \geq \frac{N+1}{2^{s-2}}\left(\frac{1}{(s-1)!} - \frac{\zeta(s)s}{(\pi s)^s}\right). \tag{2.6}$$

Combining this with the localization estimate (2.3) yields

$$\begin{aligned}
|\tilde{\sigma}_N^{(l)}(t)| &\leq \frac{(2(N+1))^l(2^s-1)\zeta(s)}{4^{s-1}(N+1)^{s-1}|t|^s}|(z^l\tilde{g}_{s-1})^{(s-1)}|_V \frac{2^s}{4(N+1)\left(\frac{1}{(s-1)!} - \frac{\zeta(s)s}{(\pi s)^s}\right)}, \\
&= \frac{2^l|(z^l\tilde{g}_{s-1})^{(s-1)}|_V}{(N+1)^{s-l}|t|^s}\frac{(2^s-1)\zeta(s)}{2^s\left(\frac{1}{(s-1)!} - \frac{\zeta(s)s}{(\pi s)^s}\right)}, \\
&= \frac{2^l|(z^l\tilde{g}_{s-1})^{(s-1)}|_V}{(N+1)^{s-l}|t|^s}\frac{(s-1)!}{\left(\frac{1}{\zeta(s)} - \frac{s(s-1)!}{(\pi s)^s}\right)}.
\end{aligned}$$

Observe, that the sequences $\frac{1}{\zeta(s)}$ and $\frac{-s(s-1)!}{(\pi s)^s}$ are monotonically increasing in s. This means, we can bound them from below by the first possible value for s, that is $s = 6$. This gives the upper bound

$$\frac{1}{\left(\frac{1}{\zeta(s)} - \frac{s(s-1)!}{(\pi s)^s}\right)} \leq \frac{1}{\left(\frac{945}{\pi^6} - \frac{720}{\pi^6 6^6}\right)} \leq 1.02,$$

which results in

$$|\tilde{\sigma}_N^{(l)}(t)| \leq \frac{1.02(s-1)!2^l|(z^l\tilde{g}_{s-1})^{(s-1)}|_V}{(N+1)^{s-l}|t|^s}.$$

In view of

$$(z^l\tilde{g}_{s-1})^{(s-1)} = \sum_{n=0}^{l}\binom{s-1}{n}(z^l)^{(n)}\tilde{g}_{s-1}^{(s-1-n)},$$

and

$$|uv|_V \leq \|u\|_\infty|v|_V + \|v\|_\infty|u|_V,$$

for two functions u and v, we get for $l \leq s-1$

$$\begin{aligned}
|(z^l\tilde{g}_{s-1})^{(s-1)}|_V &= \left|\sum_{n=0}^{l}\binom{s-1}{n}(z^l)^{(n)}\tilde{g}_{s-1}^{(s-1-n)}\right|_V, \\
&\leq \sum_{n=0}^{l}\frac{l!}{(l-n)!}\binom{s-1}{n}|z^{l-n}\tilde{g}_{s-1}^{(s-1-n)}|_V, \\
&\leq \sum_{n=0}^{l}\frac{l!}{(l-n)!}\binom{s-1}{n}\left(\|z^{l-n}\|_\infty|\tilde{g}_{s-1}^{(s-1-n)}|_V + \|\tilde{g}_{s-1}^{(s-1-n)}\|_\infty|z^{l-n}|_V\right).
\end{aligned}$$

Since $\|z^{l-n}\|_\infty = \frac{1}{2^{l-n}}$ and $|z^{l-n}|_V = \frac{1}{2^{l-n-1}}$ for $n < l$, we have

$$|(z^l\tilde{g}_{s-1})^{(s-1)}|_V \leq \sum_{n=0}^{l-1}\frac{l!}{(l-n)!}\binom{s-1}{n}\frac{1}{2^{l-n}}\left(|\tilde{g}_{s-1}^{(s-1-n)}|_V + 2\|\tilde{g}_{s-1}^{(s-1-n)}\|_\infty\right) \tag{2.7}$$

$$+ l! \binom{s-1}{l} |\tilde{g}_{s-1}^{(s-1-l)}|_V.$$

Using inequality (2.7) for the variation of the derivative of products together with Lemma A.3 and the estimate $\tan\left(\frac{\pi}{2s}\right) \leq \frac{2}{s}$ for $s \geq 6$ yields

$$\begin{aligned} |(z\tilde{g}_{s-1}(z))^{(s-1)}|_V &\leq 2^s s, \\ |(z^2\tilde{g}_{s-1}(z))^{(s-1)}|_V &\leq 2^{s-2}(4s+1), \\ |(z^3\tilde{g}_{s-1}(z))^{(s-1)}|_V &\leq 2^{s-3}(9s-2). \end{aligned}$$

and therefore the constants. □

The bounds of the previous theorem are only meaningful if ω is well separated from zero. For values of ω close to zero we will use different bounds derived from series expansion around zero. For this purpose, we need upper and lower bounds for the values of the second and fourth derivative of $\tilde{\sigma}_N$ at zero.

Lemma 2.2. *Let $s \geq 8$, $s \in 2\mathbb{N}$ and $N \geq 2s$. Using the scaled perfect B-spline $\tilde{g}_{s-1}$ as filter function leads to the following bounds*

$$c_s(N+1)^2 \leq |\tilde{\sigma}_N''(0)| \leq \tilde{c}_s(N+1)^2,$$

with

$$c_s = \frac{0.999}{2(s+1)}, \quad \tilde{c}_s = \frac{1.001}{2(s+1)},$$

and

$$d_s(N+1)^4 \leq |\tilde{\sigma}_N^{(4)}(0)| \leq \tilde{d}_s(N+1)^4,$$

with

$$d_s = \frac{3 \cdot 0.999}{4(s+2)(s+1)}, \quad \tilde{d}_s = \frac{3 \cdot 1.001}{4(s+2)(s+1)}.$$

In the case $s = 8$ we have for $N \geq 20$

$$|\tilde{\sigma}_N^{(6)}(0)| \leq 1.011 \cdot \frac{15}{8} \cdot \frac{8!}{11!} \cdot (N+1)^6. \tag{2.8}$$

Proof. We have for $m \in \mathbb{N}$

$$\begin{aligned} |\tilde{\sigma}_N^{(2m)}(0)| &= \frac{(2(N+1))^{2m}}{\|\tilde{g}_{s-1}\|_{1,N}} \sum_{k=-N}^{N} \left(\frac{k}{2(N+1)}\right)^{2m} \tilde{g}_{s-1}\left(\frac{k}{2(N+1)}\right), \\ &= \frac{(2(N+1))^{2m}}{\|\tilde{g}_{s-1}\|_{1,N}} \|z^{2m}\tilde{g}_{s-1}(z)\|_{1,N}. \end{aligned}$$

To estimate the expressions, we have to bound the discrete norms of the filter function. Using inequality (2.4), we have to calculate the L^1-norms of the functions $z^{2m}\tilde{g}_{s-1}$ on $[-\frac{1}{2}, \frac{1}{2}]$, which are given by

$$\|z^{2m}\tilde{g}_{s-1}(z)\|_1 = \frac{(2m)! \cdot s}{4^m \cdot m! \cdot 2^{s+2m-1}(s+m)!},$$

see Lemma A.2. Together with the bounds of the variations derived in Lemma A.3, we get by applying inequality (2.4)

$$|\tilde{\sigma}_N''(0)| = (2(N+1))^2 \frac{\|z^2\tilde{g}_{s-1}(z)\|_{1,N}}{\|\tilde{g}_{s-1}\|_{1,N}},$$

$$\geq (2(N+1))^2 \frac{\left(\|z^2\tilde{g}_{s-1}(z)\|_1 - \frac{2\zeta(s)}{(4\pi s)^s}|(z^2\tilde{g}_{s-1}(z))^{(s-1)}|_V\right)}{\left(\|\tilde{g}_{s-1}\|_1 + \frac{2\zeta(s)}{(4\pi s)^s}|\tilde{g}_{s-1}^{(s-1)}|_V\right)},$$

$$\geq \frac{(2(N+1))^2}{8(s+1)} \frac{1 - 2(s+1)(s-1)!\frac{\zeta(s)}{(\pi s)^s}(4s+1)}{1 + s!\frac{\zeta(s)}{(\pi s)^s}}. \tag{2.9}$$

We can bound the second quotient in (2.9) from below by its value at $s = 8$, i.e

$$\frac{1 - 2(s+1)(s-1)!\frac{\zeta(s)}{(\pi s)^s}(4s+1)}{1 + s!\frac{\zeta(s)}{(\pi s)^s}} \geq \frac{1 - \frac{25}{4 \cdot 3^8}}{1 + \frac{1}{35 \cdot 3^8 \cdot 2^5}} \geq 0.999.$$

Using the same argument, we can bound from above

$$|\tilde{\sigma}_N''(0)| \leq \frac{(2(N+1))^2}{8(s+1)} \frac{1 + 2(s+1)(s-1)!\frac{\zeta(s)}{(\pi s)^s}(4s+1)}{1 - s!\frac{\zeta(s)}{(\pi s)^s}},$$

$$\leq \frac{1.001}{2(s+1)}(N+1)^2.$$

By Lemma A.3 and inequality (2.7) as well as $\sin\left(\frac{\pi}{2s}\right) \leq \frac{\pi}{2s}$ and $\cos\left(\frac{\pi}{2s}\right)\left(2\cos\left(\frac{\pi}{s}\right) - 1\right) \geq \frac{3}{4}$ we have

$$|(z^4\tilde{g}_{s-1}(z))^{(s-1)}|_V \leq 2^{s-4}(35s - 19)$$

and with the same argumentation as before

$$|\tilde{\sigma}_N^{(4)}(0)| \leq \frac{3(N+1)^4}{4(s+2)(s+1)} \frac{1 + \frac{4}{3}\frac{\zeta(s)(s+2)!}{(\pi s)^s s}(35s - 19)}{1 - \frac{\zeta(s)s!}{(\pi s)^s}},$$

$$\leq \frac{3(N+1)^4}{4(s+2)(s+1)}1.001,$$

and

$$|\tilde{\sigma}_N^{(4)}(0)| \geq \frac{3(N+1)^4}{4(s+2)(s+1)}0.999.$$

In the case $s = 8$ we have again by using Lemma A.3 and inequality (2.7)

$$|(z^6\tilde{g}_7(z))^{(7)}|_V \leq 5.0896 \cdot 10^3.$$

The same argument as before shows

$$|\tilde{\sigma}_N^{(6)}(0)| \leq 1.011 \cdot \frac{15}{8} \cdot \frac{s!}{(s+3)!} \cdot (N+1)^6 = 1.011 \cdot \frac{15}{8} \cdot \frac{8!}{11!} \cdot (N+1)^6.$$

□

Having established the localization estimates for the trigonometric polynomial $\tilde{\sigma}_N$ and its derivatives, we are now able to state analog bounds for the kernel σ_N and its derivatives on the rotation group in the next section.

2.2 Localized Kernels on the Rotation Group

Since the kernel σ_N is of the form

$$\sigma_N(x, y) = \tilde{\sigma}_N(\omega(y^{-1}x)),$$

the derived estimates of the last section immediately yield

$$|\sigma_N(x, y)| \leq \frac{c_{0,s}}{((N+1)\omega(y^{-1}x))^s}, \tag{2.10}$$

which shows the localization of the kernel σ_N itself. The derivative kernels $X_n^y\sigma_N, X_i^x X_n^y \sigma_N$ and $X_j^x X_i^x X_n^y \sigma_N$ are no longer radial functions. Nevertheless, they obey analog localization estimates with the same constants as in the trigonometric case. Thereby, Theorem 2.3 provides estimates for the entries of the interpolation matrix in (1.17), whereas Lemma 2.4 and 2.5 give bounds for the entries of the Hessian.

Theorem 2.3. *We have for $s \in 2\mathbb{N}$, $s \geq 6$, $N \geq 2s$, $\omega(y^{-1}x) \geq \frac{\pi}{2(N+1)}$*

$$|X_n^y \sigma_N(x, y)| \leq \frac{c_{1,s}}{(N+1)^{s-1}\omega(y^{-1}x)^s},$$
$$|X_i^x X_n^y \sigma_N(x, y)| \leq \frac{c_{2,s}}{(N+1)^{s-2}\omega(y^{-1}x)^s},$$

and $c_{l,s}$ are the constants of Theorem 2.1.

Proof. We calculate the derivative kernel $X_1^y \sigma_N$. For $\omega(y^{-1}x) \notin \{0, \pi\}$, we have

$$\begin{aligned}
&\frac{\sigma_N(x, y\mathrm{e}^{t\mathcal{L}_1}) - \sigma_N(x, y)}{t} \\
&= \frac{\tilde{\sigma}_N(\omega(\mathrm{e}^{-t\mathcal{L}_1}y^{-1}x)) - \tilde{\sigma}_N(\omega(y^{-1}x))}{\omega(\mathrm{e}^{-t\mathcal{L}_1}y^{-1}x) - \omega(y^{-1}x)} \cdot \frac{\omega(\mathrm{e}^{-t\mathcal{L}_1}y^{-1}x) - \omega(y^{-1}x)}{\mathrm{tr}(\mathrm{e}^{-t\mathcal{L}_1}y^{-1}x) - \mathrm{tr}(y^{-1}x)} \\
&\quad \cdot \frac{\mathrm{tr}(\mathrm{e}^{-t\mathcal{L}_1}y^{-1}x) - \mathrm{tr}(y^{-1}x)}{t}.
\end{aligned}$$

The limits are given by

$$\lim_{t\to 0} \frac{\omega(\mathrm{e}^{-t\mathcal{L}_1}y^{-1}x) - \omega(y^{-1}x)}{\mathrm{tr}(\mathrm{e}^{-t\mathcal{L}_1}y^{-1}x) - \mathrm{tr}(y^{-1}x)} = \frac{1}{-2\sqrt{1 - (\frac{\mathrm{tr}(y^{-1}x)-1}{2})^2}}$$

and

$$\lim_{t\to 0} \frac{\mathrm{tr}(\mathrm{e}^{-t\mathcal{L}_1}y^{-1}x) - \mathrm{tr}(y^{-1}x)}{t} = ((y^{-1}x)_{32} - (y^{-1}x)_{23}).$$

Hence,

$$\begin{aligned}
X_1^y \sigma_N(x, y) &= \tilde{\sigma}_N'(\omega(y^{-1}x)) \frac{((y^{-1}x)_{32} - (y^{-1}x)_{23})}{-2\sqrt{1 - (\frac{\mathrm{tr}(y^{-1}x)-1}{2})^2}} \\
&= \tilde{\sigma}_N'(\omega(y^{-1}x)) \frac{((y^{-1}x)_{32} - (y^{-1}x)_{23})}{-2\sin(\omega(y^{-1}x))} = -\tilde{\sigma}_N'(\omega(y^{-1}x)) e_1,
\end{aligned}$$

where $e_1 = e_1(y^{-1}x)$ is the first component of the unit vector describing the rotation axis of $y^{-1}x$. In the same way, one can calculate

$$X_2^y \sigma_N(x,y) = \tilde{\sigma}'_N(\omega(y^{-1}x)) \frac{((y^{-1}x)_{31} - (y^{-1}x)_{13})}{2\sin(\omega(y^{-1}x))} = -\tilde{\sigma}'_N(\omega(y^{-1}x))e_2,$$
$$X_3^y \sigma_N(x,y) = \tilde{\sigma}'_N(\omega(y^{-1}x)) \frac{((y^{-1}x)_{12} - (y^{-1}x)_{21})}{2\sin(\omega(y^{-1}x))} = -\tilde{\sigma}'_N(\omega(y^{-1}x))e_3.$$

Also observe, that we have

$$X_n^x \sigma_N(x,y) = -X_n^y \sigma_N(x,y).$$

This gives

$$|X_n^y \sigma_N(x,y)| \leq |\tilde{\sigma}'_N(\omega(y^{-1}x))| \leq \frac{c_{1,s}}{(N+1)^{s-1}\omega(y^{-1}x)^s}.$$

These estimates are valid for all $x, y \in SO(3)$ with $\operatorname{tr}(y^{-1}x) \notin \{1,3\}$. We know that $X_n^y \sigma_N(x,y)$ is always a finite sum of products of Wigner D-functions, since each operator X_i maps a Wigner D-function to sums of Wigner D-functions, see e.g. [Chirikjian and Kyatkin, 2000]. Thus, we know for a fixed $x \in SO(3)$ that $X_n^y \sigma_N(x,y)$ exists for all $y \in SO(3)$ and is continuous, which means that by limit considerations the estimates above are also valid if $\omega(y^{-1}x) = \pi$. In the case $y = x$, we have by limit considerations $X_n^y \sigma_N(x,x) = \tilde{\sigma}'_N(0) = 0$. This leads to the estimate for the first derivatives of the kernel.

For the estimation of the second kind of kernel, we use the product rule and the calculations above to show

$$X_i^x X_n^y \sigma_N(x,y) = -X_i^x e_n(x,y)\tilde{\sigma}'_N(\omega(y^{-1}x)) - e_n(y^{-1}x) X_i^x(\tilde{\sigma}'_N(\omega(y^{-1}x))).$$

In the same way as before, we can show

$$X_i^x(\tilde{\sigma}'_N(\omega(y^{-1}x))) = \tilde{\sigma}''_N(\omega(y^{-1}x))e_i(y^{-1}x),$$

and consequently

$$X_i^x X_n^y \sigma_N(x,y) = -X_i^x e_n(x,y)\tilde{\sigma}'_N(\omega(y^{-1}x)) - e_n(y^{-1}x)\tilde{\sigma}''_N(\omega(y^{-1}x))e_i(y^{-1}x).$$

Thus, the only part we have to calculate is $X_i^x e_n(x,y)$. Again, at first we restrict ourselves to $\omega(y^{-1}x) \notin \{0,\pi\}$ and extend afterwards by continuity. We concentrate on the example $n = 1, i = 3$. We have

$$\begin{aligned}
&e_1(y^{-1}x\mathrm{e}^{t\mathcal{L}_3}) - e_1(y^{-1}x) \\
&= \left[\frac{(y^{-1}x\mathrm{e}^{t\mathcal{L}_3})_{32} - (y^{-1}x\mathrm{e}^{t\mathcal{L}_3})_{23}}{2\sin(\omega(y^{-1}x\mathrm{e}^{t\mathcal{L}_3}))} - \frac{(y^{-1}x)_{32} - (y^{-1}x)_{23}}{2\sin(\omega(y^{-1}x))}\right], \\
&= \frac{1}{2\sin(\omega(y^{-1}x\mathrm{e}^{t\mathcal{L}_3}))}\left[(y^{-1}x)_{32}\left(\cos(t) - \frac{\sin(\omega(y^{-1}x\mathrm{e}^{t\mathcal{L}_3}))}{\sin(\omega(y^{-1}x))}\right) + \dots\right. \\
&\left.\dots (y^{-1}x)_{23}\left(\frac{\sin(\omega(y^{-1}x\mathrm{e}^{t\mathcal{L}_3}))}{\sin(\omega(y^{-1}x))} - 1\right) - (y^{-1}x)_{31}\sin(t)\right].
\end{aligned}$$

Using the rule of L'Hôpital, we have

$$\lim_{t\to 0} \frac{\cos(t) - \frac{\sin(\omega(y^{-1}x\mathrm{e}^{t\mathcal{L}_3}))}{\sin(\omega(y^{-1}x))}}{t} = -e_3(y^{-1}x)\left(\frac{\cos(\omega(y^{-1}x))}{\sin(\omega(y^{-1}x))}\right),$$

where $e_3(y^{-1}x)$ denotes the third component of the unit vector representing the rotation axis of $y^{-1}x$. In the same way, we get

$$\lim_{t\to 0}\frac{\frac{\sin(\omega(y^{-1}x\mathrm{e}^{t\mathcal{L}_3}))}{\sin(\omega(y^{-1}x))}-1}{t} = e_3(y^{-1}x)\left(\frac{\cos(\omega(y^{-1}x))}{\sin(\omega(y^{-1}x))}\right).$$

Combining all this, we end up with

$$\begin{aligned} X_3^x e_1(x,y) &= \lim_{t\to 0} t^{-1}(e_1(y^{-1}x\mathrm{e}^{t\mathcal{L}_3}) - e_1(y^{-1}x)); \\ &= \frac{1}{2\sin(\omega(y^{-1}x))}\left[((y^{-1}x)_{23}-(y^{-1}x)_{32})e_3(y^{-1}x)\left(\frac{\cos(\omega(y^{-1}x))}{\sin(\omega(y^{-1}x))}\right)-(y^{-1}x)_{31}\right] \\ &= -e_1(y^{-1}x)e_3(y^{-1}x)\left(\frac{\cos(\omega(y^{-1}x))}{\sin(\omega(y^{-1}x))}\right)-\frac{(y^{-1}x)_{31}}{2\sin(\omega(y^{-1}x))}. \end{aligned}$$

Again, we use the *Rodrigues formula* for $(y^{-1}x)_{31} = (1-\cos(\omega))e_1e_3 - \sin(\omega)e_2$ and get

$$X_3^x e_1(x,y) = -\frac{e_1(y^{-1}x)e_3(y^{-1}x)(1+\cos(\omega(y^{-1}x)))}{2\sin(\omega(y^{-1}x))}+\frac{e_2(y^{-1}x)}{2}.$$

Similarly, we can calculate

$$X_2^x e_1(x,y) = -\frac{e_1(y^{-1}x)e_2(y^{-1}x)(1+\cos(\omega(y^{-1}x)))}{2\sin(\omega(y^{-1}x))}-\frac{e_3(y^{-1}x)}{2}$$

and

$$X_1^x e_1(x,y) = \frac{1+\cos(\omega(y^{-1}x))}{2\sin(\omega(y^{-1}x))}(1-e_1(y^{-1}x)^2).$$

For the other components of the rotation axis, the differentials are computed in the same way and are given by

$$\begin{aligned} X_1^x e_2(x,y) &= -\frac{e_1(y^{-1}x)e_2(y^{-1}x)(1+\cos(\omega(y^{-1}x)))}{2\sin(\omega(y^{-1}x))}+\frac{e_3(y^{-1}x)}{2}, \\ X_2^x e_2(x,y) &= \frac{1+\cos(\omega(y^{-1}x))}{2\sin(\omega(y^{-1}x))}(1-e_2(y^{-1}x)^2), \\ X_3^x e_2(x,y) &= -\frac{e_2(y^{-1}x)e_3(y^{-1}x)(1+\cos(\omega(y^{-1}x)))}{2\sin(\omega(y^{-1}x))}-\frac{e_1(y^{-1}x)}{2}, \\ X_1^x e_3(x,y) &= -\frac{e_1(y^{-1}x)e_3(y^{-1}x)(1+\cos(\omega(y^{-1}x)))}{2\sin(\omega(y^{-1}x))}-\frac{e_2(y^{-1}x)}{2}, \\ X_2^x e_3(x,y) &= -\frac{e_2(y^{-1}x)e_3(y^{-1}x)(1+\cos(\omega(y^{-1}x)))}{2\sin(\omega(y^{-1}x))}+\frac{e_1(y^{-1}x)}{2}, \\ X_3^x e_3(x,y) &= \frac{1+\cos(\omega(y^{-1}x))}{2\sin(\omega(y^{-1}x))}(1-e_3(y^{-1}x)^2). \end{aligned} \tag{2.11}$$

Observe, that we have

$$\sin(\omega) \geq \frac{2}{\pi}\omega, \quad \text{for } \omega \in (0,\pi/2]. \tag{2.12}$$

Accordingly, for $\omega \in (\frac{\pi}{2(N+1)}, \frac{\pi}{2}]$ we simply estimate

$$\left|\frac{1+\cos(\omega)}{\sin(\omega)}\tilde{\sigma}_N'(\omega)\right| \leq \pi\frac{|\tilde{\sigma}_N'(\omega)|}{|\omega|} \leq 2(N+1)|\tilde{\sigma}_N'(\omega)|.$$

For $\omega \in (\frac{\pi}{2}, \pi]$ we have

$$\left|\frac{1+\cos(\omega)}{\sin(\omega)}\right| = \left|\cot\left(\frac{\omega}{2}\right)\right| \leq 1$$

and consequently

$$\left|\frac{1+\cos(\omega)}{\sin(\omega)}\tilde{\sigma}_N'(\omega)\right| \leq |\tilde{\sigma}_N'(\omega)|.$$

Since $|e_i e_j| \leq \frac{1}{2}$ and $N \geq 2s \geq 12$, we have the estimate

$$\begin{aligned}|X_j^x X_i^y \sigma_N(x,y)| &\leq \left(\frac{1}{2}(N+1)+\frac{1}{2}\right)|\tilde{\sigma}_N'(\omega(y^{-1}x))| + \frac{1}{2}|\tilde{\sigma}_N''(\omega(y^{-1}x))|,\\ &\leq (N+1)|\tilde{\sigma}_N'(\omega(y^{-1}x))| + \frac{1}{2}|\tilde{\sigma}_N''(\omega(y^{-1}x))|.\end{aligned}$$

Using the localization result of Theorem 2.1 together with $c_{1,s} \leq \frac{1}{2}c_{2,s}$ results in

$$|X_j^x X_i^y \sigma_N(x,y)| \leq \frac{c_{1,s}+\frac{1}{2}c_{2,s}}{(N+1)^{s-2}\omega(y^{-1}x)^s} \leq \frac{c_{2,s}}{(N+1)^{s-2}\omega(y^{-1}x)^s}.$$

If $i = j$, we have

$$\begin{aligned}|X_i^x X_i^y \sigma_N(x,y)| &\leq (1-e_i^2)(N+1)|\tilde{\sigma}_N'(\omega(y^{-1}x))| + e_i^2|\tilde{\sigma}_N''(\omega(y^{-1}x))|,\\ &\leq \frac{(1-e_i^2)c_{1,s}+e_i^2 c_{2,s}}{(N+1)^{s-2}\omega(y^{-1}x)^s} \leq \frac{c_{2,s}}{(N+1)^{s-2}\omega(y^{-1}x)^s}.\end{aligned}$$

For $x = y$ we have

$$X_i^x X_i^y \sigma_N(x,x) = -\tilde{\sigma}_N''(0), \quad X_j^x X_i^y \sigma_N(x,x) = 0.$$

□

The derived bounds are useful for estimating the entries of the interpolation matrix. In addition, we need localization estimates for the entries of the Hessian matrix. We have to distinguish between two cases, namely $\omega(y^{-1}x)$ is well separated from zero, covered by Lemma 2.4, and $\omega(y^{-1}x)$ approaches zeros, which is handled in Lemma 2.5.

Lemma 2.4. *For $s \in 2\mathbb{N}$, $s \geq 6$, $N \geq 2s$, $\omega(y^{-1}x) \geq \frac{\pi}{2(N+1)}$, the entries of the Hessian matrix of σ_N and $X_k^y\sigma_N$ in normal coordinates, see* (1.9)*, obey*

$$\begin{aligned}|(H\sigma_N(x,y))_{ij}| &\leq \frac{c_{2,s}}{(N+1)^{s-3}\omega(y^{-1}x)^s},\\ |(HX_k^y\sigma_N(x,y))_{ij}| &\leq \frac{1.2\cdot c_{3,s}}{(N+1)^{s-3}\omega(y^{-1}x)^s}.\end{aligned}$$

Proof. The proof works in the same way as the proof of Theorem 2.3. First the derivatives are calculated directly via the product rule, then the according terms are grouped together in the right way, and an estimate is shown. For abbreviation, we suppress the dependence on x, y in the following. By the product rule we have

$$X_i^x X_i^x X_k^y \sigma_N = -X_i^x X_i^x e_k \tilde{\sigma}_N^{'} - (2e_i X_i^x e_k + e_k X_i^x e_i)\tilde{\sigma}_N^{''} - e_i^2 e_k \tilde{\sigma}_N^{'''}. \tag{2.13}$$

Suppose we have

$$X_i^x e_k = -e_i e_k \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \pm \frac{e_j}{2},$$

see (2.11), then the factor in front of $\tilde{\sigma}_N^{'}$ is calculated as

$$\begin{aligned} X_i^x X_i^x e_k = &\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)^2 (e_i^2 e_k - (1-e_i^2)e_k) + \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\left(\frac{e_i^2 e_k}{\sin(\omega)} \mp e_i e_j\right) \\ &\mp \frac{e_k}{4}, \end{aligned}$$

and the factor in front of $\tilde{\sigma}_N^{''}$ as

$$(2e_i X_i^x e_k + e_k X_i^x e_i) = \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\left(-2e_i^2 e_k + e_k(1-e_i^2)\right) \pm e_i e_j.$$

This means,

$$\begin{aligned} X_i^x X_i^x X_k^y \sigma_N = &\left((3e_k e_i^2 - e_k)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \mp e_i e_j\right)\tilde{\sigma}_N^{''}(\omega) \\ &+ \left((2e_k e_i^2 - e_k)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)^2 - e_k e_i^2 \left(\frac{1+\cos(\omega)}{2\sin^2(\omega)}\right)\right. \\ &\left.\pm e_i e_j \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \pm \frac{e_k}{4}\right)\tilde{\sigma}_N^{'}(\omega) \\ &- e_i^2 e_k \tilde{\sigma}_N^{'''}. \end{aligned} \tag{2.14}$$

For this reason, again by using (2.12) we have for $\omega \in (\frac{\pi}{2(N+1)}, \frac{\pi}{2}]$

$$\begin{aligned} |X_i^x X_i^x X_k^y \sigma_N| \le &\left(|3e_k e_i^2 - e_k|(N+1) + |e_i e_j|\right)|\tilde{\sigma}_N^{''}| + |e_i^2 e_k||\tilde{\sigma}_N^{'''}| \\ &+ \left(\left|e_k e_i^2 \cos(\omega) - e_k\left(\frac{1+\cos(\omega)}{2}\right)\right|(N+1)^2\right. \\ &\left.+ |e_i e_j|(N+1) + \frac{|e_k|}{4}\right)|\tilde{\sigma}_N^{'}|. \end{aligned}$$

Now, we use that $|3e_k e_i^2 - e_k| \le 1$, $|e_k e_i^2|, |e_i e_k^2| \le \frac{2}{3\sqrt{3}}$, $|e_i e_j| \le \frac{1}{2}$, and

$$\left|e_k e_i^2 \cos(\omega) - e_k\left(\frac{1+\cos(\omega)}{2}\right)\right| \le 1,$$

to derive for $N \ge 2s \ge 12$

$$|X_i^x X_i^x X_k^y \sigma_N| \le (N+1)1.04|\tilde{\sigma}_N^{''}| + \frac{2}{3\sqrt{3}}|\tilde{\sigma}_N^{'''}| + (N+1)^2 1.04|\tilde{\sigma}_N^{'}|.$$

With Theorem 2.1 and the observation that

$$\frac{c_{2,s}}{c_{3,s}} = \frac{4s+1}{9s-2} \leq \frac{1}{2},$$
$$\frac{c_{1,s}}{c_{3,s}} = \frac{2s}{9s-2} \leq \frac{1}{4},$$

we have

$$|X_i^x X_i^x X_k^y \sigma_N(x,y)| \leq \frac{1.2 \cdot c_{3,s}}{(N+1)^{s-3}\omega(y^{-1}x)^s}.$$

In the case $k = i$, we calculate

$$X_i^x X_i^x X_i^y \sigma_N = -3e_i(1-e_i^2)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\tilde{\sigma}_N^{''}(\omega) + 2e_i(1-e_i^2)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)^2 \tilde{\sigma}_N^{'}(\omega)$$
$$+ e_i(1-e_i^2)\left(\frac{1+\cos(\omega)}{2\sin^2(\omega)}\right)\tilde{\sigma}_N^{'}(\omega) - e_i^3\tilde{\sigma}_N^{'''}(\omega), \tag{2.15}$$

which yields

$$|X_i^x X_i^x X_i^y \sigma_N| \leq 3|e_i(1-e_i^2)|(N+1)\left(|\tilde{\sigma}_N^{''}| + (N+1)^2|\tilde{\sigma}_N^{'}|\right) + |e_i^3||\tilde{\sigma}_N^{'''}|,$$
$$\leq \frac{1}{(N+1)^{s-3}\omega^s}|e_i|\left(3(1-e_i^2)\left(c_{2,s}+c_{1,s}\right) + e_i^2 c_{3,s}\right),$$
$$\leq \frac{c_{3,s}}{(N+1)^{s-3}\omega^s}\underbrace{|e_i|(2.25-1.25e_i^2)}_{\leq 1.2} \leq \frac{1.2 \cdot c_{3,s}}{(N+1)^{s-3}\omega^s}.$$

This shows the estimate for the on-diagonal entries of the Hessian. For the off-diagonal entries observe, that we have for $n \neq j, i$ the following sign combination

$$X_i^x e_n = -e_i e_n \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \pm \frac{e_j}{2},$$
$$X_j^x e_n = -e_j e_n \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \mp \frac{e_i}{2},$$
$$X_j^x e_i = -e_j e_i \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \pm \frac{e_n}{2}.$$

This gives

$$X_i^x X_n^x \sigma_N \mp \frac{1}{2}X_j^x \sigma_N = e_i e_n \left(\tilde{\sigma}_N^{''}(\omega) - \frac{1+\cos(\omega)}{2\sin(\omega)}\tilde{\sigma}_N^{'}(\omega)\right)$$

and therefore, using (2.12) and Theorem 2.1,

$$|X_i^x X_n^x \sigma_N \mp \frac{1}{2}X_j^x \sigma_N| \leq |e_i e_n|\left(|\tilde{\sigma}_N^{''}(\omega)| + (N+1)|\tilde{\sigma}_N^{'}(\omega)|\right),$$
$$\leq \frac{|e_i e_n|}{(N+1)^{s-2}\omega^s}\left(c_{2,s}+c_{1,s}\right).$$

Since $\frac{c_{1,s}}{c_{2,s}} \leq \frac{1}{2}$ and $|e_i e_n| \leq \frac{1}{2}$, we get

$$|X_i^x X_n^x \sigma_N \mp \frac{1}{2}X_j^x \sigma_N| \leq \frac{c_{2,s}}{(N+1)^{s-2}\omega^s},$$

which shows the first inequality. For the second one, we calculate

$$\begin{aligned}X^x_j X^x_i X^y_n \sigma_N \mp \frac{1}{2} X^x_n X^y_n \sigma_N &= -X^x_j X^x_i e_n \tilde{\sigma}'_N(\omega) - \left(e_j X^x_i e_n + e_n X^x_j e_i + e_i X^x_j e_n\right) \tilde{\sigma}''_N(\omega) \\ &\quad - e_i e_j e_n \tilde{\sigma}'''_N(\omega) \mp \frac{1}{2}\left(-X^x_n e_n \tilde{\sigma}'_N(\omega) - e_n^2 \tilde{\sigma}''_N(\omega)\right), \\ &= -X^x_j X^x_i e_n \tilde{\sigma}'_N(\omega) - \left(e_j X^x_i e_n + e_n X^x_j e_i + e_i X^x_j e_n\right) \tilde{\sigma}''_N(\omega) \\ &\quad - e_i e_j e_n \tilde{\sigma}'''_N(\omega) \pm \frac{1}{2}\left((1-e_n^2)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\tilde{\sigma}'_N(\omega) + e_n^2 \tilde{\sigma}''_N(\omega)\right),\end{aligned}$$

with

$$\begin{aligned}X^x_j X^x_i e_n = 2 e_j e_i e_n \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)^2 &\mp \frac{e_n^2 + e_j^2 - e_i^2}{2}\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \\ + \frac{e_i e_j e_n}{\sin(\omega)}\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) &\pm \frac{1}{2}\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\end{aligned}$$

and

$$\left(e_j X^x_i e_n + e_n X^x_j e_i + e_i X^x_j e_n\right) = (-3 e_i e_n e_j)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \pm \frac{e_j^2 + e_n^2 - e_i^2}{2}.$$

Putting this together yields

$$\begin{aligned}X^x_j X^x_i X^y_n \sigma_N \mp \frac{1}{2} X^x_n X^y_n \sigma_N &= \left(3 e_i e_j e_n \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \pm \frac{e_i^2 - e_j^2}{2}\right) \tilde{\sigma}''_N(\omega) \\ &\quad - \left(2 e_i e_j e_n \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)^2 + e_i e_j e_n \left(\frac{1+\cos(\omega)}{2\sin^2(\omega)}\right)\right. \\ &\quad \left. \pm \frac{e_i^2 - e_j^2}{2}\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\right) \tilde{\sigma}'_N(\omega) - e_i e_n e_j \tilde{\sigma}'''_N(\omega). \qquad (2.16)\end{aligned}$$

Again using (2.12), we can estimate

$$\begin{aligned}|X^x_j X^x_i X^y_n \sigma_N \mp \frac{1}{2} X^x_n X^y_n \sigma_N| &\leq \left(3|e_i e_j e_n|(N+1) + \frac{|e_i^2 - e_j^2|}{2}\right)|\tilde{\sigma}''_N(\omega)| \\ &\quad + \left(3|e_i e_j e_n|(N+1)^2 + \frac{|e_i^2 - e_j^2|}{2}(N+1)\right)|\tilde{\sigma}'_N(\omega)| \\ &\quad + |e_i e_n e_j||\tilde{\sigma}'''_N(\omega)|.\end{aligned}$$

With Theorem 2.1 and $|e_i e_n e_j| \leq \left(\frac{1}{\sqrt{3}}\right)^3 \leq \frac{1}{5}$, as well as $\frac{c_{2,s}}{c_{3,s}} \leq \frac{1}{2}$ and $\frac{c_{1,s}}{c_{3,s}} \leq \frac{1}{4}$, we have for $N \geq 2s \geq 12$

$$|X^x_j X^x_i X^y_n \sigma_N \mp \frac{1}{2} X^x_n X^y_n \sigma_N| \leq \frac{c_{3,s}}{(N+1)^{s-3}\omega^s}.$$

In the same way one calculates

$$X^x_j X^x_i X^y_i \sigma_N \mp \frac{1}{2} X^x_n X^y_i \sigma_N = \left(e_j (3 e_i^2 - 1)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \mp \frac{e_i e_j}{2}\right) \tilde{\sigma}''_N(\omega)$$

$$- \left(2e_i^2 e_j \left(\frac{1+\cos(\omega)}{2\sin(\omega)} \right)^2 + e_j(1-e_i^2) \left(\frac{1+\cos(\omega)}{2\sin^2(\omega)} \right) \right.$$
$$\left. \mp \frac{e_i e_n}{2} \left(\frac{1+\cos(\omega)}{2\sin(\omega)} \right) \pm \frac{e_j}{4} \right) \tilde{\sigma}_N^{'}(\omega) - e_i^2 e_j \tilde{\sigma}_N^{'''}(\omega), \qquad (2.17)$$

leading to the estimate

$$|X_j^x X_i^x X_i^y \sigma_N \mp \frac{1}{2} X_n^x X_i^y \sigma_N| \leq \left(|e_j(3e_i^2 - 1)|(N+1) + \frac{|e_i e_j|}{2} \right) |\tilde{\sigma}_N^{''}(\omega)|$$
$$+ \left(|e_j(e_i^2 \cos(\omega) + 1)|(N+1)^2 \right.$$
$$\left. + \frac{|e_i e_n|}{2}(N+1) + \frac{|e_j|}{4} \right) |\tilde{\sigma}_N^{'}(\omega)| + |e_i^2 e_j||\tilde{\sigma}_N^{'''}(\omega)|.$$

We have $|e_j(3e_i^2 - 1)| \leq 1$, $|e_j(e_i^2 \cos(\omega) + 1)| \leq 1.1$, $|e_i e_j| \leq \frac{1}{2}$, $|e_i^2 e_j| \leq \frac{2}{3\sqrt{3}}$, and thus for $N \geq 2s \geq 12$

$$|X_j^x X_i^x X_i^y \sigma_N \mp \frac{1}{2} X_n^x X_i^y \sigma_N| \leq \frac{1.2 \cdot c_{3,s}}{(N+1)^{s-3}\omega^s}.$$

For the last inequality one finds

$$X_j^x X_i^x X_j^y \sigma_N \mp \frac{1}{2} X_n^x X_j^y \sigma_N = \left(e_i(3e_j^2 - 1) \left(\frac{1+\cos(\omega)}{2\sin(\omega)} \right) \mp \frac{e_n e_j}{2} \right) \tilde{\sigma}_N^{''}(\omega)$$
$$- \left((2e_i e_j^2 - e_i) \left(\frac{1+\cos(\omega)}{2\sin(\omega)} \right)^2 + e_i e_j^2 \left(\frac{1+\cos(\omega)}{2\sin^2(\omega)} \right) \right.$$
$$\left. + \frac{e_j e_n}{2} \left(\frac{1+\cos(\omega)}{2\sin(\omega)} \right) \right) \tilde{\sigma}_N^{'}(\omega) - e_i e_j^2 \tilde{\sigma}_N^{'''}(\omega). \qquad (2.18)$$

Using similar estimations as before, one shows

$$|X_j^x X_i^x X_j^y \sigma_N \mp \frac{1}{2} X_n^x X_j^y \sigma_N| \leq \frac{1.2 \cdot c_{3,s}}{(N+1)^{s-3}\omega^s}.$$

□

Lemma 2.5. *For $s \in 2\mathbb{N}$, $s \geq 8$, $N \geq 2s$, $\omega(y^{-1}x) \leq \frac{\delta}{N+1}$, $0 \leq \delta \leq \frac{\pi}{2}$ we have the following estimates*

$$\left| (H\sigma_N(x,y))_{ii} - \tilde{\sigma}_N^{''}(0) \right| \leq \frac{\tilde{d}_s}{2}(N+1)^2\delta^2,$$
$$|(H\sigma_N(x,y))_{ij}| \leq \frac{\tilde{d}_s}{4}(N+1)^2\delta^2, \quad i \neq j$$
$$|(HX_k^y\sigma_N(x,y))_{ij}| \leq \tilde{d}_s \left((N+1)^3\delta + \frac{1}{4}(N+1)^2\delta^2 \right) + \frac{\tilde{c}_s}{4}(N+1)\delta.$$

Proof. Since

$$X_i^x X_i^x \sigma_N = \frac{1+\cos(\omega)}{2\sin(\omega)}(1-e_i^2)\tilde{\sigma}_N^{'}(\omega) + e_i^2 \tilde{\sigma}_N^{''}(\omega),$$

where $\omega = \omega(y^{-1}x)$ again denotes the rotation angle and $e_i = e_i(y^{-1}x)$ denotes the i-th component of the rotation axis, we can write

$$\left(X_i^x X_i^x \sigma_N - \tilde{\sigma}_N''(0)\right) = (1 - e_i^2)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\tilde{\sigma}_N'(\omega) - \tilde{\sigma}_N''(\omega)\right) + \left(\tilde{\sigma}_N''(\omega) - \tilde{\sigma}_N''(0)\right). \tag{2.19}$$

The second term can be estimated using

$$(1 - \cos(k\omega)) \leq \frac{k^2\omega^2}{2}, \quad \omega \in [0, \frac{\delta}{N+1}]. \tag{2.20}$$

For this reason, we can estimate using Lemma 2.2

$$\begin{aligned}\left(\tilde{\sigma}_N''(\omega) - \tilde{\sigma}_N''(0)\right) &= \frac{1}{\|\tilde{g}_{s-1}\|_{1,N}} 2\sum_{k=1}^{N} \tilde{g}_s\left(\frac{k}{2(N+1)}\right) k^2(1 - \cos(k\omega)),\\ &\leq \frac{\omega^2}{2}|\tilde{\sigma}_N^{(4)}(0)| \leq \frac{\tilde{d}_s}{2}(N+1)^2\delta^2.\end{aligned}$$

We show that the first term in (2.19) is less or equal to zero and bounded in absolute value by the second term. Since

$$\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\tilde{\sigma}_N'(\omega) - \tilde{\sigma}_N''(\omega)\right) = \frac{2}{\|\tilde{g}_{s-1}\|_{1,N}}\sum_{k=1}^{N} \tilde{g}_s\left(\frac{k}{2(N+1)}\right) k^2\left(\cos(k\omega) - \frac{1+\cos(\omega)}{2k\sin(\omega)}\sin(k\omega)\right),$$

it is sufficient to show that for each $1 \leq k \leq N$

$$\left(\cos(k\omega) - \frac{1+\cos(\omega)}{2k\sin(\omega)}\sin(k\omega)\right) \leq 0, \quad \omega \in [0, \frac{\delta}{N+1}]. \tag{2.21}$$

First, observe that the lefthand side in (2.21) equals zero at $\omega = 0$. Now we show that the lefthand side is also monotonically decreasing. Its derivative is given by

$$-k\sin(k\omega) + \frac{1}{2}\left(\frac{1+\cos(\omega)}{\sin(\omega)}\right)\left(\frac{\sin(k\omega)}{k\sin(\omega)} - \cos(k\omega)\right). \tag{2.22}$$

To proceed, we show that for each $1 \leq k \leq N$

$$\left(\frac{1+\cos(\omega)}{\sin(\omega)}\right)\left(\frac{\sin(k\omega)}{k\sin(\omega)} - \cos(k\omega)\right) \leq k\sin(k\omega). \tag{2.23}$$

On the interval $[0, \frac{\delta}{N+1}]$ this is equivalent to

$$k\cos(\omega) - \frac{\cos(k\omega)\sin(\omega)}{\sin(k\omega)} \leq k - \frac{1}{k}. \tag{2.24}$$

The function on the left hand side equals $k - \frac{1}{k}$ for $\omega = 0$. To get the desired estimate we show that the function on the lefthand side of (2.24) attains its maximum on the interval $[0, \frac{\delta}{N+1}]$ at $\omega = 0$. The derivative of the left hand side of (2.24) is given by

$$-k\sin(\omega) + \frac{k\sin(\omega)}{\sin^2(k\omega)} - \frac{\cos(k\omega)\cos(\omega)}{\sin(k\omega)} = k\cot(k\omega)\sin(\omega)\left(\frac{\cos(k\omega)}{\sin(k\omega)} - \frac{\cos(\omega)}{k\sin(\omega)}\right).$$

We have $\frac{k\sin(\omega)}{\cos(\omega)} < \frac{\sin(k\omega)}{\cos(k\omega)}$, due to the power series representation of the tangent function, and accordingly

$$\left(\frac{\cos(k\omega)}{\sin(k\omega)} - \frac{\cos(\omega)}{k\sin(\omega)}\right) < 0. \tag{2.25}$$

This means the function given by the left hand side of (2.24) is strictly monotonic decreasing on the interval $[0, \frac{\delta}{N+1}]$. Thus, it attains its maximum at $\omega = 0$. Hence, the function in (2.22) is strictly negative, which implies that the inequality (2.21) holds. The first term of (2.19) can be bounded in absolute value by

$$\left|\frac{1+\cos(\omega)}{2\sin(\omega)}\tilde{\sigma}_N^{'}(\omega) - \tilde{\sigma}_N^{''}(\omega)\right| = \frac{2}{\|\tilde{g}_{s-1}\|_{1,N}} \sum_{k=1}^{N} \tilde{g}_s\left(\frac{k}{2(N+1)}\right) k^2 \left|\cos(k\omega) - \frac{1+\cos(\omega)}{2k\sin(\omega)}\sin(k\omega)\right|.$$

As seen before in (2.21), we already know that

$$\begin{aligned}\left|\cos(k\omega) - \frac{1+\cos(\omega)}{2k\sin(\omega)}\sin(k\omega)\right| &= \left(\frac{1+\cos(\omega)}{2k\sin(\omega)}\sin(k\omega) - \cos(k\omega)\right), \\ &= \frac{1+\cos(\omega)}{2k\sin(\omega)}\sin(k\omega) - \cos(k\omega).\end{aligned}$$

Since $\sin(k\omega) \leq k\sin(\omega)$, we see

$$\left|\cos(k\omega) - \frac{1+\cos(\omega)}{2k\sin(\omega)}\sin(k\omega)\right| \leq 1 - \cos(k\omega) \leq \frac{k^2\omega^2}{2}, \quad \omega \in [0, \frac{\delta}{N+1}], \tag{2.26}$$

which shows

$$\left|X_i^x X_i^x \sigma_N - \tilde{\sigma}_N^{''}(0)\right| \leq \frac{\tilde{d}_s}{2}(N+1)^2\delta^2.$$

Moreover,

$$\left|\frac{1+\cos(\omega)}{2\sin(\omega)}\tilde{\sigma}_N^{'}(\omega) - \tilde{\sigma}_N^{''}(\omega)\right| \leq \frac{\tilde{d}_s}{2}(N+1)^2\delta^2, \quad \omega \in [0, \frac{\delta}{N+1}]. \tag{2.27}$$

Similarly, we have

$$X_i^x X_n^x \sigma_N \mp \frac{1}{2} X_j^x \sigma_N = e_i e_n \left(\tilde{\sigma}_N^{''}(\omega) - \frac{1+\cos(\omega)}{2\sin(\omega)}\tilde{\sigma}_N^{'}(\omega)\right),$$

which yields, since $|e_i e_j| \leq \frac{1}{2}$,

$$|X_i^x X_n^x \sigma_N \mp X_j^x \sigma_N| \leq \frac{\tilde{d}_s}{4}(N+1)^2\delta^2.$$

For the third mixed derivatives one has in the case $n \neq i$

$$\begin{aligned}X_i^x X_i^x X_n^y \sigma_N = &\,(2e_n e_i^2 - e_n)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\left(\tilde{\sigma}_N^{''}(\omega) - \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\tilde{\sigma}_N(\omega)\right) \\ &+ e_n e_i^2 \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\left(\tilde{\sigma}_N^{''}(\omega) - \frac{\tilde{\sigma}_N(\omega)}{\sin(\omega)}\right) \\ &\pm e_i e_j \left(\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\tilde{\sigma}_N(\omega) - \tilde{\sigma}_N^{''}(\omega)\right) - e_i^2 e_n \tilde{\sigma}_N^{'''} \pm \frac{e_n}{4}\tilde{\sigma}_N'(\omega),\end{aligned}$$

as seen already in (2.14). Since $\frac{1+\cos(\omega)}{2} \leq 1$, we have

$$\left(\frac{1+\cos(\omega)}{\sin(\omega)}\right)\left(\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\frac{\sin(k\omega)}{k}-\cos(k\omega)\right) \leq \left(\frac{1+\cos(\omega)}{\sin(\omega)}\right)\left(\frac{\sin(k\omega)}{k\sin(\omega)}-\cos(k\omega)\right), \tag{2.28}$$
$$\leq k\sin(k\omega),$$

see (2.23). Consequently, since $|\frac{e_n}{2}|(|2e_i^2-1|+3e_i^2) \leq 1$, this yields

$$|X_i^x X_i^x X_n^y \sigma_N| \leq \tilde{d}_s\left((N+1)^3\delta + \frac{1}{4}(N+1)^2\delta^2\right) + \frac{\tilde{c}_s}{4}(N+1)\delta.$$

In the case $n = i$ we have, see (2.15),

$$X_i^x X_i^x X_i^y \sigma_N = 2e_i(1-e_i^2)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\left(\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\tilde{\sigma}_N(\omega) - \tilde{\sigma}_N''(\omega)\right)$$
$$+ e_i(1-e_i^2)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\left(\frac{\tilde{\sigma}_N(\omega)}{\sin(\omega)} - \tilde{\sigma}_N''(\omega)\right) - e_i^3\tilde{\sigma}_N'''(\omega).$$

Using $\frac{3}{2}|e_i|(1-e_i^2)+|e_i^3| \leq 1$, this results in

$$|X_i^x X_i^x X_i^y \sigma_N| \leq \tilde{d}_s(N+1)^3\delta.$$

Observe that we have for $n \neq j, i$ the following sign combination

$$X_i^x e_n = -e_i e_n\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \pm \frac{e_j}{2},$$
$$X_j^x e_n = -e_j e_n\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \mp \frac{e_i}{2},$$
$$X_j^x e_i = -e_j e_i\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right) \pm \frac{e_n}{2}.$$

Accordingly, as seen in (2.16),

$$X_j^x X_i^x X_n^y \sigma_N \mp \frac{1}{2}X_n^x X_n^y \sigma_N = 2e_i e_j e_n\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\left(\tilde{\sigma}_N''(\omega) - \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\tilde{\sigma}_N'(\omega)\right)$$
$$+ e_i e_j e_n\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\left(\tilde{\sigma}_N''(\omega) - \frac{\tilde{\sigma}_N'(\omega)}{\sin(\omega)}\right) - e_i e_n e_j \tilde{\sigma}_N'''(\omega)$$
$$\pm \frac{e_i^2 - e_j^2}{2}\left(\tilde{\sigma}_N''(\omega) - \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\tilde{\sigma}_N'(\omega)\right),$$

so we can estimate using (2.28) together with $|e_i e_n e_j| \leq \left(\frac{1}{\sqrt{3}}\right)^3 \leq \frac{1}{5}$

$$\left|X_j^x X_i^x X_n^y \sigma_N \mp \frac{1}{2}X_n^x X_n^y \sigma_N\right| \leq \tilde{d}_s\left(\frac{1}{2}(N+1)^3\delta + \frac{1}{4}(N+1)^2\delta^2\right).$$

Similarly, we have, see (2.17) and (2.18),

$$X_j^x X_i^x X_i^y \sigma_N \mp \frac{1}{2}X_n^x X_i^y \sigma_N = 2e_i^2 e_j\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\left(\tilde{\sigma}_N''(\omega) - \left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\tilde{\sigma}_N'(\omega)\right)$$

$$- e_j(1-e_i^2)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\left(\tilde{\sigma}_N''(\omega)-\frac{\tilde{\sigma}_N'(\omega)}{\sin(\omega)}\right)-e_i^2e_j\tilde{\sigma}_N'''(\omega)$$
$$\mp\frac{e_ie_n}{2}\left(\tilde{\sigma}_N''(\omega)-\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\tilde{\sigma}_N'(\omega)\right)\mp\frac{e_j}{4}\tilde{\sigma}_N'(\omega)$$

and

$$\begin{aligned}X_j^xX_i^xX_j^y\sigma_N \mp \frac{1}{2}X_n^xX_j^y\sigma_N = {} & (2e_ie_j^2-e_i)\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\left(\tilde{\sigma}_N''(\omega)-\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\tilde{\sigma}_N(\omega)\right)\\ & + e_ie_j^2\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\left(\tilde{\sigma}_N''(\omega)-\frac{\tilde{\sigma}_N(\omega)}{\sin(\omega)}\right)-e_ie_j^2\tilde{\sigma}_N'''(\omega)\\ & \pm\frac{e_ne_j}{2}\left(\tilde{\sigma}_N''(\omega)-\left(\frac{1+\cos(\omega)}{2\sin(\omega)}\right)\tilde{\sigma}_N(\omega)\right),\end{aligned}$$

which yields

$$\left|X_j^xX_i^xX_i^y\sigma_N \mp \frac{1}{2}X_n^xX_i^y\sigma_N\right| \le \tilde{d}_s\left((N+1)^3\delta+\frac{1}{8}(N+1)^2\delta^2\right)+\frac{c_s}{4}(N+1)\delta,$$
$$\left|X_j^xX_i^xX_j^y\sigma_N \mp \frac{1}{2}X_n^xX_j^y\sigma_N\right| \le \tilde{d}_s\left((N+1)^3\delta+\frac{1}{8}(N+1)^2\delta^2\right).$$

□

The previous statements, i.e. Theorem 2.3, Lemma 2.4 and Lemma 2.5, give pointwise control of the absolute value of the interpolation kernel and its various derivatives. The last lemma of this section provides bounds for summing up those pointwise expressions. In combination with Theorem 2.3 and Lemma 2.4, we use it in the next section to bound off-diagonal entries of the interpolation and the Hessian matrix.

Lemma 2.6. *Let $x_j \in \mathcal{X}$, where $\mathcal{X} \subset SO(3)$ is a discrete set obeying a separation condition of $\rho(\mathcal{X}) \ge \frac{\nu}{N+1}$ with $\nu \ge \pi$, and let $x \in SO(3)$ such that $d(x,x_j) \le \varepsilon\frac{\nu}{N+1}$, for $0 \le \varepsilon \le 1/2$. Suppose $f : SO(3)\times SO(3) \to \mathbb{C}$ obeys,*

$$|f(x,y)| \le \frac{c_f}{((N+1)\cdot\omega(y^{-1}x))^s},$$

for $x \ne y$ and $s > 3$, then

$$\sum_{x_i\in\mathcal{X}\setminus x_j}|f(x,x_i)| \le \frac{c_fa_\varepsilon}{\nu^s},$$

where $a_\varepsilon = \zeta(s-2)\cdot\min\{27\cdot(1-\varepsilon)^{-s}+124, 124\cdot(1-\varepsilon)^{-s}\}$. Here ζ denotes the Riemannian Zeta function.

Proof. For $x \in SO(3)$, with $d(x,x_j) \le \varepsilon\frac{\nu}{N+1}$ for some $x_j \in \mathcal{X}$, we define the ring about x by

$$\mathcal{S}_m := \{y \in SO(3) : \frac{\nu m}{N+1} \le d(x,y) \le \frac{\nu(m+1)}{N+1}\},$$

for $m \in \mathbb{N}$. By definition we have $\mathcal{S}_m = \emptyset$ for $\frac{m\nu}{N+1} > \pi$. Moreover, as shown in [Schmid, 2009] we can estimate the number of elements in the intersection of $\mathcal{S}_m$ with the set $\mathcal{X}\setminus\{x_j\}$ for $m \ge 1$ by

$$\mathrm{card}(\mathcal{X}\setminus\{x_j\}\cap\mathcal{S}_m) \le 48m^2+48m+28 \le 124m^2.$$

Observe, that $B_{\frac{\nu}{2(N+1)}}(x_i) \cap B_{\frac{\nu}{2(N+1)}}(x_n) = \emptyset$ for $x_i, x_n \in \mathcal{X} \setminus \{x_j\} \cap \mathcal{S}_0$ and

$$\bigcup_{x_i \in \mathcal{X} \setminus \{x_j\} \cap \mathcal{S}_0} B_{\frac{\nu}{2(N+1)}}(x_i) \subseteq B_{\frac{3\nu}{2(N+1)}}(x).$$

By the translation invariance of the Haar-measure λ, this shows

$$\operatorname{card}(\mathcal{X} \setminus \{x_j\} \cap \mathcal{S}_0) \leq \frac{\lambda\left(B_{\frac{3\nu}{2(N+1)}}(e)\right)}{\lambda\left(B_{\frac{\nu}{2(N+1)}}(e)\right)} = \frac{\frac{3\nu}{2(N+1)} - \sin(\frac{3\nu}{2(N+1)})}{\frac{\nu}{2(N+1)} - \sin(\frac{\nu}{2(N+1)})}.$$

We derive the following bound

$$\frac{3r - \sin(3r)}{r - \sin(r)} \leq 27$$

or equivalently $\sin(3r) - 27\sin(3r) + 24r \geq 0$ for $r \in [0, \pi]$. Observe, that the lefthandside equals zero for $r = 0$. For $0 < r \leq \pi$ the derivative obeys

$$3\cos(3r) - 27\cos(r) + 24 = 48\sin^4\left(\frac{r}{2}\right)(\cos(r) + 2) \geq 0,$$

which shows

$$\operatorname{card}(\mathcal{X} \setminus \{x_j\} \cap \mathcal{S}_0) \leq 27.$$

Since $d(x, x_j) \leq \varepsilon\frac{\nu}{N+1}$, we have $d(x, x_i) \geq \frac{(1-\varepsilon)\nu}{N+1}$ for $x_i \in \mathcal{X} \setminus \{x_j\} \cap \mathcal{S}_0$. Using this and the locality result (2.3), we can estimate for $s \geq 4$

$$\begin{aligned}
\sum_{x_i \in \mathcal{X} \setminus x_j} |f(x, x_i)| &\leq \sum_{x_i \in (\mathcal{X} \setminus x_j) \cap \mathcal{S}_0} \frac{c_f}{((N+1) \cdot d(x, x_i))^s} + \sum_{m=1}^{\infty} \sum_{x_i \in (\mathcal{X} \setminus x_j) \cap \mathcal{S}_m} \frac{c_f}{((N+1) \cdot d(x, x_i))^s}, \\
&\leq \frac{27c_f(1-\varepsilon)^{-s}}{\nu^s} + 124c_f \sum_{m=1}^{\infty} \frac{m^2}{(m\nu)^s}, \\
&\leq \frac{27c_f(1-\varepsilon)^{-s}}{\nu^s} + \frac{124c_f}{\nu^s} \sum_{m=1}^{\infty} \frac{1}{m^{s-2}}, \\
&\leq \frac{(27(1-\varepsilon)^{-s} + 124)c_f\zeta(s-2)}{\nu^s},
\end{aligned}$$

where the last inequality follows by the definition of the Zeta function. On the other hand, we can define the rings around x_j again by

$$\tilde{\mathcal{S}}_m := \{y \in SO(3) : \frac{(1-\varepsilon)\nu m}{N+1} \leq d(x_j, y) \leq \frac{(1-\varepsilon)\nu(m+1)}{N+1}\}.$$

Since $d(x, x_j) \leq \varepsilon\frac{\nu}{N+1}$, we have $d(x, x_j) \leq \varepsilon d(x_i, x_j)$ for $x_i \in (\mathcal{X} \setminus x_j) \cap \tilde{\mathcal{S}}_m$ and consequently $d(x, x_i) \geq d(x_i, x_j) - d(x, x_j) \geq \frac{(1-\varepsilon)\nu m}{N+1}$. Using this and the locality result (2.3), we can estimate for $s \geq 4$

$$\begin{aligned}
\sum_{x_i \in \mathcal{X} \setminus x_j} |f(x, x_i)| &\leq \sum_{m=1}^{\infty} \sum_{x_i \in (\mathcal{X} \setminus x_j) \cap \tilde{\mathcal{S}}_m} \frac{c_f}{((N+1)d(x, x_i))^s}, \\
&\leq 124c_f \sum_{m=1}^{\infty} \frac{m^2}{(1-\varepsilon)^s(m\nu)^s},
\end{aligned}$$

$$\leq \frac{124c_f}{(1-\varepsilon)^s\nu^s}\sum_{m=1}^{\infty}\frac{1}{m^{s-2}} = \frac{124c_f\zeta(s-2)}{(1-\varepsilon)^s\nu^s}.$$

□

The derived bounds for the kernel and its various derivatives are used in the next chapter to construct and validate a dual certificate, that ensures the unique solvability of the minimization problem

$$\min_{\mu\in\mathcal{M}(SO(3),\mathbb{R})}\|\mu\|_{TV}, \quad \text{subject to} \quad \mathcal{S}_N^*\mu = \mathcal{S}_N^*\mu^\star. \tag{RP}$$

As discussed in Section 1.2.2, the dual certificate is constructed as the solution of a Hermite-type interpolation problem, which we see in more detail in the next chapter.

Notes and References. *Localization estimates for trigonometric polynomials with coefficients generated by sampling a smooth function are well known, see e.g. [Mhaskar and Prestin, 2000] and [Kunis and Potts, 2007]. The paradigm of building localized kernels from orthogonal function systems, that obey estimates similar to* (2.1)*, is valid in very general settings, see [Filbir and Mhaskar, 2010]. Nevertheless, those kernels only allow for asymptotic estimates, as the appearing constants are not known explicitly and the behavior near the diagonal is not clear in general.*

Localization estimates with application to stability results in scattered data interpolation on the rotation group can be found in [Gräf and Kunis, 2008], where also weights generated by sampling Spline functions are considered. The contribution of this chapter is to provide analog estimates for the various derivatives *of a kernel of this form and to get pointwise bounds near the diagonal.*

Chapter 3

Dual Certificate and Error Estimates

The aim of this chapter is to construct a dual certificate, i.e. a function $q \in \Pi_N(SO(3))$, such that for a given set of points $\mathcal{X} = \{x_1, \dots, x_M\} \subset SO(3)$ and a given sign $u(x) \in \{-1, 1\}$, we have

$$\begin{aligned} q(x_i) &= u(x_i), \quad x_i \in \mathcal{X}, \\ |q(x)| &< 1, \quad x \in SO(3) \setminus \mathcal{X}. \end{aligned} \tag{3.1}$$

As seen in Section 1.1, constructing such a dual certificate for all possible signs is a sufficient criteria for a signed measure $\mu^\star$ with $\operatorname{supp}(\mu^\star) = \mathcal{X}$ to be the unique solution of the minimization

$$\min_{\mu \in \mathcal{M}(SO(3),\mathbb{R})} \|\mu\|_{TV}, \quad \text{subject to} \quad \mathcal{S}_N^* \mu = \mathcal{S}_N^* \mu^\star. \tag{RP}$$

Since we have to find an interpolating function $q \in \Pi_N(SO(3))$ for each possible sign, the support points $x_1, \dots, x_M$ cannot get arbitrarily close. We demand a *minimal separation* on the support points, i.e. a condition of the form

$$\rho(\mathcal{X}) = \min_{x_i, x_j \in \mathcal{X}, x_i \neq x_j} \omega(x_i^{-1} x_j) \geq \frac{\nu}{N+1}, \tag{3.2}$$

where ν is a given constant. In Section 3.1, we construct a candidate for a dual certificate using a Hermite interpolation under the separation assumption (3.2). We see in Section 3.2, that $\nu = 36$ is a sufficient assumption for the candidate function being a dual certificate, i.e. obeying the conditions (3.1).

We close this chapter considering the case, that we cannot access the low frequency information $\mathcal{S}_N^* \mu^\star$ exactly, but only corrupted by noise of the form

$$\mathcal{S}_N^*(\mu^\star + \eta),$$

with $\|\mathcal{S}^* \eta\|_{L^2(SO(3))} \leq \varepsilon$. In this case, we solve the Thikonov type minimization problem

$$\min_{\mu \in \mathcal{M}(SO(3),\mathbb{R})} \frac{1}{2} \|\mathcal{S}_N^*(\mu - \mu^\star - \eta)\|_{L^2(SO(3))}^2 + \tau \|\mu\|_{TV}. \tag{RP$_\tau$}$$

For a polynomial operator $\mathcal{K}_L$ mapping to $\Pi_L(SO(3))$ for $L \geq N$, i.e.

$$\mathcal{K}_L f(x) = \int_{SO(3)} f(y) K_L(x, y) \mathrm{d}\lambda(y),$$

choosing $\tau = \varepsilon$ yields bounds of the form

$$\|\mathcal{K}_L^*(\mu_\tau - \mu^\star)\|_\infty \leq C \cdot \|K_L\|_\infty \cdot s_{L,N}^2 \cdot \tau,$$

where μ_τ is the unique solution of the problem (RP$_\tau$) and $s_{L,N} = \frac{L}{N}$ is called *super-resolution factor*. This is the content of Section 3.3.

3.1 Solution of the Interpolation Problem

In this section, we construct a candidate for a dual certificate, i.e. a function $q \in \Pi_N(SO(3))$, such that the conditions

$$\begin{aligned} q(x_i) &= u(x_i), \quad x_i \in \mathcal{X}, \\ |q(x)| &< 1, \quad x \in SO(3) \setminus \mathcal{X} \end{aligned}$$

are fulfilled. Heuristically, the interpolation points are local extrema of the function q, which means that the first derivatives have to vanish at an interpolation point. We consequently solve the Hermite interpolation problem

$$\begin{aligned} q(x_j) &= u_j, \\ X_1 q(x_j) = X_2 q(x_j) &= X_3 q(x_j) = 0, \end{aligned}$$

for $j = 1, \ldots, M$. To find a solution of the Hermite interpolation problem in the space $\Pi_N(SO(3))$, we determine coefficients $\alpha_{j,0}, \alpha_{j,1}, \alpha_{j,2}, \alpha_{j,3}$ for $j = 1, \ldots, M$ in the kernel expansion

$$q(x) = \sum_{j=1}^{M} \alpha_{j,0}\sigma_N(x, x_j) + \alpha_{j,1} X_1^y \sigma_N(x, x_j) + \alpha_{j,2} X_2^y \sigma_N(x, x_j) + \alpha_{j,3} X_3^y \sigma_N(x, x_j),$$

satisfying

$$K\alpha := \begin{pmatrix} \sigma_N & X_1^x\sigma_N & X_2^x\sigma_N & X_3^x\sigma_N \\ X_1^y\sigma_N & X_1^x X_1^y\sigma_N & X_2^x X_1^y\sigma_N & X_3^x X_1^y\sigma_N \\ X_2^y\sigma_N & X_1^x X_2^y\sigma_N & X_2^x X_2^y\sigma_N & X_3^x X_2^y\sigma_N \\ X_3^y\sigma_N & X_1^x X_3^y\sigma_N & X_2^x X_3^y\sigma_N & X_3^x X_3^y\sigma_N \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} u \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{3.3}$$

where the entries in the matrix corresponds to blocks of the form $\sigma_N = (\sigma_N(x_i, x_j))_{i,j=1}^M$ and in the same way for the derivatives. The blocks in the vectors are given by $\alpha_k = (\alpha_{k,j})_{j=1}^M$, for $k = 0, 1, 2, 3$, and $u = (u_j)_{j=1}^M$. In the case this matrix is invertible, we have that q satisfies the Hermite interpolation conditions. Moreover, by construction of the kernel σ_N, the function q is always a polynomial of degree at most N. For abbreviation, we write

$$\sigma_{ij} = X_i^x X_j^y \sigma_N, \quad i, j = 1, \ldots, 3.$$

We have to show, that the block matrix

$$K = \begin{pmatrix} K_0 & \tilde{K}_1 \\ K_1 & K_2 \end{pmatrix},$$

with blocks given by

$$\begin{aligned} K_0 &= \sigma_{00} = \sigma_N, \\ K_1 &= \begin{bmatrix} \sigma_{01} & \sigma_{02} & \sigma_{03} \end{bmatrix}^T = \begin{bmatrix} X_1^y\sigma_N & X_2^y\sigma_N & X_3^y\sigma_N \end{bmatrix}^T, \\ \tilde{K}_1 &= \begin{bmatrix} \sigma_{10} & \sigma_{20} & \sigma_{30} \end{bmatrix} = \begin{bmatrix} X_1^x\sigma_N & X_2^x\sigma_N & X_3^x\sigma_N \end{bmatrix}, \\ K_2 &= \begin{bmatrix} \sigma_{11} & \sigma_{21} & \sigma_{31} \\ \sigma_{12} & \sigma_{22} & \sigma_{32} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} = \begin{bmatrix} X_1^x X_1^y\sigma_N & X_2^x X_1^y\sigma_N & X_3^x X_1^y\sigma_N \\ X_1^x X_2^y\sigma_N & X_2^x X_2^y\sigma_N & X_3^x X_2^y\sigma_N \\ X_1^x X_3^y\sigma_N & X_2^x X_3^y\sigma_N & X_3^x X_3^y\sigma_N \end{bmatrix} \end{aligned}$$

is invertible. To do this, we use an two step block inversion to show that both the matrix K_2 and its Schur complement $K/K_2 = K_0 - \tilde{K}_1 K_2^{-1} K_1$ are invertible. To show the invertibility of K_2, we split up K_2 in the first step furthermore into blocks as

$$K_2 = \begin{pmatrix} K_{2,0} & \tilde{K}_{2,1} \\ K_{2,1} & K_{2,2} \end{pmatrix},$$

with

$$\begin{aligned} K_{2,0} &= \sigma_{11}, \\ K_{2,1} &= \begin{bmatrix} \sigma_{12} & \sigma_{13} \end{bmatrix}^T, \\ \tilde{K}_{2,1} &= \begin{bmatrix} \sigma_{21} & \sigma_{31} \end{bmatrix}, \\ K_{2,2} &= \begin{bmatrix} \sigma_{22} & \sigma_{32} \\ \sigma_{23} & \sigma_{33} \end{bmatrix}. \end{aligned}$$

This shows that K_2 is invertible, if $K_{2,2}$ is invertible and its Shur complement in K_2 given by

$$S = K_2/K_{2,2} = K_{2,0} - \tilde{K}_{2,1} K_{2,2}^{-1} K_{2,1}$$

is invertible. For the invertibility of $K_{2,2}$, we proof the invertibility of

$$\sigma_{33} = X_3^x X_3^y \sigma_N$$

and its Schur complement in $K_{2,2}$ given by

$$T = K_{2,2}/\sigma_{33} = \sigma_{22} - \sigma_{32} \left(\sigma_{33}\right)^{-1} \sigma_{23}.$$

Having this, we go backwards determining the inverse of K_2 and in the end of K. For this purpose, we use that a matrix A is invertible if

$$\|I - A\|_\infty < 1,$$

where $\|A\|_\infty = \max_i \sum_j |a_{i,j}|$. In this case the norm of the inverse is bounded by

$$\|A^{-1}\|_\infty \leq \frac{1}{1 - \|I - A\|_\infty},$$

see Appendix C. In the following Lemma we bound the norms of the corresponding entries in the kernel matrix K.

Lemma 3.1. *If the separation condition* (3.2) *is satisfied, we have for any* $s \geq 6$ *even,* $N \geq 2s$ *with* $C_{i,s} = 124 c_{i,s} \zeta(s-2)$ *and* $c_s = \frac{0.999}{2(s+1)}$ *the estimates*

$$\begin{aligned} &\|I - \sigma_{00}\|_\infty \leq \frac{C_{0,s}}{\nu^s}, \quad \left\|\sigma_{00}^{-1}\right\|_\infty \leq \frac{1}{1 - \frac{C_{0,s}}{\nu^s}}, \\ &\|\sigma_{0i}\|_\infty, \quad \|\sigma_{i0}\|_\infty \leq \frac{C_{1,s}(N+1)}{\nu^s}, \quad \|\sigma_{ij}\|_\infty \leq \frac{C_{2,s}(N+1)^2}{\nu^s}, \quad \textit{for } i \neq j, i, j \neq 0, \\ &\left\|-\tilde{\sigma}_N''(0) I - \sigma_{ii}\right\|_\infty \leq \frac{C_{2,s}(N+1)^2}{\nu^s}, \quad \left\|\sigma_{ii}^{-1}\right\|_\infty \leq \frac{1}{c_s(N+1)^2\left(1 - \frac{C_{2,s}}{c_s \nu^s}\right)}, \end{aligned}$$

where the constants $c_{i,s}$ are given in Theorem 2.1.

Proof. The proof follows directly from applying Lemma 2.6 together with the bound for $|\tilde{\sigma}_N''(0)|$ given in Lemma 2.2. □

Theorem 3.2. *Suppose the separation condition* (3.2) *is satisfied for some* $\nu \geq \pi$*, such that for* $s \geq 8$ *even and* $N \geq 2s$*, there is a constant* $b > 4$*, with*

$$\nu^s \geq b \frac{C_{2,s}}{c_s}, \tag{3.4}$$

where the constant c_s *is given in Lemma 2.2 and* $C_{2,s}$ *in Lemma 3.1. Then the matrix* K *is invertible and for* $u, b_1, b_2, b_3 \in \mathbb{C}^{M\times 1}$ *such that*

$$\begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = K^{-1} \begin{pmatrix} u \\ b_1 \\ b_2 \\ b_3 \end{pmatrix},$$

we have

$$\begin{aligned}
\|\alpha_0\|_\infty &\leq C(b,s) \left(4(b-3)\|u\|_\infty + \frac{2}{(N+1)} (\|b_1\|_\infty + \|b_2\|_\infty + \|b_3\|_\infty) \right), \\
\|\alpha_1\|_\infty &\leq C(b,s) \left(\frac{2\|u\|_\infty}{(N+1)} + \frac{4}{c_s(N+1)^2} \Big((b-2)\|b_1\|_\infty + \|b_2\|_\infty + \|b_3\|_\infty \Big) \right), \\
\|\alpha_2\|_\infty &\leq C(b,s) \left(\frac{2\|u\|_\infty}{(N+1)} + \frac{4}{c_s(N+1)^2} \Big((b-2)\|b_2\|_\infty + \|b_1\|_\infty + \|b_3\|_\infty \Big) \right), \\
\|\alpha_3\|_\infty &\leq C(b,s) \left(\frac{2\|u\|_\infty}{(N+1)} + \frac{4}{c_s(N+1)^2} \Big((b-2)\|b_3\|_\infty + \|b_2\|_\infty + \|b_3\|_\infty \Big) \right),
\end{aligned}$$

with $C(b,s) = \frac{1}{4(b-3)-c_s}$*. Moreover, if* $|u_i| = 1$*,* $\|u\|_\infty \leq 1$ *and* $b_1 = b_2 = b_3 = 0$ *we have the bound*

$$|\alpha_{0,i}| \geq 1 - \frac{c_s}{4(b-3)-c_s}.$$

Proof. In this proof, the quotient $\frac{C_{2,s}}{c_s \nu^s}$ appears quite often, so we will denote it for abbreviation by

$$a_1 := \frac{C_{2,s}}{c_s \nu^s}.$$

It represents the quotient of the off-diagonal upper bound and the on-diagonal lower bound. The assumption of the theorem then reads as

$$a_1 \leq \frac{1}{b},$$

with $b > 4$. Oberve, that we automatically have $b > 3 + \frac{c_s}{2}$ for all $s \in 2\mathbb{N}$. Using Lemma 3.1, we can estimate

$$\left\|\sigma_{33}^{-1}\right\|_\infty \leq \frac{1}{c_s(N+1)^2\,(1-a_1)}$$

and

$$\begin{aligned}
\left\|\tilde{\sigma}_N''(0)I - K_{2,2}/\sigma_{33}\right\|_\infty &\leq \left\|\tilde{\sigma}_N''(0)I - \sigma_{22}\right\|_\infty + \|\sigma_{32}\|_\infty \left\|\sigma_{33}^{-1}\right\|_\infty \|\sigma_{23}\|_\infty, \\
&\leq \frac{C_{2,s}(N+1)^2}{\nu^s} \left(1 + \frac{C_{2,s}}{c_s\nu^s - C_{2,s}}\right).
\end{aligned}$$

This means

$$\left\| I - \frac{K_{2,2}/\sigma_{33}}{\tilde{\sigma}_N''(0)} \right\|_\infty \leq \frac{1}{|\tilde{\sigma}_N''(0)|} \frac{C_{2,s}(N+1)^2}{\nu^s} \left(1 + \frac{C_{s,2}}{c_s\nu^s - C_{2,s}}\right).$$

For the expression on the right hand side, we get

$$\frac{C_{2,s}}{c_s\nu^s}\left(1+\frac{C_{2,s}}{c_s\nu^s - C_{2,s}}\right) = \frac{C_{2,s}}{c_s\nu^s - C_{2,s}} = \frac{a_1}{1-a_1}. \tag{3.5}$$

Since $a_1 < \frac{1}{b}$ with $b > 3 + \frac{c_s}{2}$, we have the bound

$$\frac{a_1}{1-a_1} \leq \frac{1}{b-1} < 1 \tag{3.6}$$

and consequently

$$\left\|(K_{2,2}/\sigma_{33})^{-1}\right\|_\infty \leq \frac{(N+1)^{-2}(b-1)}{c_s(b-2)}.$$

This shows the invertibility of $K_{2,2}$. Accordingly, this yields with $T = K_{2,2}/\sigma_{33}$ the representation

$$(K_{2,2})^{-1} = \begin{pmatrix} T^{-1} & -T^{-1}\sigma_{32}\left(\sigma_{33}\right)^{-1} \\ -\left(\sigma_{33}\right)^{-1}\sigma_{23}T^{-1} & \left(\sigma_{33}\right)^{-1} + \left(\sigma_{33}\right)^{-1}\sigma_{23}T^{-1}\sigma_{32}\left(\sigma_{33}\right)^{-1} \end{pmatrix}. \tag{3.7}$$

In the next step we show the invertibility of the Schur complement of $K_{2,2}$ in K_2, which is given by $K_2/K_{2,2} = K_{2,0} - \tilde{K}_{2,1}K_{2,2}^{-1}K_{2,1}$. By the quotient formula for Schur complements, see Lemma C.2, we can express $K_2/K_{2,2}$ as

$$K_2/K_{2,2} = (K_2/\sigma_{33})/(K_{2,2}/\sigma_{33}).$$

Thus, we have to look at the matrix K_2/σ_{33}. Using the alternative partition of K_2, given by

$$K_2 = \begin{pmatrix} A & B \\ C & \sigma_{33} \end{pmatrix},$$

with

$$\begin{aligned} A &= \begin{bmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}, \\ B &= \begin{bmatrix} \sigma_{31} & \sigma_{32} \end{bmatrix}^T, \\ C &= \begin{bmatrix} \sigma_{13} & \sigma_{23} \end{bmatrix}, \end{aligned}$$

shows that we have

$$\begin{aligned} K_2/\sigma_{33} &= A - B\left(\sigma_{33}\right)^{-1}C, \\ &= \begin{pmatrix} \sigma_{11} - \sigma_{31}\left(\sigma_{33}\right)^{-1}\sigma_{13} & \sigma_{21} - \sigma_{31}\left(\sigma_{33}\right)^{-1}\sigma_{23} \\ \sigma_{12} - \sigma_{32}\left(\sigma_{33}\right)^{-1}\sigma_{13} & \sigma_{22} - \sigma_{32}\left(\sigma_{33}\right)^{-1}\sigma_{23} \end{pmatrix}, \\ &=: \begin{pmatrix} \mathcal{K}_{2,0} & \tilde{\mathcal{K}}_{2,1} \\ \mathcal{K}_{2,1} & K_{2,2}/\sigma_{33} \end{pmatrix}. \end{aligned} \tag{3.8}$$

This means

$$\begin{aligned} K_2/K_{2,2} &= (K_2/\sigma_{33})/(K_{2,2}/\sigma_{33}) \\ &= \sigma_{11} - \sigma_{31}\left(\sigma_{33}\right)^{-1}\sigma_{13} - \tilde{\mathcal{K}}_{2,1}\left(K_{2,2}/\sigma_{33}\right)^{-1}\mathcal{K}_{2,1}. \end{aligned}$$

So we can estimate using Lemma 3.1 and the derived bound for $(K_{2,2}/\sigma_{33})^{-1}$,

$$\left\|-\tilde{\sigma}_N''(0)I - K_2/K_{2,2}\right\|_\infty \leq \left\|\tilde{\sigma}_N''(0)I - \sigma_{11}\right\| + \left\|\sigma_{31}\left(\sigma_{33}\right)^{-1}\sigma_{13}\right\|_\infty$$

$$+ \left\| \tilde{\mathcal{K}}_{2,1} \left(K_{2,2}/\sigma_{33}\right)^{-1} \mathcal{K}_{2,1} \right\|_\infty , \\ \leq c_s(N+1)^2 \frac{a_1}{1-2a_1}.$$

For this reason,

$$\left\| I - \frac{K_2/K_{2,2}}{-\tilde{\sigma}_N''(0)} \right\|_\infty \leq \frac{a_1}{1-2a_1}.$$

Again, we have

$$\frac{a_1}{1-2a_1} \leq \frac{1}{b-2} < 1,$$

which gives the invertibility of $K_2/K_{2,2}$ with

$$\left\| (K_2/K_{2,2})^{-1} \right\|_\infty \leq \frac{(N+1)^{-2}(b-2)}{c_s(b-3)}.$$

This shows the invertibility of K_2. If we denote $S = K_2/K_{2,2}$, then the inverse is given by

$$K_2^{-1} = \begin{pmatrix} S^{-1} & -S^{-1}\tilde{K}_{2,1}K_{2,2}^{-1} \\ -K_{2,2}^{-1}K_{2,1}S^{-1} & K_{2,2}^{-1} + K_{2,2}^{-1}K_{2,1}S^{-1}\tilde{K}_{2,1}K_{2,2}^{-1} \end{pmatrix},$$

which has the blockwise representation

$$K_2^{-1} = \begin{pmatrix} \mathcal{A}_{1,1} & \mathcal{A}_{1,2} & \mathcal{A}_{1,3} \\ \mathcal{A}_{2,1} & \mathcal{A}_{2,2} & \mathcal{A}_{2,3} \\ \mathcal{A}_{3,1} & \mathcal{A}_{3,2} & \mathcal{A}_{3,3} \end{pmatrix}, \tag{3.9}$$

where

$$\begin{aligned} &\mathcal{A}_{1,1} = S^{-1}, \quad \mathcal{A}_{2,1} = -T^{-1}\mathcal{K}_{2,1}\mathcal{A}_{1,1}, \quad \mathcal{A}_{1,2} = -\mathcal{A}_{1,1}\tilde{\mathcal{K}}_{2,1}T^{-1} \\ &\mathcal{A}_{1,3} = -(\mathcal{A}_{1,1}\sigma_{31} + \mathcal{A}_{1,2}\sigma_{32})\sigma_{33}^{-1}, \quad \mathcal{A}_{3,1} = -\sigma_{33}^{-1}(\sigma_{13}\mathcal{A}_{1,1} + \sigma_{23}\mathcal{A}_{2,1}) \\ &\mathcal{A}_{2,2} = T^{-1}(id - \mathcal{K}_{2,1}\mathcal{A}_{1,2}), \quad \mathcal{A}_{3,2} = -\sigma_{33}^{-1}(\sigma_{23}\mathcal{A}_{2,2} + \sigma_{1,3}\mathcal{A}_{1,2}) \\ &\mathcal{A}_{2,3} = -(\mathcal{A}_{2,2}\sigma_{32} + \mathcal{A}_{2,1}\sigma_{31})\sigma_{33}^{-1}, \quad \mathcal{A}_{3,3} = \sigma_{33}^{-1}(id - (\sigma_{23}\mathcal{A}_{2,3} + \sigma_{13}\mathcal{A}_{1,3})), \end{aligned}$$

and $\mathcal{K}_{2,1}, \tilde{\mathcal{K}}_{2,1}$ are given in (3.8). This leads to the following norm bounds

$$\|\mathcal{A}_{i,j}\|_\infty \leq \begin{cases} \frac{(N+1)^{-2}(b-2)}{c_s(b-3)}, & i = j, \\ \frac{(N+1)^{-2}}{c_s(b-3)}, & i \neq j. \end{cases}$$

In the last step, we apply the same procedure to show the invertibility of $R = K/K_2$. So, as seen before, we use the quotient rule

$$K/K_2 = (K/K_{2,2})/(K_2/K_{2,2}).$$

To calculate $K/K_{2,2}$ we split K into blocks as

$$K = \begin{pmatrix} A & B \\ C & K_{2,2} \end{pmatrix},$$

with

$$A = \begin{bmatrix} \sigma_{00} & \sigma_{10} \\ \sigma_{01} & \sigma_{11} \end{bmatrix},$$
$$B = \begin{bmatrix} \sigma_{20} & \sigma_{30} \\ \sigma_{21} & \sigma_{31} \end{bmatrix},$$
$$C = \begin{bmatrix} \sigma_{02} & \sigma_{12} \\ \sigma_{03} & \sigma_{13} \end{bmatrix},$$

which leads to

$$K/K_{2,2} = A - B\left(K_{2,2}\right)^{-1} C.$$

A lengthy calculation shows that we can write

$$K/K_{2,2} = A - B\left(K_{2,2}\right)^{-1} C = \begin{pmatrix} \mathcal{K}_0 & \tilde{\mathcal{K}}_1 \\ \mathcal{K}_1 & K_2/K_{2,2} \end{pmatrix},$$

with

$$\begin{aligned} \mathcal{K}_0 &= \Big(\sigma_{00} - \sigma_{30}\left(\sigma_{33}\right)^{-1} \sigma_{03}\Big) - \tilde{\mathcal{C}}_{2,1} T^{-1} \mathcal{C}_{2,1}, \\ \tilde{\mathcal{K}}_1 &= \tilde{\mathcal{G}}_{2,1} - \tilde{\mathcal{C}}_{2,1} T^{-1} \mathcal{K}_{2,1}, \\ \mathcal{K}_1 &= \mathcal{G}_{2,1} - \tilde{\mathcal{K}}_{2,1} T^{-1} \mathcal{C}_{2,1}, \end{aligned}$$

where $\mathcal{K}_{2,1}$, $\tilde{\mathcal{K}}_{2,1}$ are given by (3.8) and

$$\begin{aligned} \mathcal{C}_{2,1} &= \sigma_{02} - \sigma_{32}\left(\sigma_{33}\right)^{-1} \sigma_{03}, \\ \tilde{\mathcal{C}}_{2,1} &= \sigma_{20} - \sigma_{30}\left(\sigma_{33}\right)^{-1} \sigma_{23}, \\ \mathcal{G}_{2,1} &= \sigma_{01} - \sigma_{31}\left(\sigma_{33}\right)^{-1} \sigma_{03}, \\ \tilde{\mathcal{G}}_{2,1} &= \sigma_{10} - \sigma_{30}\left(\sigma_{33}\right)^{-1} \sigma_{13}. \end{aligned}$$

This yields

$$K/K_2 = \mathcal{K}_0 - \tilde{\mathcal{K}}_1\left(K_2/K_{2,2}\right)^{-1} \mathcal{K}_1,$$

and therefore

$$\|I - K/K_2\|_\infty \leq \|I - \mathcal{K}_0\|_\infty + \|\tilde{\mathcal{K}}_1\|_\infty \|\mathcal{K}_1\|_\infty \left\|\left(K_2/K_{2,2}\right)^{-1}\right\|_\infty.$$

Observe, that we have the bounds

$$\|\mathcal{C}_{2,1}\|_\infty, \|\tilde{\mathcal{C}}_{2,1}\|_\infty, \|\mathcal{G}_{2,1}\|_\infty, \|\tilde{\mathcal{G}}_{2,1}\|_\infty \leq \frac{C_{1,s}(N+1)}{\nu^s} \frac{1}{1-a_1}.$$

Using this, we can derive

$$\begin{aligned} \|I - \mathcal{K}_0\|_\infty &\leq \|I - \sigma_{00}\|_\infty + \|\sigma_{30}\|_\infty \|\sigma_{03}\|_\infty \left\|\left(\sigma_{33}\right)^{-1}\right\|_\infty + \|\mathcal{C}_{2,1}\|_\infty \|\tilde{\mathcal{C}}_{2,1}\|_\infty \|T^{-1}\|_\infty, \\ &\leq \frac{C_{0,s}}{\nu^s} + \left(\frac{C_{1,s}}{\nu^s}\right)^2 \frac{2}{c_s(1-2a_1)}, \end{aligned}$$

and similarly

$$\|\tilde{\mathcal{K}}_1\|_\infty, \quad \|\mathcal{K}_1\|_\infty \leq \frac{C_{1,s}(N+1)}{\nu^s} \frac{1}{1-2a_1}.$$

This results in

$$\|I - K/K_2\|_\infty \le \frac{C_{0,s}}{\nu^s} + \left(\frac{C_{1,s}}{\nu^s}\right)^2 \frac{3}{c_s(1-3a_1)}.$$

Since $2C_{0,s} \le C_{1,s} \le \frac{C_{2,s}}{2}$ and $b > 3 + \frac{c_s}{2}$, we can estimate

$$\frac{C_{0,s}}{\nu^s} + \left(\frac{C_{1,s}}{\nu^s}\right)^2 \frac{3}{c_s(1-3a_1)} \le \frac{1}{4}c_s\frac{a_1}{1-3a_1} < \frac{c_s}{4(b-3)} < 1,$$

and K/K_2 is invertible with

$$\left\|(K/K_2)^{-1}\right\|_\infty \le 1 + \frac{c_s}{4(b-3)-c_s}.$$

This completes the last step and gives the invertibility of K. With $R = K/K_2$ we have

$$K^{-1} = \begin{pmatrix} R^{-1} & -R^{-1}\tilde{K}_1K_2^{-1} \\ -K_2^{-1}K_1R^{-1} & K_2^{-1} + K_2^{-1}K_1R^{-1}\tilde{K}_1K_2^{-1}\end{pmatrix}.$$

Using the representations (3.9) of the inverse of K_2, a lengthy calculation shows the following block-wise representation

$$K^{-1} = \begin{pmatrix} \mathcal{B}_{1,1} & \mathcal{B}_{1,2} & \mathcal{B}_{1,3} & \mathcal{B}_{1,4} \\ \mathcal{B}_{2,1} & \mathcal{B}_{2,2} & \mathcal{B}_{2,3} & \mathcal{B}_{2,4} \\ \mathcal{B}_{3,1} & \mathcal{B}_{3,2} & \mathcal{B}_{3,3} & \mathcal{B}_{3,4} \\ \mathcal{B}_{4,1} & \mathcal{B}_{4,2} & \mathcal{B}_{4,3} & \mathcal{B}_{4,4}\end{pmatrix},$$

with

$$\begin{aligned}&\mathcal{B}_{1,1} = R^{-1}, \quad \mathcal{B}_{1,2} = -R^{-1}\tilde{\mathcal{K}}_1S^{-1},\\ &\mathcal{B}_{2,1} = -S^{-1}\mathcal{K}_1R^{-1}, \quad \mathcal{B}_{2,2} = S^{-1}(id + \mathcal{K}_1R^{-1}\tilde{\mathcal{K}}_1S^{-1})\end{aligned}$$

and

$$\begin{aligned}&\mathcal{B}_{1,3} = -(\mathcal{B}_{1,2}\tilde{\mathcal{K}}_{2,1} + \mathcal{B}_{1,1}\tilde{\mathcal{C}}_{2,1})T^{-1}, \quad \mathcal{B}_{1,4} = -(\mathcal{B}_{1,1}\sigma_{30} + \mathcal{B}_{1,2}\sigma_{31} + \mathcal{B}_{1,3}\sigma_{32})\sigma_{33}^{-1}\\ &\mathcal{B}_{3,1} = -T^{-1}(\mathcal{K}_{2,1}\mathcal{B}_{2,1} + \mathcal{C}_{2,1}\mathcal{B}_{1,1}), \quad \mathcal{B}_{4,1} = -\sigma_{33}^{-1}(\sigma_{03}\mathcal{B}_{1,1} + \sigma_{13}\mathcal{B}_{2,1} + \sigma_{23}\mathcal{B}_{3,1}),\\ &\mathcal{B}_{2,3} = -(\mathcal{B}_{22}\tilde{\mathcal{K}}_{2,1} + \mathcal{B}_{2,1}\tilde{\mathcal{C}}_{2,1})T^{-1}, \quad \mathcal{B}_{3,2} = -T^{-1}(\mathcal{K}_{2,1}\mathcal{B}_{22} + \mathcal{C}_{2,1}\mathcal{B}_{1,2}),\\ &\mathcal{B}_{3,3} = T^{-1}(id - (\mathcal{K}_{2,1}\mathcal{B}_{2,3} + \mathcal{C}_{2,1}\mathcal{B}_{1,3})), \quad \mathcal{B}_{4,2} = -\sigma_{33}^{-1}(\sigma_{23}\mathcal{B}_{3,2} + \sigma_{13}\mathcal{B}_{2,2} + \sigma_{03}\mathcal{B}_{1,2}),\\ &\mathcal{B}_{2,4} = -(\mathcal{B}_{2,3}\sigma_{32} + \mathcal{B}_{2,2}\sigma_{31} + \mathcal{B}_{2,1}\sigma_{30})\sigma_{33}^{-1}, \quad \mathcal{B}_{3,4} = -(\mathcal{B}_{3,1}\sigma_{30} + \mathcal{B}_{3,2}\sigma_{31} + \mathcal{B}_{3,3}\sigma_{32})\sigma_{33}^{-1},\\ &\mathcal{B}_{4,3} = -\sigma_{33}^{-1}(\sigma_{03}\mathcal{B}_{1,3} + \sigma_{13}\mathcal{B}_{2,3} + \sigma_{23}\mathcal{B}_{3,3}), \quad \mathcal{B}_{4,4} = \sigma_{33}^{-1}(id - (\sigma_{03}\mathcal{B}_{14} + \sigma_{13}\mathcal{B}_{24} + \sigma_{23}\mathcal{B}_{34})).\end{aligned}$$

Accordingly, we can derive the following norm bounds

$$\begin{aligned}\|\mathcal{B}_{1,1}\|_\infty &\le 1 + \frac{c_s}{4(b-3)-c_s}, \quad \|\mathcal{B}_{1,2}\|_\infty, \|\mathcal{B}_{2,1}\|_\infty \le \frac{2(N+1)^{-1}}{(4(b-3)-c_s)},\\ \|\mathcal{B}_{2,2}\|_\infty &\le \frac{(4(b-2)-c_s)(N+1)^{-2}}{c_s(4(b-3)-c_s)},\\ \|\mathcal{B}_{1,3}\|_\infty, \|\mathcal{B}_{3,1}\|_\infty &\le \frac{2(b-3)(N+1)^{-1}}{(b-2)(4(b-3)-c_s)} + \frac{(N+1)^{-2}}{(b-2)(4(b-3)-c_s)} \le \frac{2(N+1)^{-1}}{(4(b-3)-c_s)},\\ \|\mathcal{B}_{1,4}\|_\infty, \|\mathcal{B}_{4,1}\|_\infty &\le \frac{2(b-3)(N+1)^{-1}}{(b-1)(4(b-3)-c_s)} + \frac{(b-5)(N+1)^{-2}}{(b-1)(4(b-3)-c_s)} + \frac{(N+1)^{-3}}{2(b-2)(b-1)(4(b-3)-c_s)},\end{aligned}$$

$$\leq \frac{2(N+1)^{-1}}{(4(b-3)-c_s)},$$

$$\|\mathcal{B}_{2,3}\|_\infty, \|\mathcal{B}_{3,2}\|_\infty \leq \frac{(N+1)^{-2}}{(b-2)(4(b-3)-c_s)} + \frac{(4(b-2)-c_s)(N+1)^{-2}}{c_s(b-2)(4(b-3)-c_s)} \leq \frac{4(N+1)^{-2}}{c_s(4(b-3)-c_s)},$$

$$\|\mathcal{B}_{3,3}\|_\infty \leq \frac{(4(b-2)-c_s)(N+1)^{-2}}{c_s(4(b-3)-c_s)},$$

$$\|\mathcal{B}_{2,4}\|_\infty, \|\mathcal{B}_{4,2}\|_\infty \leq \frac{4(N+1)^{-2}}{c_s(4(b-3)-c_s)}, \quad \|\mathcal{B}_{3,4}\|_\infty, \|\mathcal{B}_{4,3}\|_\infty \leq \frac{4(N+1)^{-2}}{c_s(4(b-3)-c_s)},$$

$$\|\mathcal{B}_{4,4}\|_\infty \leq \frac{(4(b-2)-c_s)(N+1)^{-2}}{c_s(4(b-3)-c_s)}.$$

Using these norm bounds, we get the bound

$$\begin{aligned}
\|\alpha_0\| &= \|\mathcal{B}_{1,1}u + \mathcal{B}_{1,2}b_1 + \mathcal{B}_{1,3}b_2 + \mathcal{B}_{1,4}b_3\|_\infty, \\
&\leq \|\mathcal{B}_{1,1}\|_\infty\|u\|_\infty + \|\mathcal{B}_{1,2}\|_\infty\|b_1\|_\infty + \|\mathcal{B}_{1,3}\|_\infty\|b_2\|_\infty + \|\mathcal{B}_{1,4}\|_\infty\|b_3\|_\infty, \\
&\leq \frac{1}{4(b-3)-c_s}(4(b-3)\|u\|_\infty + 2(N+1)^{-1}(\|b_1\|_\infty + \|b_2\|_\infty + \|b_3\|_\infty))
\end{aligned}$$

and in a similar way the desired bounds for $\alpha_1, \alpha_2, \alpha_3$. Moreover, if $|u_i| = 1$, $\|u\|_\infty \leq 1$ and $b_1 = b_2 = b_3 = 0$, we can estimate

$$\begin{aligned}
|\alpha_{0,i}| &= \left|\left(\left(I - \left(I - (K/K_2)^{-1}\right)\right)u\right)_i\right|, \\
&= \left|u_i - \left(\left(I - (K/K_2)^{-1}\right)u\right)_i\right|, \\
&\geq \left||u_i| - \left|\left(\left(I - (K/K_2)^{-1}\right)u\right)_i\right|\right|, \\
&\geq \left|1 - \left|(I - K/K_2)(K/K_2)^{-1}u_i\right|\right|.
\end{aligned}$$

Since $b > 3 + \frac{c_s}{2}$, we have

$$\begin{aligned}
\left|(I - K/K_2)(K/K_2)^{-1}u_i\right| &\leq \|I - K/K_2\|_\infty\|(K/K_2)^{-1}\|_\infty, \\
&\leq \frac{c_s}{4(b-3)-c_s} < 1,
\end{aligned}$$

and therefore

$$|\alpha_{0,i}| \geq 1 - \frac{c_s}{4(b-3)-c_s}.$$

□

Corollary 3.3. *Suppose the interpolation points $\mathcal{X} = \{x_1, \ldots, x_M\}$ obey the separation condition*

$$\min_{x_i \neq x_j} \omega(x_j^{-1}x_i) \geq \frac{36}{N+1} \tag{3.10}$$

for $N \geq 20$. Then the interpolation problem (3.3) *has a unique solution, such that the coefficients obey*

$$\|\alpha_0\|_\infty \leq 1 + 6\cdot 10^{-4}, \quad \|\alpha_j\|_\infty \leq \frac{0.02 + 2\cdot 10^{-5}}{(N+1)}, \quad j = 1,2,3,$$

and

$$|\alpha_{0,i}| \geq 1 - 6\cdot 10^{-4}.$$

Proof. It can be checked that the condition (3.4) is fulfilled for the parameters $\nu = 36$, $b = 28$ and $s = 8$. □

3.2 Bound for the Interpolant

Next to the interpolation conditions, guaranteed by Corollary 3.3, we have to show that the interpolant q fulfills the second assumption in (3.1), namely $|q(x)| < 1$ for x not being an interpolation point. We split the proof into those $x \in SO(3)$, which are close to an interpolation point, which is Lemma 3.4, and those which are well separated, governed by Lemma 3.5.

Lemma 3.4. *Suppose the separation condition* (3.10) *is satisfied for* $N \geq 20$. *Then for all* $x \in SO(3)$, *such that there is a* x_m *with* $\omega(x_m^{-1}x) \leq \frac{\pi}{2(N+1)}$, *we have for the interpolating function* q *of Corollary 3.3*

$$|q(x)| < 1.$$

Proof. The proof is based on a concavity respectively a convexity argument. We show that in the prescribed neighbourhood of an interpolation point x_m the Hessian is negative definite in the case $u_m = 1$ and positive definite in the case $u_m = -1$, using the Theorem of Gerschgorin. For this, first observe that the Hessian matrix for a function f is given by

$$Hf = \begin{pmatrix} X_1X_1f & X_1X_2f & X_1X_3f \\ X_2X_1f & X_2X_2f & X_2X_3f \\ X_3X_1f & X_3X_2f & X_3X_3f \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 0 & X_3f & -X_2f \\ -X_3f & 0 & X_1f \\ X_2f & -X_1f & 0 \end{pmatrix}. \tag{3.11}$$

If we apply this to the function constructed from the interpolation problem

$$q(x) = \sum_{j=1}^{M} \alpha_{j,0}\sigma_N(x, x_j) + \alpha_{j,1}X_1^y\sigma_N(x, x_j) + \alpha_{j,2}X_2^y\sigma_N(x, x_j) + \alpha_{j,3}X_3^y\sigma_N(x, x_j),$$

the diagonal entries of the matrix are given by

$$\begin{aligned} X_i^xX_i^xq(x) = \sum_{j=1}^{M} & \alpha_{j,0}X_i^xX_i^x\sigma_N(x, x_j) + \alpha_{j,1}X_i^xX_i^xX_1^y\sigma_N(x, x_j) + \alpha_{j,2}X_i^xX_i^xX_2^y\sigma_N(x, x_j) \\ & + \alpha_{j,3}X_i^xX_i^xX_3^y\sigma_N(x, x_j). \end{aligned}$$

For the estimates of the entries of the Hessian, we use Lemma 2.4, 2.5, 2.6, 3.2 and 2.2 with the following parameters $s = 8$, $\nu = 36$, $b = 28$, $\delta = \frac{\pi}{2}$, $c_8 = \frac{0.999}{18}$, $\tilde{c}_8 = \frac{1.001}{18}$ and $\tilde{d}_8 = \frac{3\cdot 1.001}{40\cdot 9}$. We assume that $u_m = 1$, since the estimates for $u_m = -1$ are completely analog. The first step is to show, that the diagonal entries are negative. For this, we estimate

$$\begin{aligned} X_i^xX_i^xq(x) \leq \alpha_{0,m}X_i^xX_i^x\sigma_N(x, x_m) &+ \sum_{n=1}^{3} \|\alpha_n\|_\infty \left|X_i^xX_i^xX_n^y\sigma_N(x, x_m)\right| \\ &+ \|\alpha_0\|_\infty \sum_{x_j \neq x_m} \left|X_i^xX_i^x\sigma_N(x, x_j)\right| + \sum_{n=1}^{3} \|\alpha_n\|_\infty \sum_{x_j \neq x_m} \left|X_i^xX_i^xX_n^y\sigma_N(x, x_j)\right|. \end{aligned}$$

The first term can be estimated using the bounds of Lemma 2.2 and Theorem 3.2

$$\begin{aligned} \alpha_{0,m}X_i^xX_i^x\sigma_N(x, x_m) &= \alpha_{0,m}X_i^xX_i^x\sigma_N(x, x) + \alpha_{0,m}\left(X_i^xX_i^x\sigma_N(x, x_m) - X_i^xX_i^x\sigma_N(x, x)\right), \\ &= \alpha_{0,m}\tilde{\sigma}_N^{''}(0) + \alpha_{0,m}\left(X_i^xX_i^x\sigma_N(x, x_m) - \tilde{\sigma}_N^{''}(0)\right), \\ &\leq -c_8(N+1)^2\left(1 - \frac{c_8}{4(b-3) - c_8}\right) \end{aligned}$$

$$+\left(1+\frac{c_8}{4(b-3)-c_8}\right)\left(X_i^x X_i^x \sigma_N(x,x_m)-\tilde{\sigma}_N''(0)\right).$$

By Lemma 2.5, we have the bound

$$\left(X_i^x X_i^x \sigma_N(x,x_m)-\tilde{\sigma}_N''(0)\right)\le \frac{\tilde{d}_8}{2}(N+1)^2\delta^2, \quad \omega\in[0,\frac{\delta}{N+1}],$$

which yields

$$\begin{aligned}\alpha_{0,m}X_i^x X_i^x \sigma_N(x,x_m) &\le -(N+1)^2\left(1+\frac{c_8}{4(b-3)-c_8}\right)c_8\left(1-\frac{c_8}{2(b-3)}-\frac{\delta^2\tilde{d}_8}{2c_8}\right)\\ &\le -4.518\cdot 10^{-2}\cdot(N+1)^2.\end{aligned}$$

In addition, we have again using Lemma 2.5 and Theorem 3.2

$$\begin{aligned}\sum_{n=1}^{3}\|\alpha_n\|_\infty \left|X_i^x X_i^x X_n^y \sigma_N(x,x_m)\right| &\le \tilde{d}_8\frac{6(N+1)^2\delta}{4(b-3)-c_8}\left(1+\frac{\delta}{4\cdot 21}+\frac{\tilde{c}_8}{4\tilde{d}_8\cdot 21^2}\right),\\ &\le 8.044\cdot 10^{-4}\cdot(N+1)^2,\end{aligned}$$

as well as, using Lemma 2.6 with $a_{\delta/\nu}=\min\{\frac{27}{124}(1-\frac{\delta}{\nu})^{-8}+1,(1-\frac{\delta}{\nu})^{-8}\}$ and the assumption $\nu^s>b\frac{C_{2,8}}{c_8}$,

$$\begin{aligned}\|\alpha_0\|_\infty\sum_{x_j\neq x_m}\left|X_i^x X_i^x \sigma_N(x,x_j)\right| &\le \left(1+\frac{c_8}{4(b-3)-c_8}\right)\frac{c_8 a_{\delta/\nu}(N+1)^2}{b},\\ &\le 2.601\cdot 10^{-3}\cdot(N+1)^2\end{aligned}$$

and

$$\begin{aligned}\sum_{n=1}^{3}\|\alpha_n\|_\infty\sum_{x_j\neq x_m}\left|X_i^x X_i^x X_n^y \sigma_N(x,x_j)\right| &\le 3\cdot\frac{c_{3,8}}{c_{2,8}}\cdot\frac{2c_s a_{\frac{\delta}{\nu}}}{b(4(b-3)-c_8)},\\ &\le 3.31\cdot 10^{-4}\cdot(N+1)^2.\end{aligned}$$

Inserting the parameters, we find

$$X_i^x X_i^x q(x)\le -4.144\cdot 10^{-2}\cdot(N+1)^2.$$

For the off-diagonal entries of the Hessian matrix we have

$$\begin{aligned}X_j^x X_i^x q \mp \frac{1}{2}X_n^x q &= \sum_{j=1}^{M}\alpha_{j,0}\left(X_j^x X_i^x \sigma_N(x,x_j)\mp\frac{1}{2}X_n^x\sigma_N(x,x_j)\right)\\ &\quad+\sum_{k=1}^{3}\alpha_{j,k}\left(X_j^x X_i^x X_k^y\sigma_N(x,x_j)\mp\frac{1}{2}X_n^x X_k^x\sigma_N(x,x_j)\right)\end{aligned}$$

and therefore we can estimate

$$|X_j^x X_i^x q\mp\frac{1}{2}X_n^x q|\le\|\alpha_0\|_\infty\left|X_j^x X_i^x\sigma_N(x,x_m)\mp\frac{1}{2}X_n^x\sigma_N(x,x_m)\right|$$

$$+\sum_{k=1}^{3}\|\alpha_k\|_\infty\left|X_j^xX_i^xX_k^y\sigma_N(x,x_m)\mp\frac{1}{2}X_n^xX_k^x\sigma_N(x,x_m)\right|$$
$$+\|\alpha_0\|_\infty\sum_{x_j\neq x_m}\left|X_j^xX_i^x\sigma_N(x,x_j)\mp\frac{1}{2}X_n^x\sigma_N(x,x_j)\right|$$
$$+\sum_{k=1}^{3}\|\alpha_k\|_\infty\sum_{x_j\neq x_m}\left|X_j^xX_i^xX_k^y\sigma_N(x,x_j)\mp\frac{1}{2}X_n^xX_k^x\sigma_N(x,x_j)\right|.$$

Lemma 2.5 and Theorem 3.2 yield

$$\|\alpha_0\|_\infty\left|X_j^xX_i^x\sigma_N(x,x_m)\mp\frac{1}{2}X_n^x\sigma_N(x,x_m)\right|\leq\left(1+\frac{c_8}{4(b-3)-c_8}\right)(N+1)^2\frac{\tilde{d}_8}{4}\delta^2,$$
$$\leq 5.149\cdot 10^{-3}\cdot(N+1)^2,$$

as well as

$$\sum_{k=1}^{3}\|\alpha_k\|_\infty\left|X_j^xX_i^xX_k^y\sigma_N(x,x_m)\mp\frac{1}{2}X_n^xX_k^x\sigma_N(x,x_m)\right|$$
$$\leq\tilde{d}_8\frac{6(N+1)^2\delta}{4(b-3)-c_8}\left(1+\frac{\delta}{4\cdot 21}+\frac{\tilde{c_8}}{4\tilde{d}_8\cdot 21^2}\right),$$
$$\leq 8.044\cdot 10^{-4}\cdot(N+1)^2.$$

Furthermore, we apply Lemma 2.4 to derive

$$\|\alpha_0\|_\infty\sum_{x_j\neq x_m}\left|X_j^xX_i^x\sigma_N(x,x_j)\mp\frac{1}{2}X_n^x\sigma_N(x,x_j)\right|$$
$$\leq\left(1+\frac{c_8}{4(b-3)-c_8}\right)\frac{c_8a_{\delta/\nu}(N+1)^2}{b},$$
$$\leq 2.601\cdot 10^{-3}\cdot(N+1)^2$$

and

$$\sum_{k=1}^{3}\|\alpha_k\|_\infty\sum_{x_j\neq x_m}\left|X_j^xX_i^xX_k^y\sigma_N(x,x_j)\mp\frac{1}{2}X_n^xX_k^x\sigma_N(x,x_j)\right|$$
$$\leq 3\cdot\frac{c_{3,8}}{c_{2,8}}\cdot\frac{2c_sa_{\frac{\delta}{\nu}}}{b(4(b-3)-c_8)},$$
$$\leq 3.31\cdot 10^{-4}\cdot(N+1)^2.$$

Inserting the parameters results in

$$|X_j^xX_i^xq\mp\frac{1}{2}X_n^xq|\leq 8.886\cdot 10^{-3}\cdot(N+1)^2.$$

Since

$$|X_i^xX_i^xq(x)|>2|X_j^xX_i^xq\mp\frac{1}{2}X_n^xq|,$$

and $X_i^xX_i^xq(x)<0$ for $i=1,2,3$, we see that the Hessian matrix is negative definite at x with $\lambda_{\max}(Hq(x))\leq -2.36\cdot 10^{-2}\cdot(N+1)^2$.

The definitness shows $q(x) < 1$. To show $q(x) > -1$ observe, that

$$q(x) \geq \alpha_{0,m}\sigma_N(x,x_m) - \sum_{k=1}^{3} \|\alpha_k\|_\infty |X_k^y \sigma_N(x,x_m)| - \|\alpha_0\|_\infty \sum_{x_j \neq x_m} |\sigma_N(x,x_j)|$$
$$- \sum_{k=1}^{3} \|\alpha_k\|_\infty \sum_{x_j \neq x_m} |X_k^x \sigma_N(x,x_j)|.$$

We have, using the Taylor expansion of cosine at zero,

$$\alpha_{0,m}\sigma_N(x,x_m) \geq \left(1 - \frac{c_8}{4(b-3)-c_8}\right) \cdot \left(1 - \frac{\tilde{c}_8}{2}\delta^2\right) \geq 0.93$$

and

$$\|\alpha_k\|_\infty |X_k^y \sigma_N(x,x_m)| \leq \frac{2\tilde{c}_8\delta}{4(b-3)-c_8} \leq 1.8 \cdot 10^{-3}.$$

Using Lemma 2.6 yields

$$\|\alpha_0\|_\infty \sum_{x_j \neq x_m} |\sigma_N(x,x_j)| \leq \left(1 + \frac{c_8}{4(b-3)-c_8}\right) \cdot \frac{c_8 a_{\delta/\nu}}{4b} \leq 6.501 \cdot 10^{-4},$$
$$\|\alpha_k\|_\infty \sum_{x_j \neq x_m} |X_k^x \sigma_N(x,x_j)| \leq \frac{2}{4(b-3)-c_8} \cdot \frac{c_8 a_{\delta/\nu}}{2b} \leq 2.601 \cdot 10^{-5}.$$

Combining this results in

$$q(x) \geq 0.92.$$

The case $u_i = -1$ is completely analog and one has $\lambda_{\min}(Hq(x)) \geq 2.36 \cdot 10^{-2} \cdot (N+1)^2$ and $q(x) \leq -0.92$.

□

Lemma 3.5. *Under the assumptions of Lemma 3.4, we have that for all $x \in SO(3)$ with $\omega(x_m^{-1}x) \geq \frac{\pi}{2(N+1)}$ for all $x_m \in \mathcal{X}$ the interpolating function q of Corollary 3.3 fulfills*

$$|q(x)| < 1.$$

Proof. We can estimate

$$|q(x)| \leq \|\alpha_0\|_\infty |\sigma_N(x,x_m)| + \sum_{k=1}^{3} \|\alpha_k\|_\infty |X_k^y \sigma_N(x,x_m)|$$
$$+ \|\alpha_0\|_\infty \sum_{x_j \neq x_m} |\sigma_N(x,x_j)| + \sum_{k=1}^{3} \|\alpha_k\|_\infty \sum_{x_j \neq x_m} |X_k^x \sigma_N(x,x_j)|. \tag{3.12}$$

First, assume that there is an x_m, such that $\omega(x_m^{-1}x) \leq \frac{2.45\pi}{N+1}$. Using the Taylor expansion of the cosine function at zero, we have with Lemma 2.2

$$\sigma_N(x,x_m) \geq 1 - \frac{|\tilde{\sigma}_N^{(2)}(0)|}{2}\omega^2 + \frac{|\tilde{\sigma}_N^{(4)}(0)|}{24}\omega^4 - \frac{|\tilde{\sigma}_N^{(6)}(0)|}{6!}\omega^6,$$

$$\geq 1 - \frac{1.001}{36}(N+1)^2\omega^2 + \frac{0.999}{24 \cdot 120}(N+1)^4\omega^4 - \frac{1.011}{6! \cdot 11 \cdot 48}(N+1)^6\omega^6.$$

It can be shown that the polynomial

$$1 - \frac{1.001}{36}t^2 + \frac{0.999}{24 \cdot 120}t^4 - \frac{1.011}{6! \cdot 11 \cdot 48}t^6$$

is positive for $t \in [0, 2.45\pi]$, since its derivative is strictly negative on this interval and the polynomial is positive evaluated at $t = 2.45\pi$. Therefore,

$$\begin{aligned} |\sigma_N(x, x_m)| &= \sigma_N(x, x_m), \\ &\leq 1 - \frac{|\tilde{\sigma}_N^{(2)}(0)|}{2}\omega^2 + \frac{|\tilde{\sigma}_N^{(4)}(0)|}{24}\omega^4, \\ &\leq 1 - \frac{1.001}{36}(N+1)^2\omega^2 + \frac{0.999}{24 \cdot 120}(N+1)^4\omega^4. \end{aligned} \tag{3.13}$$

The right hand side of (3.13) is strictly monotonic decreasing for $\omega \in \left[\frac{\pi}{2(N+1)}, \frac{t_0}{(N+1)}\right]$ with $t_0 = 2\sqrt{10 \cdot \frac{1.001}{0.999}}$ and strictly increasing for $\omega \in \left[\frac{t_0}{(N+1)}, \frac{2.45\pi}{(N+1)}\right]$. Furthermore, we can estimate

$$|X_k^y \sigma_N(x, x_m)| \leq \tilde{c}_s(N+1)t, \tag{3.14}$$

as well as using Lemma 2.6

$$\begin{aligned} \sum_{x_j \neq x_m} |\sigma_N(x, x_j)| &\leq \frac{c_s a_{t/\nu}}{4b}, \\ \sum_{x_j \neq x_m} |X_k^x \sigma_N(x, x_j)| &\leq \frac{c_s a_{t/\nu}(N+1)}{2b}, \end{aligned}$$

for $t = (N+1)\omega$. Inserting the values $b = 28$, $\nu = 36$ and $s = 8$ for the bounds of the coefficients in Theorem 3.2 results in

$$\begin{aligned} |q(x)| &\leq 0.96, \quad \text{for } \omega(x_m^{-1}x) \in \left[\frac{\pi}{2(N+1)}, \frac{t_0}{(N+1)}\right], \\ |q(x)| &\leq 0.60, \quad \text{for } \omega(x_m^{-1}x) \in \left[\frac{t_0}{(N+1)}, \frac{2.45\pi}{(N+1)}\right]. \end{aligned}$$

If there is a x_m such that $\frac{2.45\pi}{N+1} \leq \omega(x_m^{-1}x) \leq \frac{18}{N+1}$, we can estimate in a similar way, but instead of the Taylor expansion we use the asymptotic bound of Theorem 2.3, i.e.

$$|\sigma_N(x, x_m)| \leq \frac{c_{0,8}}{(N+1)^8\omega^8}.$$

For the derivative we use the Bernstein inequality (1.14) to get

$$|X_k^y \sigma_N(x, x_m)| \leq (N+1)\|\sigma_N(\cdot, x_m)\|_\infty = (N+1).$$

This results in

$$|q(x)| \leq 0.95,$$

for $\frac{2.45\pi}{N+1} \leq \omega(x_m^{-1}x) \leq \frac{18}{N+1}$. In the case $\omega(x_j^{-1}x) \geq \frac{18}{(N+1)}$ for all x_j, we derive with Theorem 3.2 and estimates similar to those of Lemma 2.6

$$\begin{aligned} |q(x)| &\leq \|\alpha_0\|_\infty \sum_{x_j} |\sigma_N(x,x_j)| + \sum_{k=1}^{3} \|\alpha_k\|_\infty \sum_{x_j} |X_k^y \sigma_N(x,x_j)|, \\ &\leq \frac{c_s a_{1/2}}{4(b-3)-cs} \leq 0.032. \end{aligned}$$

□

The combination of Corollary 3.3, Lemma 3.4 and Lemma 3.5 gives the main result of this section.

Theorem 3.6. *Suppose the points $\mathcal{X} = \{x_1,\dots,x_M\}$ obey a separation distance of $\rho(\mathcal{X}) \geq \frac{36}{N+1}$ for $N \geq 20$. Then for each sign combination $u_i \in \{-1,1\}$, there is a $q \in \Pi_N$ such that*

$$\begin{aligned} q(x_i) &= u_i, \quad \text{for } x_i \in \mathcal{X}, \\ |q(x)| &< 1, \quad \text{for } x \in SO(3) \setminus \mathcal{X}. \end{aligned}$$

As seen in Chapter 1, the existence of a dual certificate guarantees, that the measure $\mu^\star$ is the unique solution of the total variation minimization. For completeness, we will state this as a corollary.

Corollary 3.7. *Suppose the support of the signed measure $\mu^\star$ obeys the separation condition*

$$\min_{x\neq y} \omega(y^{-1}x) \geq \frac{36}{N+1}, \quad x,y \in \operatorname{supp}(\mu^\star),$$

for $N \geq 20$, then $\mu^\star$ is the unique real solution of the minimization problem

$$\min_{\mu\in\mathcal{M}(SO(3),\mathbb{R})} \|\mu\|_{TV}, \quad \text{subject to} \quad \mathcal{S}_N^*\mu = \mathcal{S}_N^*\mu^\star. \tag{RP}$$

Proof. Theorem 3.6 guarantuees the existence of a dual certificate. Hence, by Theorem 1.7, $\mathcal{S}_N^*$ has the null-space property with respect to $\operatorname{supp}(\mu^\star)$ and Theorem 1.6 shows that $\mu^\star$ is the unique real solution. □

On the spectral side, this recovery result means, that we can recover high frequency information of the measure $\mu^\star$ from its low frequency moments. Which means, we get access to

$$\langle \mu^\star, D_{k,m}^l \rangle, \quad |k|,|m| \leq l$$

for all $l \in \mathbb{N}$, and therefore can extrapolate the spectrum exactly from its low frequency parts. In other words, we can construct the polynomial approximation $\mathcal{S}_L^*$ to the measure $\mu^\star$ for all $L \geq N$, i.e. $\mathcal{S}_L^*\mu = \mathcal{S}_L^*\mu^\star$ for the solution μ of the total variation minimization (RP). If we cannot measure the low frequency information of $\mu^\star$ exactly, this is no longer possible. Instead, we introduce an error by extrapolating the spectrum measured by the function

$$\mathcal{S}_L^*\mu^\star - \mathcal{S}_L^*\mu$$

for any approximation μ to the sought measure $\mu^\star$, that we construct via any process, and the higher frequency $L \geq N$. Clearly, the induced extrapolation error, or *super-resolution error*, should depend on the relation between N and L. This is the content of the next section.

3.3 Super-Resolution Error Estimates

In the last section, we concentrated on the construction of a dual certificate, that guarantees the exact recovery of the sought measure in the case of exact data. Nevertheless, one cannot assume that the given moments of the measure are exact, but corrupted by noise. In this case, we cannot hope to recover the sought measure exactly, but only up to an error, which needs to be controlled. It is therefore convenient to substitute the minimization (RP) with a regularized version of it. In this section, we consider the Thikonov-type problem

$$\min_{\mu \in \mathcal{M}(SO(3),\mathbb{R})} \frac{1}{2}\|\mathcal{S}_N^*(\mu - \mu^\star - \eta)\|^2_{L^2(SO(3))} + \tau\|\mu\|_{TV}. \tag{RP$_\tau$}$$

We choose a deterministic noise model, i.e. the noise η satisfies

$$\|\mathcal{S}_N^*\eta\|_{L^2(SO(3))} \le \varepsilon,$$

and $\varepsilon > 0$ is called noise level. Since the functional

$$J_\tau(\mu) = \frac{1}{2}\|\mathcal{S}_N^*(\mu - \mu^\star - \eta)\|^2_{L^2(SO(3))} + \tau\|\mu\|_{TV}$$

is weakly* lower-semicontinuous and strictly convex, we know that there is a unique solution of (RP$_\tau$), which we will denote by μ_τ. As seen in Chapter 1, see Theorem 1.9, we already know, that μ_τ will converge to $\mu^\star$ in the weak* topology, as long as the regularization paramter τ is coupled adequately to the noise level ε.

In the following, we consider the extrapolation error induced by μ_τ, i.e. we would like to control the difference

$$\mathcal{S}_L^*\mu^\star - \mathcal{S}_L^*\mu_\tau,$$

with $L \ge N$. In the trigonometric setting, the first estimates in this direction can be found in [Candés and Fernandez-Granda, 2013]. The authors give bounds for the L^1-error of the convolution of the difference measure with a high frequency Fejér kernel. Very recently, see [Li, 2017], those estimates where generalized to L^∞ estimates, using a more general class of kernels.

We build on the work [Li, 2017] to show analog estimates for polynomial kernels on the rotation group. Let K_L be a polynomial positive semi-definite kernel of the form

$$K_L(x,y) = \sum_{l=0}^{L} a_l \sum_{k,m=-l}^{l} D^l_{k,m}(x)\overline{D^l_{k,m}(y)}, \tag{3.15}$$

with $a_l > 0$ and $a_l \ge a_{l+1}$. As seen in chapter 2, $K_L(x,y) = \tilde{K}_L(\omega(y^{-1}x))$ with

$$\tilde{K}_L(\omega) = \sum_{k=-L}^{L} b_k \mathrm{e}^{ik\omega}$$

and $b_k = \sum_{l\ge|k|} a_l$. Clearly, this includes the Dirichlet kernel $D_L(x,y)$ as well as the kernel constructed from the B-spline filters seen in Chapter 2. With this kernel we define the approximation operator $\mathcal{K}_L : L^2(SO(3)) \to C(SO(3))$,

$$\mathcal{K}_L f(x) = \int_{SO(3)} f(y)K_L(x,y)\mathrm{d}\lambda(y).$$

The bound on the difference $\mathcal{K}_L^*(\mu_\tau - \mu^\star)$ depends on the quotient of N and L given by

$$s_{L,N} = \frac{L}{N},$$

which is called *super-reolution factor*. The aim of this section is to show the estimate

$$\|\mathcal{K}_L^*(\mu_\tau - \mu^\star)\|_\infty \leq C \cdot \|\tilde{K}_L\|_\infty \cdot s_{L,N}^2 \cdot \tau.$$

It relies on two pillars. First, the volume with respect to the difference measure $|\mu^\star - \mu_\tau|$ of sets close to the support of $\mu^\star$ and its complement can be controlled by τ, which is shown in Lemma 3.8 and is independent of the chosen kernel K_L. Second, any kernel K_L of the form (3.15) can be quadratically approximated near the support of $\mu^\star$ with a low frequency function $f \in \Pi_N$, which is the statement of Lemma 3.9. Both statements are built upon the estimates shown in Chapter 2 as well as the construction of the dual certificate seen in the previous section. Consequently, we will assume that the support set $\mathcal{X} = \{x_j\}$ of the measure $\mu^\star$ obeys the separation condition

$$\min_{x_i \neq x_j} \omega(x_j^{-1} x_i) \geq \frac{36}{N+1}, \tag{3.16}$$

for $N \geq 20$.

Lemma 3.8. *Suppose the support $\mathcal{X} = \{x_j\}$ of the measure $\mu^\star$ obeys the separation condition* (3.16) *and μ_τ is the unique solution of the minimization*

$$\min_{\mu \in \mathcal{M}(SO(3),\mathbb{R})} \frac{1}{2}\|\mathcal{S}_N^*(\mu - \mu^\star - \eta)\|_{L^2(SO(3))}^2 + \tau\|\mu\|_{TV},$$

with $\|\mathcal{S}_N^* \eta\|_{L^2(SO(3))} \leq \tau$. Then

$$\begin{aligned} \int_{\left(\bigcup_j B_r(x_j)\right)^c} \mathrm{d}|\mu_\tau - \mu^\star| &\leq 100 \cdot \tau, \\ \sum_j \int_{B_r(x_j)} \omega(x_j^{-1}x)^2 \mathrm{d}|\mu_\tau - \mu^\star|(x) &\leq 213 \cdot \tau \cdot (N+1)^{-2}, \end{aligned} \tag{3.17}$$

for $r = \frac{\pi}{2(N+1)}$.

Proof. First, we show that the solution of the regularized problem admits the following property

$$\begin{aligned} \|\mu_\tau\|_{TV} &\leq \|\mu^\star\|_{TV} + \frac{\tau}{2}, \\ \|\mathcal{S}_N^*(\mu_\tau - \mu^\star)\|_{L^2(SO(3))} &\leq 2\tau. \end{aligned} \tag{3.18}$$

Since μ_τ is a solution of the regularized problem, we have

$$\tau\|\mu_\tau\|_{TV} \leq \tau\|\mu^\star\|_{TV} + \frac{1}{2}\|\mathcal{S}_N^*(\mu^\star - \mu^\star - \eta)\|_{L^2(SO(3))}^2 \leq \tau\|\mu^\star\|_{TV} + \frac{\tau^2}{2},$$

which shows the first inequality. Again, since μ_τ is a minimizer of

$$\min_{\mu \in SO(3)} \tau\|\mu\|_{TV} + J(\mu),$$

with $J(\mu) = \frac{1}{2}\|\mathcal{S}_N^*(\mu - \mu^\star - \eta)\|_{L^2(SO(3))}^2$, we know by Fermat's rule that

$$0 \in \partial(\tau\|\cdot\|_{TV} + J(\cdot))(\mu_\tau),$$

and consequently by the Moreau-Rockafellar Theorem that

$$\tau^{-1} J' \in \partial\|\cdot\|_{TV}(\mu_\tau),$$

where J' denotes the Fréchet derivative of J at μ_τ. We have

$$\begin{aligned} J'(\mu) &= \mathrm{Re}\left(\langle S_N^* \mu, S_N^*(\mu_\tau - \mu^\star - \eta)\rangle_{L^2}\right), \\ &= \mathrm{Re}\left(\langle \mu, S_N^*(\mu_\tau - \mu^\star - \eta)\rangle\right). \end{aligned}$$

Since $\tau^{-1} J' \in \partial \|\mu_\tau\|_{TV}$, we know that

$$1 = \tau^{-1}\|J'\|_{\mathcal{M}'(SO(3))} \geq \tau^{-1}\|S_N^*(\mu_\tau - \mu^\star - \eta)\|_\infty.$$

This yields,

$$\|S_N^*(\mu_\tau - \mu^\star)\|_{L^2(SO(3))} \leq \|S_N^*(\mu_\tau - \mu^\star - \eta)\|_{L^2(SO(3))} + \|S_N^* \eta\|_{L^2(SO(3))} \leq 2\tau.$$

For abbreviation, we now set $v = \mu_\tau - \mu^\star$. The polar decomposition of ν yields a function sign_ν, such that $\nu = \mathrm{sign}_\nu \cdot |\nu|$. By Theorem 3.6, we find a function $q \in \Pi_N$ with $\|q\|_\infty \leq 1$, which interpolates sign_ν on $\mathcal{X}$. Moreover, by inspection of the proof of Lemma 3.5, we have that

$$|q(x)| \leq 0.96 = 1 - 0.04, \quad \text{for } x \in \left(\bigcup_j B_r(x_j)\right)^c.$$

For $x \in B_r(x_j)$, we can expand q locally in a Taylor series given by

$$\begin{aligned} q(x) &= q(x_j \exp(\omega(x_j^{-1}x)\log(x_j^{-1}x))), \\ &= q(x_j) + \frac{\omega(x_j^{-1}x)^2}{2} e(x_j^{-1}x)^T Hq(x_j \exp(\xi \log(x_j^{-1}x))) e(x_j^{-1}x), \end{aligned}$$

with $|\xi| \leq \omega(x_j^{-1}x)$. Close to an interpolation point x_j, the proof of Lemma 3.4 shows that the Hessian is negative definite in the case $x_j = 1$ and positive definite in the case $x_j = -1$. We therefore have,

$$|q(x)| \leq \begin{cases} 1 + \frac{\omega(x_j^{-1}x)^2}{2}\tau_{max}(Hq(x_j \exp(\xi \log(x_j^{-1}x)))), & x_j = 1, \\ 1 + \frac{\omega(x_j^{-1}x)^2}{2}\tau_{min}(Hq(x_j \exp(\xi \log(x_j^{-1}x)))), & x_j = -1, \end{cases}$$

and in both cases the absolute value of the eigenvalues can be bounded by $2.36 \cdot 10^{-2} \cdot (N+1)^2$, yielding

$$|q(x)| \leq 1 - 1.18 \cdot 10^{-2} \cdot (N+1)^2 \omega(x_j^{-1}x)^2, \quad \text{for } x \in \bigcup_j B_r(x_j).$$

With these properties of q we have

$$\begin{aligned} \int_{\mathcal{X}} \mathrm{d}|\nu| &= \int_{\mathcal{X}} q(x)\mathrm{d}\nu(x), \\ &\leq \left|\int_{SO(3)} q(x)\mathrm{d}\nu(x)\right| + \left|\int_{\left(\bigcup_j B_r(x_j)\right)^c} q(x)\mathrm{d}\nu(x)\right| + \left|\sum_j \int_{B_r(x_j)\setminus\{x_j\}} q(x)\mathrm{d}\nu(x)\right|, \\ &\leq \left|\int_{SO(3)} q(x)\mathrm{d}\nu(x)\right| + \int_{\mathcal{X}^c} \mathrm{d}|\nu| - 0.04 \cdot \int_{\left(\bigcup_j B_r(x_j)\right)^c} \mathrm{d}|\nu| \\ &\quad - 1.18 \cdot 10^{-2} \cdot (N+1)^2 \sum_j \int_{\bigcup_j B_r(x_j)} \omega(x_j^{-1}x)^2 \mathrm{d}|\nu|(x). \end{aligned}$$

Rearranging yields

$$\sum_j \int_{\bigcup_j B_r(x_j)} \omega(x_j^{-1}x)^2 \mathrm{d}|\nu|(x)$$
$$\leq 85 \cdot (N+1)^{-2} \cdot \left(\left| \int_{SO(3)} q(x)\mathrm{d}\nu(x) \right| + \int_{\mathcal{X}^c} \mathrm{d}|\nu| - \int_{\mathcal{X}} \mathrm{d}|\nu| \right)$$

and

$$\int_{\left(\bigcup_j B_r(x_j)\right)^c} \mathrm{d}|\nu|$$
$$\leq 40 \cdot \left(\left| \int_{SO(3)} q(x)\mathrm{d}\nu(x) \right| + \int_{\mathcal{X}^c} \mathrm{d}|\nu| - \int_{\mathcal{X}} \mathrm{d}|\nu| \right).$$

For the first term on the right hand side, we have using (3.18)

$$\left| \int_{SO(3)} q(x)\mathrm{d}\nu(x) \right| = |\langle \nu, q\rangle| = |\langle \nu, S_N q\rangle| = |\langle S_N^* \nu, q\rangle_{L^2}|,$$
$$\leq \|S_N^*\nu\|_{L^2(SO(3))} \cdot \|q\|_{L^2(SO(3))} \leq 2\tau.$$

For the other terms on the right hand side, observe that since $\mu^\star$ is supported in $\mathcal{X}$, the inverse triangle inequality shows

$$\int_{\mathcal{X}} \mathrm{d}|\mu^\star| + \int_{\mathcal{X}^c} \mathrm{d}|\nu| - \int_{\mathcal{X}} \mathrm{d}|\nu| \leq \|\mu^\star + \nu\|_{TV} = \|\mu_\tau\|_{TV}.$$

Since $\|\mu_\tau\|_{TV} \leq \|\mu^\star\|_{TV} + \tau/2$, it follows, that

$$\int_{\mathcal{X}^c} \mathrm{d}|\nu| - \int_{\mathcal{X}} \mathrm{d}|\nu| \leq \frac{\tau}{2}.$$

This means,

$$\left(\left| \int_{SO(3)} q(x)\mathrm{d}\nu(x) \right| + \int_{\mathcal{X}^c} \mathrm{d}|\nu| - \int_{\mathcal{X}} \mathrm{d}|\nu| \right) \leq 2.5\tau$$

and the estimates (3.17) follows. □

Lemma 3.9. *Suppose the points $\mathcal{X} = \{x_j\}$ obey the separation condition* (3.16)*. Then there is a function $f \in \Pi_N$, with $\|f\|_\infty \leq 23 \cdot s_{L,N} \cdot \|\tilde{K}_L\|_\infty$, which fulfills for all $x \in SO(3)$*

$$|f(yx) - K_L(x,y)| \leq (1.5 s_{L,N}^2 + 153 s_{L,N}) \cdot \|\tilde{K}_L\|_\infty (N+1)^2 \omega(x_j^{-1}y),$$

for all $y \in B_r(x_j)$ and all $x_j \in \mathcal{X}$ with $r = \frac{\pi}{2(N+1)}$.

Proof. The proof is based on Taylor expansion locally around each $x_j \in \mathcal{X}$. Under the separation condition on the points x_j, we know by Theorem 3.2, that there is for each $x \in SO(3)$ a function $f_x \in \Pi_N$, which fulfils the following interpolation conditions

$$f_x(x_j) = K_L(x, x_j),$$
$$X_i f_x(x_j) = (\nabla K_L(x, x_j))_i = -\tilde{K}_L'(\omega(x_j^{-1}x)) e_i(x_j^{-1}x), \quad i = 1,2,3,$$

for $x_j \in \mathcal{X}$. Indeed, we again set the parameters to $s = 8$ and $b = 28$ and f has the form

$$f_x(y) = \sum_k \left(\alpha_{0,k}\sigma_N(y, x_k) + \sum_{n=1}^{3} \alpha_{n,k} X_n^y \sigma_N(y, x_k) \right),$$

such that

$$\begin{aligned} \|\alpha_0\|_\infty &\leq \frac{106 \cdot s_{L,N}}{100 - c_8} \|\tilde{K}_L\|_\infty \leq 1.07 \cdot s_{L,N} \cdot \|\tilde{K}_L\|_\infty, \\ \|\alpha_n\|_\infty &\leq \frac{s_{L,N}}{(N+1)(100 - c_8)} \|\tilde{K}_L\|_\infty \left(2 + \frac{112}{c_8}\right), \\ &\leq 21 \cdot s_{L,N} \cdot \frac{\|\tilde{K}_L\|_\infty}{(N+1)}, \quad n = 1, 2, 3, \end{aligned} \tag{3.19}$$

with $c_8 = \frac{0.999}{18}$. Thus, Taylor expansion around $x_j \in \mathcal{X}$ yields

$$|f_x(y) - K_L(x, y)| \leq \frac{\omega(x_j^{-1}y)^2}{2} \left| e(x_j^{-1}y)^T (Hf_x(z_j) - HK_L(x, z_j)) e(x_j^{-1}y) \right|,$$

with $z_j = x_j \exp(\xi_j \log(x_j^{-1}y))$ and $|\xi_j| \leq \omega(x_j^{-1}y)$, meaning $\omega(x_j^{-1}z_j) \leq \omega(x_j^{-1}y) \leq \frac{\pi}{2(N+1)}$. Now we estimate the spectral radius of $Hf_x(z_j)$ respectively $HK_L(x, z_j)$ using the Theorem of Gershgorin. For the diagonal entries of $Hf_x(z_j)$, we get

$$\begin{aligned} X_i X_i f_x(z_j) = \alpha_{0,j} X_i^x X_i^x \sigma_N(z_j, x_j) + \sum_{n=1}^{3} \alpha_{n,j} X_i^x X_i^x X_n^y \sigma_N(z_j, x_j) \\ \sum_{k \neq j} \left(\alpha_{0,k} X_i^x X_i^x \sigma_N(z_j, x_k) + \sum_{n=1}^{3} \alpha_{n,k} X_i^x X_i^x X_n^y \sigma_N(z_j, x_k) \right). \end{aligned}$$

As seen in Lemma 2.5, we have for the interpolation point close to z_j

$$\begin{aligned} |X_i^x X_i^x \sigma_N(z_j, x_j)| &\leq |\tilde{\sigma}_N''(0)| \leq \tilde{c}_8 (N+1)^2 \leq 0.06 \cdot (N+1)^2, \\ |X_i^x X_i^x X_n^y \sigma_N(z_j, x_j)| &\leq \tilde{d}_8 \left((N+1)^3 \frac{\pi}{2} + \frac{1}{4}(N+1)^2 \frac{\pi^2}{4} \right) + \frac{\tilde{c}_8}{4}(N+1)\frac{\pi}{2} \leq 1.61 \cdot (N+1)^3, \end{aligned}$$

with $\tilde{c}_8 = \frac{1.001}{18}$ and $\tilde{d}_8 = \frac{3 \cdot 1.001}{360}$. Moreover, applying Lemma 2.6 with the parameters $s = 8$, $\nu = 36$, $b = 28$ and $\varepsilon = \frac{\pi}{2 \cdot \nu}$ yields for the remaining interpolation points

$$\begin{aligned} \sum_{k \neq j} |X_i^x X_i^x \sigma_N(z_j, x_k)| &\leq \frac{c_8 a_\varepsilon (N+1)^2}{124 \cdot \zeta(6) \cdot b} \leq 3 \cdot 10^{-3} \cdot (N+1)^2, \\ \sum_{n=1}^{3} \sum_{k \neq j} |X_i^x X_i^x X_n^y \sigma_N(z_j, x_k)| &\leq \frac{8.4 \cdot c_8 a_\varepsilon (N+1)^3}{124 \cdot \zeta(6) \cdot b} \leq 3 \cdot 10^{-2} \cdot (N+1)^3. \end{aligned}$$

Combining these estimates results in

$$|X_i X_i f_x(z_j)| = \|\alpha_0\|_\infty |X_i^x X_i^x \sigma_N(z_j, x_j)| + \sum_{n=1}^{3} \|\alpha_n\|_\infty |X_i^x X_i^x X_n^y \sigma_N(z_j, x_j)|$$

$$\sum_{k\neq j}\left(\|\alpha_0\|_\infty |X_i^x X_i^x \sigma_N(z_j,x_k)| + \sum_{n=1}^{3}\|\alpha_n\|_\infty |X_i^x X_i^x X_n^y \sigma_N(z_j,x_k)|\right),$$
$$\leq 102\cdot s_{L,N}\cdot \|\tilde{K}_L\|_\infty \cdot (N+1)^2.$$

In the same way, we can establish estimates for the off-diagonal entries of $Hf_x(z_j)$, given by

$$X_jX_if_x(y) \mp \frac{1}{2}X_nf_x(y) = \sum_k \alpha_{j,0}\left(X_j^xX_i^x\sigma_N(y,x_k) \mp \frac{1}{2}X_n^x\sigma_N(y,x_k)\right)$$
$$+\sum_{n=1}^{3}\alpha_{n,k}\left(X_j^xX_i^xX_k^y\sigma_N(y,x_k) \mp \frac{1}{2}X_n^xX_k^x\sigma_N(y,x_k)\right).$$

For the interpolation point near z_j we have

$$|X_j^xX_i^x\sigma_N(x,x_j) \mp \frac{1}{2}X_n^x\sigma_N(x,x_j)| \leq 0.03\cdot(N+1)^2,$$
$$|X_j^xX_i^xX_k^y\sigma_N(x,x_j) \mp \frac{1}{2}X_n^xX_k^x\sigma_N(x,x_j)| \leq 1.61\cdot(N+1)^3$$

and for the points separated from x_j

$$\sum_{k\neq j}|X_j^xX_i^x\sigma_N(x,x_k) \mp \frac{1}{2}X_n^x\sigma_N(x,x_k)| \leq 3\cdot 10^{-3}\cdot(N+1)^2,$$
$$\sum_{n=1}^{3}\sum_{k\neq j}|X_j^xX_i^xX_k^y\sigma_N(x,x_k) \mp \frac{1}{2}X_n^xX_k^x\sigma_N(x,x_k)| \leq 3\cdot 10^{-2}\cdot(N+1)^3.$$

Again, this yields

$$|X_jX_if_x \mp \frac{1}{2}X_nf_x| \leq 102\cdot s_{L,N}\cdot\|\tilde{K}_L\|_\infty\cdot(N+1)^2.$$

Using the Theorem of Gershgorin, we therefore have the bound

$$\left|e(x_j^{-1}y)^T Hf_x(z_j)e(x_j^{-1}y)\right| \leq 306\cdot s_{L,N}\cdot\|\tilde{K}_L\|_\infty\cdot(N+1)^2.$$

In a similar way, we proceed with the spectral radius of $HK_L(x,z_j)$. We have to distinguish between two cases, $\omega(x^{-1}z_j) > \frac{\pi}{2(N+1)}$ and $\omega(x^{-1}z_j) \leq \frac{\pi}{2(N+1)}$. In the first case, we know by inspection of the proofs of Theorem 2.3 and Lemma 2.4 that

$$|X_i^xX_i^yK_L(x,z_j)| \leq (1-e_i^2)(N+1)|\tilde{K}_L^{'}(\omega(x^{-1}z_j))| + e_i^2|\tilde{K}_L^{''}(\omega(x^{-1}z_j))|,$$
$$|X_i^xX_n^xK_L(x,z_j) \mp \frac{1}{2}X_j^xK_L(x,z_j)| \leq \frac{1}{2}\left(|\tilde{K}_L^{''}(\omega(x^{-1}z_j))| + (N+1)|\tilde{K}_L^{'}(\omega(x^{-1}z_j))|\right),$$

where $e_i = e_i(x^{-1}z_j)$ is the i-th component of the rotation axis of $x^{-1}z_j$. Since $\tilde{K}_L$ is a trigonometric polynomial of degree L, so are $\tilde{K}_L'$ and $\tilde{K}_L''$ and we have by the Bernstein inequality for trigonometric polynomials

$$\|\tilde{K}_L''\|_\infty \leq L\|\tilde{K}_L'\|_\infty \leq L^2\|\tilde{K}_L\|_\infty.$$

This yields

$$|X_i^xX_i^yK_L(x,z_j)| \leq s_{L,N}^2\cdot(N+1)^2\|\tilde{K}_L\|_\infty,$$

$$|X_i^x X_n^x K_L(x, z_j) \mp \frac{1}{2} X_j^x K_L(x, z_j)| \leq s_{L,N}^2 \cdot (N+1)^2 \|\tilde{K}_L\|_\infty,$$

and consequently

$$\left| e(x_j^{-1}y)^T H K_L(x, z_j) e(x_j^{-1}y) \right| \leq 3 \cdot s_{L,N}^2 \cdot (N+1)^2 \|\tilde{K}_L\|_\infty.$$

In the second case, i.e. $\omega(x^{-1}z_j) \leq \frac{\pi}{2(N+1)}$, we have, as seen in the proof of Lemma 2.5,

$$|X_i^x X_i^x K_L(x, z_j)| \leq |\tilde{K}_L''(0)|,$$
$$|X_i^x X_n^x K_L(x, z_j) \mp X_j^x K_L(x, z_j)| \leq \frac{1}{2} |\tilde{D}_M''(0) - \tilde{K}_L''(\omega)| \leq |\tilde{K}_L''(0)|,$$

which yields the same bound as for the first case. Combining these results yields

$$|f_x(y) - K_L(x, y)| \leq (1.5 s_{L,N}^2 + 153 s_{L,N}) \cdot \|\tilde{K}_L\|_\infty (N+1)^2 \omega(x_j^{-1}y).$$

In addition, we see that $f_x(y) = f_{\mathrm{u}}(yx)$, where $u \in SO(3)$ is the identity. Setting $f = f_u$, it remains to show the bound for its maximal value. We proceed in the same way as in Lemma 3.5. We split the argument into three cases in order to bound $|f(x)|$ for $x \in SO(3)$. Namely, first there is an interpolation point x_j such that $\omega(x_j^{-1}x) \leq \frac{2.45\pi}{(N+1)}$, second we have $\frac{2.45\pi}{(N+1)} < \omega(x_j^{-1}x) \leq \frac{18}{N+1}$ and third all interpolation points are sufficiently separated, i.e $\omega(x_k^{-1}x) > \frac{18}{N+1}$ for all $x_k \in \mathcal{X}$. For the first case, we have

$$|\sigma_N(x, x_j)| \leq 1,$$
$$|X_n^y \sigma_N(x, x_j)| \leq \tilde{c}_8 (N+1)^2 \omega(x_j^{-1}x) \leq \frac{(N+1)}{2},$$

as well as using Lemma 2.6

$$\sum_{x_k \neq x_j} |\sigma_N(x, x_k)| \leq 2 \cdot 10^{-3},$$
$$\sum_{x_k \neq x_j} |X_n^y \sigma_N(x, x_k)| \leq 4 \cdot 10^{-3} \cdot (N+1).$$

Together with the bounds (3.19) on the coefficients of f we get

$$|f(x)| \leq 12 \cdot s_{L,N} \cdot \|\tilde{K}_L\|_\infty.$$

For the second case, we have using the bound of Theorem 2.3

$$|\sigma_N(x, x_j)| \leq 0.86,$$

as well as using the Bernstein inequality (1.14)

$$|X_n^y \sigma_N(x, x_j)| \leq (N+1) \cdot \|\sigma_N(\cdot, x_j)\|_\infty = (N+1),$$

Again, applying Lemma 2.6 results in

$$\sum_{x_k \neq x_j} |\sigma_N(x, x_k)| \leq 3 \cdot 10^{-2},$$
$$\sum_{x_k \neq x_j} |X_k^x \sigma_N(x, x_k)| \leq 6 \cdot 10^{-2} \cdot (N+1),$$

which yields

$$|f(x)| \leq 23 s_{L,N} \cdot \|\tilde{K}_L\|_\infty.$$

In the last case, i.e. $\omega(x_k^{-1}x) > \frac{18}{N+1}$ for all $x_k \in \mathcal{X}$, we simply use the estimates of Lemma 2.6 for the set $\mathcal{X} \cup \{x\}$ to derive

$$|f(x)| \leq 4 s_{L,N} \cdot \|\tilde{K}_L\|_\infty.$$

□

After establishing the results from the previous two lemmas, we can now state the main result of this section, i.e. bounding the super-resolution error in the L^∞ norm by a constant depending on the kernel and the super-resolution factor times the noise level.

Theorem 3.10. *Suppose the support $\mathcal{X} = \{x_j\}$ of the measure $\mu^\star$ obeys the separation condition* (3.16)*. Then the unique solution μ_τ, $\tau > 0$, of the minimization*

$$\min_{\mu \in \mathcal{M}(SO(3),\mathbb{R})} \frac{1}{2}\|\mathcal{S}_N^*(\mu - \mu^\star - \eta)\|^2_{L^2(SO(3))} + \tau\|\mu\|_{TV},$$

with $\|\mathcal{S}_N^*\eta\|_{L^2(SO(3))} \leq \tau$ fulfills for $L \geq N$

$$\|\mathcal{K}_L^*(\mu^\star - \mu_\tau)\|_\infty \leq 320 \cdot p(s_{L,N}) \cdot \|\tilde{K}_L\|_\infty \cdot \tau, \tag{3.20}$$

with $p(s_{L,N}) = (s_{L,N}^2 + 110 s_{L,N} + 1)$.

Proof. Again, we write $\nu = \mu_\tau - \mu^\star$ for abbreviation. For $x \in SO(3)$ we have, using Lemma 3.8,

$$\begin{aligned}
|\mathcal{K}_L^*\nu(x)| &= \left| \int_{SO(3)} K_L(x,y)\mathrm{d}\nu(y) \right|, \\
&\leq \left| \sum_j \int_{B_r(x_j)} K_L(x,y)\mathrm{d}\nu(y) \right| + \sup_{y \in SO(3)} |K_L(x,y)| \int_{\left(\bigcup_j B_r(x_j)\right)^c} \mathrm{d}|\nu|, \\
&\leq \left| \sum_j \int_{B_r(x_j)} K_L(x,y)\mathrm{d}\nu(y) \right| + 100 \cdot \|\tilde{K}_L\|_\infty \cdot \tau.
\end{aligned}$$

For the first term, we use the low frequency function $f \in \Pi_N$ of Lemma 3.9 to derive

$$\begin{aligned}
&\left| \sum_j \int_{B_r(x_j)} K_L(x,y)\mathrm{d}\nu(y) \right| \\
&\leq \sum_j \int_{B_r(x_j)} |f(yx) - K_L(x,y)|\mathrm{d}|\nu|(y) + \left| \sum_j \int_{B_r(x_j)} f(yx)\mathrm{d}\nu(y) \right|, \\
&\leq (1.5 s_{L,N}^2 + 153 s_{L,N}) \cdot \|\tilde{K}_L\|_\infty (N+1)^2 \sum_j \int_{B_r(x_j)} \omega(x_j^{-1}y)^2 \mathrm{d}\nu(y) \\
&\quad + \left| \sum_j \int_{B_r(x_j)} f(yx)\mathrm{d}\nu(y) \right|.
\end{aligned} \tag{3.21}$$

With Lemma 3.8, we see that

$$
\begin{aligned}
&(1.5s_{L,N}^2 + 153s_{L,N}) \cdot \|\tilde{K}_L\|_\infty (N+1)^2 \sum_j \int_{B_r(x_j)} \omega(x_j^{-1}y)^2 \mathrm{d}\nu(y) \\
&\leq 213 \cdot (1.5s_{L,N}^2 + 153s_{L,N}) \cdot \|\tilde{K}_L\|_\infty \cdot \tau.
\end{aligned}
$$

It remains to estimate the second term of the right hand side of (3.21). We have,

$$
\begin{aligned}
\left|\sum_j \int_{B_r(x_j)} f(yx)\mathrm{d}\nu(y)\right| &\leq \left|\int_{\left(\bigcup_j B_r(x_j)\right)^c} f(yx)\mathrm{d}\nu(y)\right| + \left|\int_{SO(3)} f(yx)\mathrm{d}\nu(y)\right|, \\
&\leq \|f\|_\infty \int_{\left(\bigcup_j B_r(x_j)\right)^c} \mathrm{d}|\nu|(y) + \left|\int_{SO(3)} f(yx)\mathrm{d}\nu(y)\right|.
\end{aligned}
$$

Again, the first term can be estimated using Lemma 3.8 and the supremum norm bound on the function f derived in Lemma 3.9, i.e.

$$
\|f\|_\infty \int_{\left(\bigcup_j B_r(x_j)\right)^c} \mathrm{d}|\nu|(y) \leq 23 \cdot 10^2 \cdot s_{L,N} \cdot \|\tilde{K}_L\|_\infty \cdot \tau.
$$

For the second term, observe, since the function $f_x = f(\cdot x) \in \Pi_N$ for all $x \in SO(3)$, that

$$
\begin{aligned}
\left|\int_{SO(3)} f(yx)\mathrm{d}\nu(y)\right| &= |\langle \nu, f_x \rangle| = |\langle \nu, \mathcal{S}_N f_x \rangle| = |\langle \mathcal{S}_N^* \nu, f_x \rangle_{L^2}|, \\
&\leq \|f\|_{L^2} \|\mathcal{S}_N^* \nu\|_{L^2} \leq \|f\|_\infty \|\mathcal{S}_N^* \nu\|_{L^2}, \\
&\leq 23 \cdot s_{L,N} \cdot \|\tilde{K}_L\|_\infty \cdot \|\mathcal{S}_N^* \nu\|_{L^2}.
\end{aligned}
$$

As seen in the proof of Lemma 3.8, see (3.18), we know that $\|\mathcal{S}_N^* \nu\|_{L^2} \leq 2\tau$ and therefore

$$
\left|\int_{SO(3)} f(yx)\mathrm{d}\nu(y)\right| \leq 46 \cdot s_{L,N} \cdot \|\tilde{K}_L\|_\infty \cdot \tau.
$$

Thus

$$
\left|\sum_j \int_{B_r(x_j)} f(yx)\mathrm{d}\nu(y)\right| \leq 23 \cdot 102 \cdot s_{L,N} \cdot \|\tilde{K}_L\|_\infty \cdot \tau
$$

and combining this with the estimate (3.21) yields

$$
\left|\sum_j \int_{B_r(x_j)} K_L(x,y)\mathrm{d}\nu(y)\right| \leq 320 \cdot (s_{L,N}^2 + 110s_{L,N}) \cdot \|\tilde{K}_L\|_\infty \cdot \tau.
$$

Finally, we therefore have the bound

$$
\|\mathcal{K}_L^* \nu\|_\infty \leq 320 \cdot (s_{L,N}^2 + 110s_{L,N} + 1) \cdot \|\tilde{K}_L\|_\infty \cdot \tau.
$$

□

Observe, that by sending τ to zero in an appropriate way and L to infinity , this results in the noise free recovery case.

Notes and References. *Measuring the error of the recovery in the noisy case by measuring the distance in higher frequencies was first considered for trigonometric moments in [Candés and Fernandez-Granda, 2013]. Next to* L^1 *error estimates for a deterministic noise model, the authors considered a Gaussian noise model. We believe, that those kind of estimates can be transferred to signals on the rotation group.*

The derived estimates of this section are in the spirit of the estimates in [Li, 2017] for the case of trigonometric moments. Nevertheless, the generalization to the super-resolution problem on the rotation group heavily relies on the localization estimates of Chapter 2 and the construction of interpolating functions in Section 3.1.

Chapter 4

Numerical Solution of the Optimization Problem

In the previous chapter, we have seen that the sought measure is uniquely determined as the solution of a total variation minimization on the space of signed Borel measures. Nevertheless, the minimization takes place in an infinite-dimensional vector space and is therefore not feasible at a first sight.

In this chapter, we present two different approaches to tackle the minimization problem

$$\min_{\mu \in \mathcal{M}(SO(3))} \|\mu\|_{TV}, \quad \text{subject to} \quad \mathcal{S}_N^* \mu = \mathcal{S}_N^* \mu^\star, \tag{RP}$$

respectively the Thikonov-type problem

$$\min_{\mu \in \mathcal{M}(SO(3))} \frac{1}{2} \|\mathcal{S}_N^*(\mu - \mu^\star - \eta)\|_{L^2(SO(3))}^2 + \tau \|\mu\|_{TV}. \tag{RP$_\tau$}$$

In the first approach, we solve a semi-definite relaxation of the dual problems, given by

$$\max_{f \in \Pi_N} \operatorname{Re}\langle f, \mathcal{S}_N^* \mu^\star \rangle, \quad \text{subject to} \quad \|f\|_\infty \leq 1, \tag{dRP}$$

respectively

$$\max_{f \in \Pi_N} \operatorname{Re}\langle \mathcal{S}_N^*(\mu^\star + \eta), f \rangle - \tau \|f\|_{L^2(SO(3))}^2, \quad \text{subject to} \quad \|f\|_\infty \leq 1. \tag{dRP$_\tau$}$$

Observe, that, although the objective of the pre-dual problem is finite-dimensional, the side condition is an infinite-dimensional constraint. We substitute this infinite-dimensional constraint with a *sufficient* finite-dimensional constraint, following the works [Candés and Fernandez-Granda, 2013] for trigonometric moments and [Bendory et al., 2015b] for spherical harmonics.

In the second approach, w discretize the primal problems beforehand and solve the corresponding finite-dimensional optimization problem. In the analysis of the convergence of this process, we build on results stated in [Tang et al., 2013].

4.1 Semi-Definite Formulation

To start this section, we recap the super-resolution problem in the case of Fourier moments, as it was solved in [Candés and Fernandez-Granda, 2014]. It is based on a *Bounded Real Lemma*, derived from Gramian parametrizations of non-negative polynomials and the Fejér-Riesz theorem. A good overview on this topic can be found in [Dumitrescu, 2007]. For the sake of completeness, we present the essential ingredients to formulate our first algorithm.

4.1.1 Recovery from Trigonometric Moments and the Bounded Real Lemma

The minimization problem, given by

$$\min_{\mu\in\mathcal{M}(\mathbb{T})} \|\mu\|_{TV}, \quad \text{subject to} \quad \hat{\mu}(k) = y_k, \quad \text{for } |k| \le N, \tag{TP}$$

is an infinite-dimensional optimization problem over the whole measure space $\mathcal{M}(\mathbb{T})$ and thus not directly tractable. The convex pre-dual problem to (TP), is given by

$$\max_{c\in\mathbb{C}^{(2N+1)}} \mathrm{Re}(\langle c, y\rangle), \quad \text{such that} \quad \|F_N c\|_\infty \le 1, \tag{dTP}$$

where $F_N c(t) = \sum_{k=-N}^{N} c_k e^{ikt}$ is the Fourier summation operator. In the pre-dual problem, the objective is finite-dimensional, but the constraints are infinite-dimensional. Since $\mathrm{ran}(F_N)$ is finite-dimensional, the pre-dual problem always has a solution. Moreover, due to a Slater condition, we now that the duality gap is zero, see Theorem D.4, which implies any solution $c^\star \in \mathbb{C}^{2N+1}$ obeys

$$\mathrm{Re}(\langle c^\star, y\rangle) = \mathrm{Re}(\langle c^\star, F_N^* \mu^\star\rangle) = \mathrm{Re}(\langle \mu^\star, F_N c^\star\rangle) = \|\mu^\star\|_{TV},$$

where $\mu^\star$ is a solution to the primal problem (TP). In the case $\mu^\star$ is discrete, this leads to the interpolation

$$F_N c^\star(x_i) = \mathrm{sign}_{\mu^\star}(x_i),$$

as seen in (1.6). This means, the support set of $\mu^\star$ is contained in the set of zeros of

$$1 - |F_N c^\star(t)|^2.$$

In the univariate case, one can now obtain the zeros by computing the unit magnitude eigenvalues of the companion matrix. Afterwards, one can obtain the amplitudes by solving a least squares problem. Observe, that the side condition is equivalent to the non-negativity of the trigonometric polynomial

$$R(t) = 1 - |F_N c(t)|^2. \tag{4.1}$$

The Fejér-Riesz theorem states, that a trigonometric polynomial R of degree n, i.e.

$$R(t) = \sum_{k=-n}^{n} r_k \mathrm{e}^{-ikt}, \quad t \in \mathbb{T},$$

is non-negative, if and only if there is a causal trigonometric polynomial H, i.e. it has the form

$$H(t) = \sum_{k=0}^{n} h_k \mathrm{e}^{-ikt}, \tag{4.2}$$

such that

$$R(t) = H(t)\cdot\overline{H(t)} = |H(t)|^2. \tag{4.3}$$

If we write

$$H(t) = \varphi_n(t)^H h,$$

where

$$\varphi_n(t) = \begin{pmatrix} 1 \\ \mathrm{e}^{it} \\ \vdots \\ \mathrm{e}^{int} \end{pmatrix}, \quad h = \begin{pmatrix} h_0 \\ h_1 \\ \vdots \\ h_n \end{pmatrix},$$

then equation (4.3) reads as

$$R(t) = \varphi(t)^H h h^H \varphi(t).$$

On the other hand, each trigonometric polynomial of degree n of the form

$$R(t) = \varphi_n(t)^H Q \varphi_n(t), \tag{4.4}$$

where $Q \in \mathbb{C}^{(n+1)\times(n+1)}$ is hermitian positive semi-definite is non-negative. The matrix Q is called a *Gramian matrix* representing R. There is the following connection between the coefficients of R and any Gramian matrix representing R.

Theorem 4.1 ([Dumitrescu, 2007, Thm. 2.3]). *If $R(t) = \sum_{k=-n}^{n} r_k \mathrm{e}^{-ikt}$ and Q is a Gramian matrix representing R, then*

$$r_k = \operatorname{tr}(\Theta_k Q) = \sum_{i=\max(0,k)}^{\min(n+k,n)} Q_{i,i-k}, \quad k = -n, \dots, n, \tag{4.5}$$

where Θ_k is the elementary Toeplitz matrix with ones on the k-th diagonal and zeros elsewhere.

This is known as trace parametrization of a non-negative trigonometric polynomial. Now, again consider the case

$$R(t) = 1 - |F_N c(t)|^2.$$

This can be written in the form

$$R(t) = 1 - |\mathrm{e}^{-iNt} \cdot F_N c(t)|^2,$$

which means $R(t) = 1 - |H_0(t)|^2$, where

$$\begin{aligned} H_0(t) &= \mathrm{e}^{-iNt} \cdot F_N c(t), \\ &= \sum_{k=0}^{2N} (h_0)_k \mathrm{e}^{-ikt} = \varphi(t)^H h_0, \end{aligned}$$

with $(h_0)_k = c_{k-N}$ is a monic polynomial of the form (4.2), with $n = 2N$. The Fejér-Riesz Theorem now states, that there has to be a positive semi-definite Gramian Q_0 representing the polynomial identically to one, i.e.

$$\delta_{0,k} = \operatorname{tr}(\Theta_k Q_0), \quad k = -N, \dots, N, \tag{4.6}$$

sucht that

$$R(t) = 1 - |H_0(t)|^2 = \varphi(t)^H (Q_0 - h_0 h_0^H) \varphi(t).$$

This means $R(t) \geq 0$, if and only if

$$Q = Q_0 - h_0 h_0^H$$

is positive semidefinite. Observe, that given the matrix

$$\begin{pmatrix} Q_0 & h_0 \\ h_0^H & 1 \end{pmatrix}, \tag{4.7}$$

the matrix Q is the Schur complement of the block '1' in the matrix (4.7). Therefor, Q is positive semi-definite if and only if the matrix (4.7) is positive semi-definite, see Lemma C.1. Concluding, one has that

$R(t) = 1 - |F_N c(t)|^2$ is non-negative, if and only if, there is a positive semi-definite hermitian matrix $Q_0 \in \mathbb{C}^{(2N+1)\times(2N+1)}$ obeying

$$\begin{pmatrix} Q_0 & c \\ c^H & 1 \end{pmatrix} \succcurlyeq 0, \quad \text{such that} \quad \operatorname{tr}(\Theta_k Q_0) = \delta_{0,k}, \quad k = -N, \dots, N. \tag{4.8}$$

Consequently, the constraint $\|F_N c\|_\infty \leq 1$ in (dTP) can equivalently be written as (4.8). Thus, the dual program becomes

$$\max_{Q_0, c} \operatorname{Re}(\langle c, y \rangle), \quad \text{subject to} \quad (4.8),$$

and is therefore a finite-dimensional semi-definite program with $(2N+1)^2/2$ variables. As seen, the sufficiency is based on trace parametrizations and the necessity utilizes in addition that every univariate non-negative trigonometric polynomial is a sum of squares, which is the assertion of the Fejér-Riesz theorem.

In higher dimension, the equality of non-negative and sum of squares polynomials does no longer hold true. Thus, we cannot hope to find an equivalent characterization of the supremum norm constraint in dimensions higher than one. Nevertheless, the construction of a sufficient condition utilizing trace parametrizations is still possible by replacing the non-negativity constraint by a sum of squares assumption. This is known as *sum of squares relaxation*. For this purpose, following [Dumitrescu, 2007], consider multivariate trigonometric polynomials of the form

$$R(\mathbf{t}) = \sum_{\mathbf{k}=-n}^{n} r_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{t}}, \quad \mathbf{t} \in \mathbb{T}^d,$$

where $\mathbf{k} \in \mathbb{Z}^d$ denotes a multiindex $\mathbf{k} = (k_1, \dots, k_d)$, $\mathbf{t} = (t_1, \dots, t_d)$ and $\mathbf{k} = -n, \dots, n$ is meant elementwise. Positive orthant polynomials are of the form

$$H(\mathbf{t}) = \sum_{\mathbf{k}=0}^{n} h_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{t}}, \quad \mathbf{t} \in \mathbb{T}.$$

More generally, to state trace parametrizations of constant trigonometric polynomials, we need the notion of a halfspace. A set $\mathcal{H} \subset \mathbb{Z}^d$ is called *halfspace*, if $\mathcal{H} \cap (-\mathcal{H}) = \{0\}$, $\mathcal{H} \cup (-\mathcal{H}) = \mathbb{Z}^d$ and $\mathcal{H} + \mathcal{H} \subset \mathcal{H}$. A standard way to construct a halfspace is given iteratively. We start with $\mathcal{H}_1 = \mathbb{N}$ and say that $\mathbf{k} \in \mathcal{H}_d$, if either $k_d > 0$, or $k_d = 0$ and $(k_1, \dots, k_{d-1}) \in \mathcal{H}_{d-1}$. We say, that the trigonometric polynomial R is a *sum of squares*, if

$$R(\mathbf{t}) = \sum_{l=1}^{\nu} |H_l(\mathbf{t})|^2, \tag{4.9}$$

where H_l are positive orthant polynomials. Observe, that the maximal degree m of the trigonometric polynomials H_l are allowed to be higher than the degree of R, i.e. $m \geq n$. This additional degree of freedom allows for the following. Every *positive* trigonometric polynomial is a sum of squares polynomial, see [Dumitrescu, 2007, Thm. 3.5]. To get a more comprehensive representation in analogy to (4.4), set the $(m+1)$-dimensional vector

$$\Phi_m(\mathbf{t}) = \varphi_m(t_d) \otimes \cdots \otimes \varphi_m(t_1),$$

where $\otimes$ denotes the Kronecker product and $\varphi_m(t_j) = (1, \dots, e^{imt_j})^T$, for $m \geq n$. This is only a suitable enumeration of the exponentials. In two dimensions, for example, this would read as

$$\Phi_m(\mathbf{t}) = \varphi_m(t_2) \otimes \varphi_m(t_1) = \begin{pmatrix} 1 \cdot \varphi_m(t_1) \\ e^{it_2} \cdot \varphi_m(t_1) \\ \vdots \\ e^{imt_2} \cdot \varphi_m(t_1) \end{pmatrix}.$$

Using this notation, every positive orthant trigonometric polynomial can be written in the form

$$H(t) = \Phi_m(t)^H \mathbf{h},$$

where $\mathbf{h} \in \mathbb{C}^{(m+1)^d}$ contains the coefficients of H according to the ordering of the indices in Φ_m. This means, every sum of squares polynomial R can be written as

$$R(t) = \Phi_m(t)^H Q \Phi_m(t), \tag{4.10}$$

where $Q \in \mathbb{C}^{(m+1)^d \times (m+1)^d}$ is hermitian positive semi-definite and m is the maximal degree appearing in the representation (4.9). Again, the important part is a trace parametrization.

Theorem 4.2 ([Dumitrescu, 2007, Thm. 3.13]). *Suppose that* $R(\mathbf{t}) = \sum_{\mathbf{k}=-n}^{n} r_{\mathbf{k}} \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{t}}$ *and* $Q \in \mathbb{C}^{(m+1)^d \times (m+1)^d}$ *such that* (4.10) *holds. Then*

$$\tilde{r}_{\mathbf{k}} = \mathrm{tr}(\Theta_{\mathbf{k}} Q),$$

where

$$\Theta_{\mathbf{k}} = \Theta_{k_d} \otimes \cdots \otimes \Theta_{k_1},$$

and Θ_{k_j} *again denotes the elementary Toeplitz matrix with ones on the* k_j*-th diagonal and zeros elsewhere and* $\tilde{r} \in \mathbb{C}^{(m+1)^d}$ *denotes the continuation of* $r \in \mathbb{C}^{(m+1)^d}$ *with zeros.*

Again, consider the case that $R(\mathbf{t}) = 1 - |H_0(\mathbf{t})|^2$, where $H_0(\mathbf{t})$ is a positive orthant polynomial of degree n with coefficients $\mathbf{h}_0$. Then a sufficient condition for the non-negativity of R can be stated as follows. Find a trace parametrization of the trigonometric polynomial identically one, i.e. there is $Q_0 \in \mathbb{C}^{(m+1)^d \times (m+1)^d}$ obeying

$$\delta_{0,\mathbf{k}} = \mathrm{tr}(\Theta_{\mathbf{k}} Q_0), \quad \mathbf{k} \in \mathcal{H},\ \mathbf{k} \leq m,$$

where $\mathcal{H}$ is a halfspace, such that $Q = Q_0 - \tilde{\mathbf{h}}_0 \tilde{\mathbf{h}}_0^H \succcurlyeq 0$ and $\tilde{\mathbf{h}}_0 \in \mathbb{C}^{(m+1)^d}$ is the continuation of $\mathbf{h}_0$ with zeros. Indeed, in this case

$$R(\mathbf{t}) = 1 - |H_0(\mathbf{t})|^2 = \Phi_m(t)^H (Q_0 - \tilde{\mathbf{h}}_0 \tilde{\mathbf{h}}_0^H) \Phi_m(t) \geq 0.$$

As in the one-dimensional case, the condition

$$\|F_N c\|_\infty \leq 1,$$

where

$$F_N c(\mathbf{t}) = \sum_{\mathbf{k}=-N}^{N} c_{\mathbf{k}} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{t}},$$

is equivalently to the non-negativity of

$$R(\mathbf{t}) = 1 - |\mathrm{e}^{-iN\mathbf{t}} \cdot F_N c(\mathbf{t})|^2 = 1 - |H_0(\mathbf{t})|^2,$$

where $H_0(\mathbf{t}) = \sum_{\mathbf{k}=0}^{2N} c_{\mathbf{k}-N} \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{t}}$ is a positive orthant polynomial. The relaxation of the side condition consists of replacing the non-negativity condition with a sums of squares representation of $R(\mathbf{t})$ with squares of degree $M \geq N$, to derive a finite-dimensional *sufficient* criteria. More concrete, as seen before, the supremum norm bound is fulfilled, if there is $Q_0 \in \mathbb{C}^{(2M+1)^d \times (2M+1)^d}$ for $M \geq N$ and a halfspace $\mathcal{H}$, obeying

$$\begin{pmatrix} Q_0 & \tilde{c} \\ \tilde{c}^H & 1 \end{pmatrix} \succcurlyeq 0, \quad \text{such that} \quad \mathrm{tr}(\Theta_{\mathbf{k}} Q_0) = \delta_{0,\mathbf{k}}, \quad \mathbf{k} \in \mathcal{H},\ -M \leq \mathbf{k} \leq M,$$

where again $\tilde{c} \in \mathbb{C}^{(2M+1)^d}$ is the continuation of c by zeros. This is a generalization of (4.8) to higher dimensions, with the difference that this condition is only *sufficient*, but not *necessary*. In addition, there is an extra relaxation parameter $M \geq N$, coming from the assumption, that the polynomial $R(\mathbf{t}) = 1 - |\mathrm{e}^{-iN\mathbf{t}} \cdot F_N c(\mathbf{t})|^2$ is a sum of squares, with squares of degree M. Summarizing, to derive the non-negativity of the trigonometric polynomial one replaces the non-negativity with a sum of squares condition.

In the next section, we will see, that we only need to relax the non-negativity conditions on an algebraic subset of the torus. To be as close as possible to the original non-negativity condition, we make use of a slight generalization of the previous relaxation for globally non-negative polynomials, which is known as the *Bounded Real Lemma*. For reasons of completeness, we comprehensively state the necessary results from [Dumitrescu, 2007]. An algebraic subset $\mathcal{D} \subset \mathbb{T}^d$ has the form

$$\mathcal{D} = \{\mathbf{t} \in \mathbb{T}^d : D_l(\mathbf{t}) \geq 0, l = 1, \dots, L\}, \tag{4.11}$$

where $D_l(\mathbf{t}) = \sum_{\mathbf{k}} d_{l,\mathbf{k}} \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{t}}$ are given trigonometric polynomials. If a trigonometric polynomial $R(\mathbf{t}) = \sum_{\mathbf{k}=-n}^{n} r_{\mathbf{k}} \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{t}}$ has the representation

$$R(\mathbf{t}) = S_0(\mathbf{t}) + \sum_{l=1}^{L} D_l(\mathbf{t}) S_l(\mathbf{t}), \tag{4.12}$$

where S_l, $l = 0, \dots, L$, are sum of squares polynomials, then $R(\mathbf{t}) \geq 0$ for $\mathbf{t} \in \mathcal{D}$. Again, one assumes an upper bound on the degree of the polynomials S_l. If $\deg(R) = n$ and $\deg(S_0) = m$, with $m \geq n$, then one choice is $\deg(S_l) = m - \deg(D_l)$, such that the products $D_l S_l$ have degree m, see [Dumitrescu, 2007, Rem. 4.17]. In this case, the following trace representation is valid and can be derived from the previous trace parametrization, see [Dumitrescu, 2007, Thm. 4.15].

$$\tilde{r}_{\mathbf{k}} = \operatorname{tr}(\Theta_{\mathbf{k}} Q_0 + \sum_{l=1}^{L} \Psi_{l,\mathbf{k}} Q_l), \quad \mathbf{k} \in \mathcal{H},\ -m \leq \mathbf{k} \leq m,$$

where

$$\Psi_{l,\mathbf{k}} = \sum_{\mathbf{j}+\mathbf{l}=\mathbf{k}} d_{l,\mathbf{j}} \Theta_{\mathbf{l}}. \tag{4.13}$$

The matrices $Q_l \in \mathbb{C}^{(\deg(S_l)+1)^d \times (\deg(S_l)+1)^d}$, $l = 0, \dots, L$ are Gramians associated to the polynomials S_l and the dimension of $\Theta_{\mathbf{l}}$, appearing in $\Psi_{l,\mathbf{k}}$ is given accordingly. From this, one gets an sufficient condition for a norm bound on the set $\mathcal{D}$, that is known as *Bounded Real Lemma*.

Lemma 4.3 (Bounded Real Lemma, [Dumitrescu, 2007, Cor. 4.25])**.** *Suppose $H(\mathbf{t}) = \sum_{\mathbf{k}=0}^{n} h_{\mathbf{k}} \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{t}}$ is a positive orthant trigonometric polynomial and $\mathcal{D}$ is a frequency domain as given in* (4.11)*. If there exist hermitian positive semi-definite matrices Q_l, $l = 0, \dots, L$, and a halfspace $\mathcal{H}$ for $m \geq n$, obeying*

$$\begin{pmatrix} Q_0 & \tilde{h} \\ \tilde{h}^H & 1 \end{pmatrix} \succcurlyeq 0, \quad \textit{such that} \quad \operatorname{tr}(\Theta_{\mathbf{k}} Q_0 + \sum_{l=1}^{L} \Psi_{l,\mathbf{k}} Q_l) = \delta_{0,\mathbf{k}}, \quad \mathbf{k} \in \mathcal{H},\ \mathbf{k} \leq m,$$

where $\Psi_{l,\mathbf{k}}$ are given in (4.13)*, then*

$$|H(\mathbf{t})| \leq 1, \quad \textit{for } \mathbf{t} \in \mathcal{D}.$$

In the next section, we will utilize the Bounded Real Lemma to derive our first reconstruction algorithm in the case of Wigner D-moments.

4.1.2 Recovery from Wigner D-moments

In this section, we formulate a finite-dimensional relaxation for the total variation minimization in the case of Wigner D-moments. It is based on a special representation of the Wigner D-functions in combination with the Bounded Real Lemma from the previous section.

As seen before, under a suitable separation condition for the support of the sought measure $\mu^\star$ it can be recovered as the solution of the minimization problem

$$\min_{\mu \in \mathcal{M}(SO(3))} \|\mu\|_{TV}, \quad \text{subject to} \quad \mathcal{S}_N^* \mu = g, \tag{RP}$$

where

$$\mathcal{S}_N^* \mu(x) = \sum_{l=0}^{N} (2l+1) \sum_{k,m=-l}^{l} \langle \mu, D_{k,m}^l \rangle D_{k,m}^l(x)$$

is the partial summation operator with respect to the Wigner D-functions and $g = \mathcal{S}_N^* \mu^\star$. Since the objective space is infinite-dimensional, the optimization problem is not feasible. The convex pre-dual of the minimization problem (RP) is given by

$$\max_{f \in L^2(SO(3))} \operatorname{Re}\langle g, \mathcal{S}_N f \rangle, \quad \text{subject to} \quad \|\mathcal{S}_N f\|_\infty \leq 1,$$

or equivalently

$$\max_{f \in \Pi_N} \operatorname{Re}\langle g, f \rangle, \quad \text{subject to} \quad \|f\|_\infty \leq 1, \tag{dRP}$$

where

$$\Pi_N = \operatorname{ran}(\mathcal{S}_N) = \operatorname{span}\{D_{k,m}^l : l \leq N, -l \leq k, m \leq l\}.$$

Let $f^\star$ be a solution of the pre-dual problem, which always exists. Again, by strong duality we know that

$$\operatorname{Re}\langle g, f^\star \rangle = \operatorname{Re}\langle \mu^\star, \mathcal{S}_N f^\star \rangle = \|\mu^\star\|_{TV},$$

if $\mu^\star$ is a solution of the primal problem. In the case $\mu^\star$ is discrete, we have that

$$f^\star(x_i) = \operatorname{sign}_{\mu^\star}(x_i), \quad x_i \in \operatorname{supp}(\mu^\star)$$

meaning the support points are a subset of the zeros of the function $1 - |f^\star(x)|^2 \in \Pi_{2N}$. To replace the norm constraint in the dual problem (dRP) with a finite-dimensional condition, we use Lemma 4.3 from the previous section. In order to point out this connection, remember from (1.11), that in the Euler angle parametrization $(\alpha, \beta, \gamma) \in (0, 2\pi] \times (0, \pi] \times (0, 2\pi]$ we have the representation of the Wigner D-functions

$$D_{k,m}^l(\alpha, \beta, \gamma) = \mathrm{e}^{-ik\alpha} \left(\sum_{j=-l}^{l} \hat{d}_{km}^l(j) \mathrm{e}^{-ij\beta} \right) \mathrm{e}^{-im\gamma},$$

with $\hat{d}_{km}^l(j) = i^{m+k}(-1)^j d_{-j,k}^l(0) d_{m,-j}^l(0)$. Thus, if $f \in \Pi_N$, we have in the Euler angle parametrization $(\alpha, \beta, \gamma) \in (0, 2\pi] \times (0, \pi] \times (0, 2\pi]$

$$\begin{aligned} f(x(\alpha, \beta, \gamma)) &= \sum_{l=0}^{N} (2l+1) \sum_{k,m=-l}^{l} \langle f, D_{k,m}^l \rangle D_{k,m}^l(\alpha, \beta, \gamma), \\ &= \sum_{j,k,m=-N}^{N} \left(\sum_{l=\max(|j|,|k|,|m|)}^{N} \hat{d}_{km}^l(j)(2l+1) \langle f, D_{k,m}^l \rangle \right) \mathrm{e}^{-i(j\beta + k\alpha + m\gamma)}, \end{aligned}$$

or equivalently for $(\alpha,\beta,\gamma) \in (-\pi,\pi] \times (-\pi,0] \times (-\pi,\pi]$

$$f(x(\alpha+\pi,\beta+\pi,\gamma+\pi)) = \sum_{j,k,m=-N}^{N} \left(\sum_{l=\max(|j|,|k|,|m|)}^{N} (-1)^{j+k+m} \hat{d}^l_{km}(j)(2l+1)\langle f, D^l_{k,m}\rangle \right) \cdot \mathrm{e}^{-i(j\beta+k\alpha+m\gamma)},$$

which means on the set $(-\pi,\pi] \times (-\pi,0] \times (-\pi,\pi]$ f is equal to a trigonometric polynomial. For abbreviation, define the mapping $T_M f : \Pi_N \to \mathbb{C}^{(M+1)^d}$,

$$T_M f = \begin{cases} \sum_{l=\max(|j|,|k|,|m|)}^{N} (-1)^{j+k+m} \hat{d}^l_{km}(j)(2l+1)\langle f, D^l_{k,m}\rangle, & \max(|j|,|k|,|m|) \le N, \\ 0, & \text{else}, \end{cases}$$

and the algebraic set

$$\mathcal{D} = [-\pi,\pi] \times [-\pi,0] \times [-\pi,\pi] = \left\{(\alpha,\beta,\gamma) \in \mathbb{T}^3 : -\sin(\beta) \ge 0\right\}.$$

We can write $-\sin(\beta) = \frac{i}{2}(\mathrm{e}^{i\beta} - \mathrm{e}^{-i\beta})$, thus, in the context of Lemma 4.3, we define $\Psi_{\mathbf{k}} \in \mathbb{C}^{(2M)^3 \times (2M)^3}$ for $\mathbf{k} = (j,k,m)$ as

$$\Psi_{\mathbf{k}} = \frac{i}{2}\Theta_{\mathbf{k}-\begin{pmatrix}1\\0\\0\end{pmatrix}} - \frac{i}{2}\Theta_{\mathbf{k}+\begin{pmatrix}1\\0\\0\end{pmatrix}}, \tag{4.14}$$

with the convention that $\Theta_{\mathbf{l}} = 0$, if one of the entries of $\mathbf{l}$ exceeds $2M-1$ in absolute value. Concluding, we have the following Lemma.

Lemma 4.4. *Suppose $f \in \Pi_N$. If for $M \ge N$ there are hermitian positive semi-definite matrices $Q_0 \in \mathbb{C}^{(2M+1)^3 \times (2M+1)^3}$ and $Q_1 \in \mathbb{C}^{(2M)^3 \times (2M)^3}$ and a halfspace $\mathcal{H}$ such that*

$$\begin{aligned} &\operatorname{tr}(\Theta_{\mathbf{k}} Q_0 + \Psi_{\mathbf{k}} Q_1) = \delta_{0,\mathbf{k}}, \quad \mathbf{k} \in \mathcal{H},\ -2M \le \mathbf{k} \le 2M, \\ &\begin{pmatrix} Q_0 & T_M f \\ T_M f^H & 1 \end{pmatrix} \succcurlyeq 0, \end{aligned} \tag{4.15}$$

where $\Psi_{\mathbf{k}}$ is given in (4.14)*, then $\|f\|_\infty \le 1$.*

Proof. The lemma follows immediately from Lemma 4.3. □

In order to formulate our first algorithm, we substitute the infinite-dimensional constraint in the dual problem (dRP) with the finite-dimensional constraint (4.15). We would like to mention again, that in contrast to the case of the univariate Fourier moments this is not an equivalent formulation but a relaxation of the pre-dual problem, which depends on the following assumption.

Assumption. *The solution $f^\star$ of the pre-dual problem* (dRP) *fulfills the following. The trigonometric polynomial with the coefficients $T_M f$ has a sums of squares representation of the form*

$$F_N T_M f^\star(\mathbf{t}) = S_0(\mathbf{t}) - \sin(t_2) S_1(\mathbf{t}), \quad \mathbf{t} \in \mathbb{T}^3, \tag{SOS}$$

where S_0 is a sum of squares polynomial of degree M and S_1 is a sum of squares polynomial of degree $M-1$.

Under which conditions on the primal solution $\mu^\star$ this assumption is valid, remains an open and difficult problem. Nevertheless, we have the following.

Theorem 4.5. *Let $f^\star \in \Pi_N$ be a solution of the pre-dual problem* (dRP)*. If $f^\star$ fulfills the Assumption* (SOS) *for $M \geq N$, then it is the solution of the problem*

$$\max_{f\in\Pi_N,Q_0,Q_1} \operatorname{Re}\langle g, f\rangle, \quad \textit{subject to} \quad (4.15), \qquad (\text{dRP}_{\text{rel}})$$

where $Q_0 \in \mathbb{C}^{(2M+1)^3\times(2M+1)^3}$ and $Q_1 \in \mathbb{C}^{(2M)^3\times(2M)^3}$ have to be hermitian positive semi-definite.

In contrast to the pre-dual problem (dRP), the optimization problem (dRP_{rel}) is inherently finite-dimensional. In the case of the pre-dual problem (dRP_τ) to the Thikonov-type problem (RP_τ), we substitute it with the minimization

$$\max_{f\in\Pi_N,Q_0,Q_1} \operatorname{Re}\langle g, f\rangle - \tau\|f\|^2_{L^2(SO(3))}, \quad \text{subject to} \quad (4.15). \qquad (\text{dRP}_{\tau,\text{rel}})$$

To solve the convex optimization problem, we use MATLAB® in combination with CVX, a free package for specifying and solving convex programs, see [Grant and Boyd, 2014] and [Grant and Boyd, 2008]. We set the preferences of CVX to call the open source solver SeDuMi, see [Sturm, 1999], at the highest precision. This solver uses a primal-dual interior point method to approximate a solution of (dRP_{rel}) from the interior of the feasible set and we will use it as a '*black box*' solver. For further details see [Sturm, 1997] and [Sturm, 1999]. For the same argument as in [Candés and Fernandez-Granda, 2014, Sec. 4], it is highly unlikely, that the result of the interior point method gives a constant function equally to one in absolute value. After solving the program (dRP_{rel}), we determine the minima of the function

$$p(x) = 1 - |f^\star(x)|^2.$$

This part is a problem of its own interest. We describe one possible solution in the following. To find all local minimizer of the function $p = 1 - |f^\star(\cdot)|^2$, we will apply a simultaneous conjugate gradient method with inexact line search using the *Wolfe condition*, as described in [Gräf and Hielscher, 2015]. For initialization, we choose randomly P initial points. For completeness, we briefly state the conjugate gradient algorithm as procedure 1. The expressions $\nabla p(x)$ and $Hp(x)$, that have to be computed, have a local representation in Euler angles, which makes it possible to use fast Fourier techniques for the simultaneous computation at different points, see [Potts et al., 2009], [Keiner and Vollrath, 2012]. For a detailed discussion on this, including convergence rates, see [Gräf and Hielscher, 2015] and [Gräf, 2013a]. For our computations, we use an implementation of this algorithm from the C++ Library *LORM* [Gräf, 2013b], provided by the authors of [Gräf and Hielscher, 2015].

Procedure 1: CG method with inexact line search

Input: $p \in C^2(SO(3)), x^{(0)} \in SO(3)$
Parameters: $0 < c_1 < \frac{1}{2}, 0 < c_1 < c_2, tol_{cg} > 0$
initialization: $g^{(0)} = \nabla p(x^{(0)}), d^{(0)} = -\nabla p(x^{(0)}),\ r = \|\nabla p(x^{(0)})\|_2, k = 0;$
while $r > tol_{cg}$ **do**

```
//Step size via Wolfe condition
```

choose $\alpha^{(k)}$ maximal such that

$$p(\gamma_{x^{(k)},d^{(k)}}(\alpha^{(k)})) - p(x^{(k)}) \le c_1 \alpha^{(k)} {d^{(k)}}^T \nabla p(x^{(k)}),$$
$${d^{(k)}}^T \nabla p(\gamma_{x^{(k)},d^{(k)}}(\alpha^{(k)})) \ge c_2 \nabla p(x^{(k)});$$

```
//Updating the argument and the conjugate direction
```

$$x^{(k+1)} = \gamma_{x^{(k)},d^{(k)}}(\alpha^{(k)});$$
$$g^{(k+1)} = \nabla p(x^{(k+1)});$$
$$\tilde{d}^{(k)} = \dot{\gamma}_{x^{(k)},d^{(k)}}(\alpha^{(k)});$$

```
//Step size for descent direction update
```

$$\beta^{(k)} = \begin{cases} \frac{{g^{(k+1)}}^T Hp(x^{(k+1)})\tilde{d}^{(k)}}{{\tilde{d}^{(k)}}^T Hp(x^{(k+1)})\tilde{d}^{(k)}}, & {\tilde{d}^{(k)}}^T Hp(x^{(k+1)})\tilde{d}^{(k)} \neq 0, \\ 0, & \text{else;} \end{cases}$$
$$d^{(k)} = -g^{(k)} + \beta^{(k)}\tilde{d}^{(k)};$$

$x^\star = x^{(k+1)};$
$r = \frac{1}{L}\sum_{i=1}^{L} |\nabla p(x_i^k)|;$
$k = k + 1;$

end
Output: $x^\star \in SO(3)$

We can now formulate our first reconstruction algorithm, Algorithm 1, and test it in several numerical experiments.

Experiment 1 (Noise-free recovery)**.** *In the first experiment, we provide some reconstruction examples of noiseless recovery. To illustrate the result of the semi-definite program, involved in Algorithm 1, we simulate a point measure*

$$\mu^\star = \sum_{i=1}^{6} c_i \delta_{x_i},$$

where $x_i = x_i(\alpha_i, \beta_i, \gamma_i)$ *are six randomly distributed support points and* c_i *are randomly generated amplitudes, given in table 4.1. The support points are chosen, such that the minimal separation obeys*

$$\min_{i \neq j} \omega(x_i^{-1} x_j) > 2.01.$$

From this test measure we generate moments up to degree $N = 2$ *and solve the semi-definite program* ($\mathrm{dRP_{rel}}$) *with* $M = N$*. To see the difference between the given low frequency information* $\mathcal{S}_2^* \mu^\star$ *and the solution* $f^\star$ *of the optimization* ($\mathrm{dRP_{rel}}$)*, we plot the squared absolute value of both functions on three slices along the second Euler angle. This is illustrated in figure 4.1.*

Algorithm 1: Super-resolution on $SO(3)$ via SDP

Input: low-frequency approximation $g \in \Pi_N$ of $\mu^\star$
Parameters: relaxation parameter $M \geq N$, regularization parameter τ, $P \in \mathbb{N}$, $tol > 0$
initialization: $\mathcal{X} = \emptyset$
do

Solve for $f^\star \in \Pi_N$, $Q_0 \in \mathbb{C}^{(2M+1)^3 \times (2M+1)^3}$, $Q_1 \in \mathbb{C}^{(2M)^3 \times (2M)^3}$

$$\max_{f \in \Pi_N, Q_0, Q_1} \operatorname{Re}\langle g, f\rangle - \tau \|f\|^2_{L^2(SO(3))}, \quad \text{subject to} \quad (4.15), \qquad (\text{dRP}_{\tau,\text{rel}})$$

using an interior point method;

Randomly choose P initial points $\{x_1, \ldots, x_P\} \subset SO(3)$;

for $l = 1, \ldots, P$ **do**
  Find $x_l^\star$ via Procedure 1 with initialization point x_l;
  if $(1 - |f^\star(x_l^\star)| < tol)$ **then**
    $\mathcal{X} = \mathcal{X} \cup \{x_l^\star\}$;
  end
end
Set $\nu = \sum_{x_i \in \mathcal{X}} c_i \delta_{x_i}$, with c_i such that

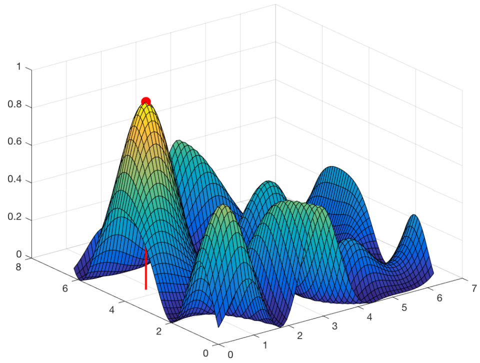
$$\nu = \operatorname{argmin} \|S_N^* \nu - g\|_2;$$

end;
Output: $\nu \in \mathcal{M}(SO(3))$

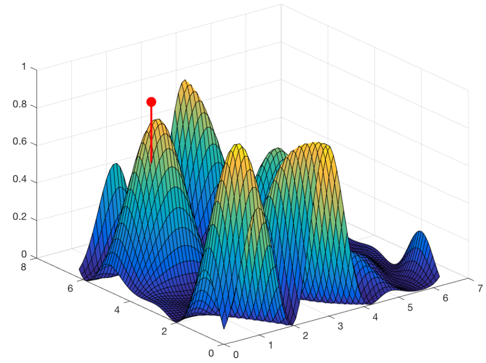
---

i	α_i	β_i	γ_i	c_i
1	0.8277	1.0964	4.3325	-0.6911
2	0.6007	1.5683	0.8292	0.7511
3	3.9386	1.7086	3.7938	0.7696
4	4.1735	2.2986	0.8480	0.3185
5	2.7859	0.4716	0.4551	-0.9460
6	1.6728	2.4994	3.0875	0.9161

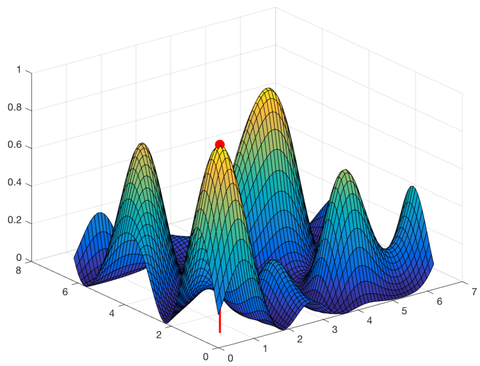
Table 4.1: The support points and amplitudes of the test measure $\mu^\star$.



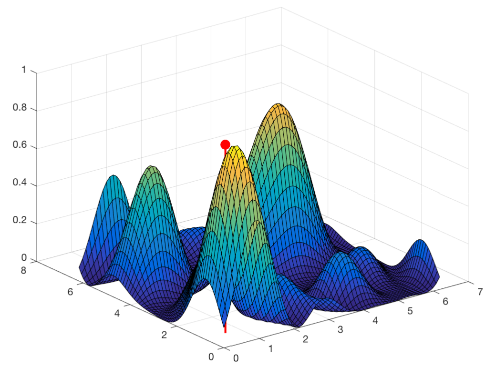
(a) $f^\star$ at slice β_1



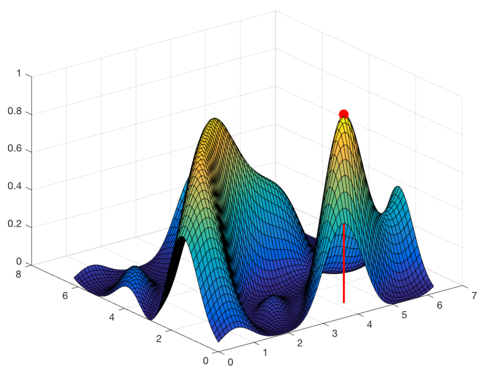
(b) $\mathcal{S}_2^*\mu^\star$ at slice β_1



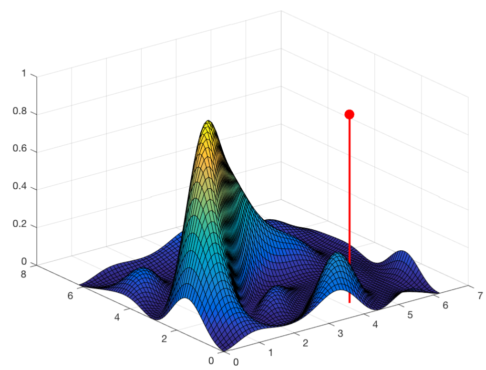
(c) $f^\star$ at slice β_2



(d) $\mathcal{S}_2^*\mu^\star$ at slice β_2



(e) $f^\star$ at slice β_4



(f) $\mathcal{S}_2^*\mu^\star$ at slice β_4

Figure 4.1: The squared absolute value of the solution $f^\star$ of (dRP) against the scaled squared absolute value of the given low frequency information $\mathcal{S}_2^*\mu^\star$ on different slices along the second Euler angle. The support of the measure $\mu^\star$ is indicated by the red bars.

One observes, that the solution $f^\star$ of the minimization acts as a dual certificate, i.e. it equals one in absolute value at the unknown support points of $\mu^\star$, whereas the low frequency approximation has several local maxima that cannot be distinguished from the unknown support. Even if one restricts to the slices in the second Euler angels of the unknown support points of $\mu^\star$, the biggest local extrema of $\mathcal{S}_2^*\mu^\star$ are located elsewhere and the true support point may even not be a local extremum of $\mathcal{S}_2^*\mu^\star$.

After computing the solution $f^\star$, we randomly choose $P = 20000$ initial points and apply the simultaneous conjugate gradient method described in Procedure 1 to the function $1-|f^\star|^2$. We set the tolerance to $tol = 10^{-8}$ and identify the resulting minima x_i^{rec}, which fulfill $1-|f^\star(x_i^{rec})|^2 < tol$, as support points. Finally, we solve for the amplitudes in

$$\mu^{rec} = \sum_i c_i^{rec}\delta_{x_i^{rec}}$$

via the least squares minimization

$$c^{rec} = \operatorname{argmin} \|S_2^*\mu^{rec} - S_2^*\mu^\star\|_2.$$

With this procedure, we obtain six points x_i^{rec}, such that the maximal distance to the points x_i fulfills

$$\max_i \min_j \omega(x_i^{-1}x_j^{rec}) < 8.6 \cdot 10^{-7},$$

and the correctly ordered amplitudes obey $\max_i |c_i - c_i^{rec}| < 7.7 \cdot 10^{-8}$. We plot the result in figure 4.2.

Moreover, we can apply Algorithm 1 to complexed valued measures. Some reconstruction examples for different number of unknown points and values of N are given in figure 4.3.

Experiment 2 (Influence of the separation)**.** *In this experiment, we give a glimpse on the influence of the separation distance of the support points. Remember, the support points $\{x_i\}$ are supposed to obey a separation condition of the form*

$$\omega(x_i^{-1}x_j) \geq \frac{\nu}{N}.$$

As seen in Chapter 3, $\nu = 36$ is a sufficient criteria. Nevertheless, the actual minimal *constant may be much smaller, which the following experiment indicates. For this, we partition the interval $[1, N\cdot\pi]$, i.e. we choose $1 = \nu_0 < \nu_1 < \cdots < \nu_n = N\cdot\pi$ and generate for each subinterval $[\nu_i, \nu_{i+1}]$ twenty sets of two points, that are separated by $\frac{\nu_i}{N}$ but not by $\frac{\nu_{i+1}}{N}$, and choose random complex amplitudes. Afterwards, we compute the moments for $N = 1$ and $N = 2$ and apply Algorithm 1 with $M = N$, $\tau = 0$ and $L = 20000$. We say that the recovery is successful, if*

$$\max_i \min_j \omega(x_i^{-1}x_j^{rec}) < 10^{-4},$$

where x_j^{rec} are the recovered points. To underpin the observations, we partition the interval $[2.5, 3.5]$ and generate for each subinterval ten sets of two points separated by the corresponding distance and apply Algorithm 1 to moments of order $N = 3$. One can observe, that in these cases for $\frac{\nu}{N}$, where ν is slightly larger than 3, we get exact recovery using Algorithm 1. The measured recovery rate is plotted in figure 4.4. Due to the high complexity of solving the semidefinite program ($\mathrm{dRP_{rel}}$)*, testing Algorithm 1 for higher order moments is sophisticated. To consolidate the transition to exact recovery at a constant slightly larger than 3, one has to test even on higher available moments, which demands for the use of customized solvers for the semidefinite program* ($\mathrm{dRP_{rel}}$) *to make the computation feasible, which goes beyond the scope of this thesis. We leave this for future research.*

Remark 4.6 (Relaxation parameter)**.** *In Experiment 2, we choose the minimal relaxation parameter, i.e. $M = N$. We find, that increasing M to $M = N+1$ does not change the recovery rate, at least for $N = 1$ and $N = 2$. The influences of the relaxation parameter on the recovery is closely related to the validity of the Assumption* (SOS)*. How to chose the parameter depending on the measure to recover is an open problem.*

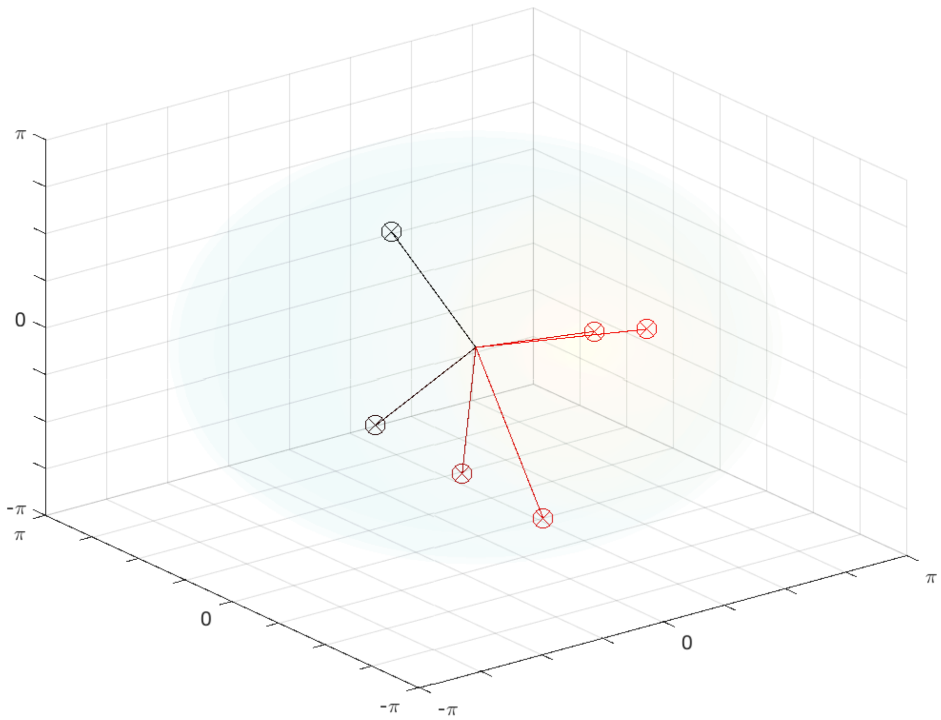
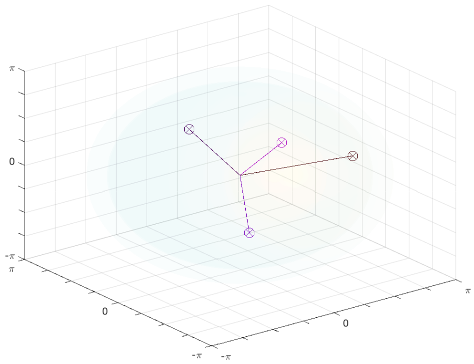
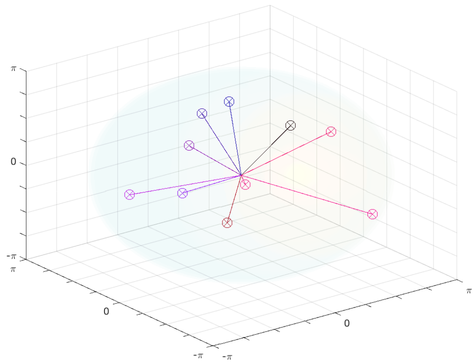


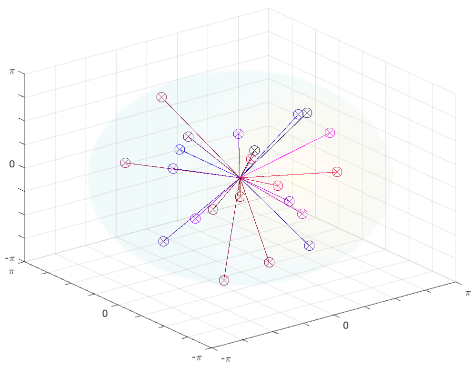
Figure 4.2: Reconstruction of the measure $\mu^\star$, defined in table 4.1. The representation is done in the axis-angle parametrization. The unknown points x_i are marked with circles, located on the top of a ray, whose direction correspond to the axis given as a point on the sphere, plotted in light grey, and its length to the angle. The unknown amplitudes are color coded in different shadows of red corresponding to the value of the amplitude. The reconstructions x_i^{rec} by algorithm 1 are marked by little small crosses and the reconstructed amplitudes c_i^{rec} are color coded as described before.



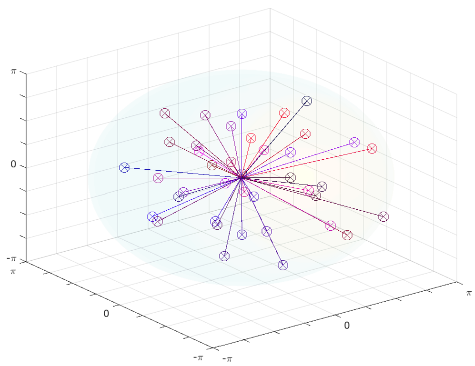
(a) 4 points with separation 3, $N = 1$



(b) 10 points with separation 1.9, $N = 2$

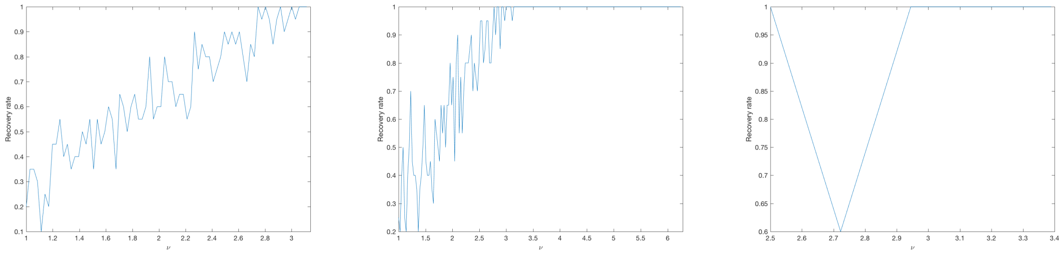


(c) 22 points with separation 1.3, $N = 3$



(d) 40 points with separation 1.0, $N = 4$

Figure 4.3: Reconstruction examples for different number of unknown points and values of N. The representation is done in the axis-angle parametrization. The unknown points are marked with circles, located on the top of a ray, whose direction correspond to the axis and its length to the angle. The unknown amplitudes are colour coded , the real part in red and the imaginary in blue. The reconstructions by algorithm 1 are marked by little small crosses. All examples show a nearly perfect recovery.



(a) Recovery rate for $N = 1$.

(b) Recovery rate for $N = 2$.

(c) Recovery rate for $N = 3$.

Figure 4.4: The recovery rate of Algorithm 1 for different values of N. On the x-axis, we plotted the constant appearing in the separation condition $\frac{\nu}{N}$, whereas on the y-axis we have the recovery rate.

Experiment 3 (Noisy data)**.** *In this experiment, we test the behavior of Algorithm 1 on noisy data. For this, we randomly generate sets of eight points on the rotation group, randomly generate corresponding amplitudes and compute moments of order $N = 2$. We normalize the amplitudes, such that we always have $\|\mathcal{S}_N^* \mu\|_2 = 1$. The generated point sets are separated by $\frac{4}{N}$, as Experiment 2 suggests that we can expect exact noiseless recovery in this case. We choose nine different values of τ ranging from 0.5 to 0.01 and randomly disturb the low frequency information with noise $\mathcal{S}_N^* \eta$, such that*

$$\|\mathcal{S}_N^* \eta\|_{L^2} \leq \tau.$$

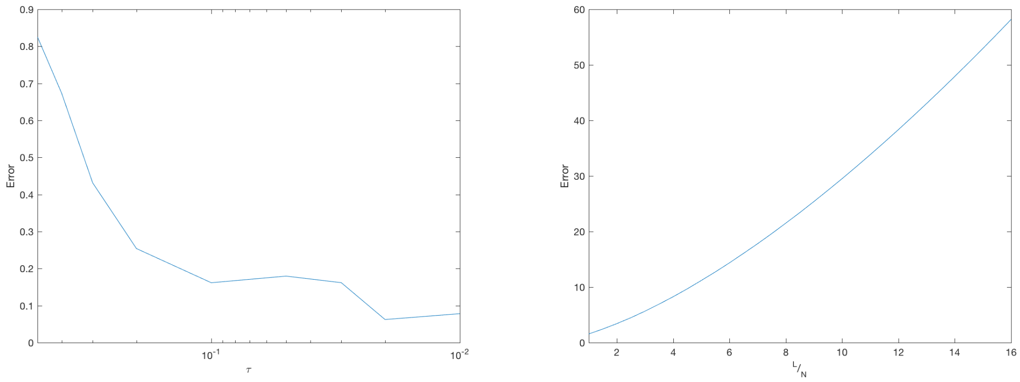
After generating the data, we apply Algorithm 1 for the different values of τ and different data sets with the parameters $M = N$ and $P = 30000$. Again, we first evaluate the recovery by measuring the maximal distance

$$\max_i \min_j \omega(x_i^{-1} x_j^{rec}), \tag{4.16}$$

where x_j^{rec} are the recovered points and x_j denote the true support points. We plot the recovery error averaged over the twenty sets in Figure 4.5(a).

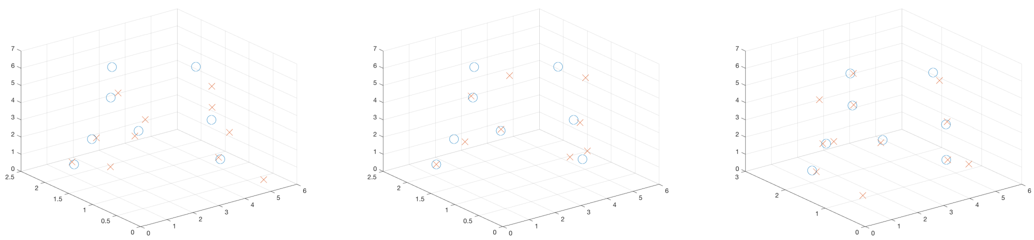
To measure the error induced by the super-resolution process, we plot the quantity $\|\mathcal{S}_L^*(\mu^\star - \mu_\tau)\|_2$ in Figure 4.5(b) for values of L ranging from 2 to 32. Here, $\mu^\star$ is the sought measure and μ_τ is the measure recovered from the noisy measurements. Observe, that the theoretical L^∞ bound of Section 3.3 also yields a corresponding L^2-estimate. The quadratic growth of the error with respect to the super-resolution factor can be observed in Figure 4.5(b).

Lastly, we plot the recovered measures for different values of the noiselevel τ for one test measure in Figure 4.6. One observes, that the noise on the measurements induces artificial support points close to the true support of the sought measure.

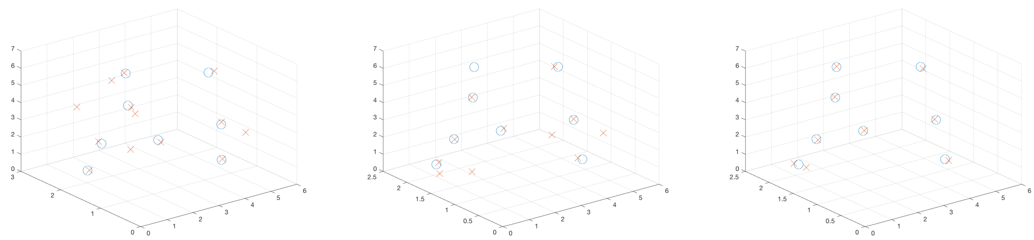


(a) Recovery error as defined in (4.16). (b) Super-resolution error $\|\mathcal{S}_L^*(\mu^\star - \mu_\tau)\|_2$.

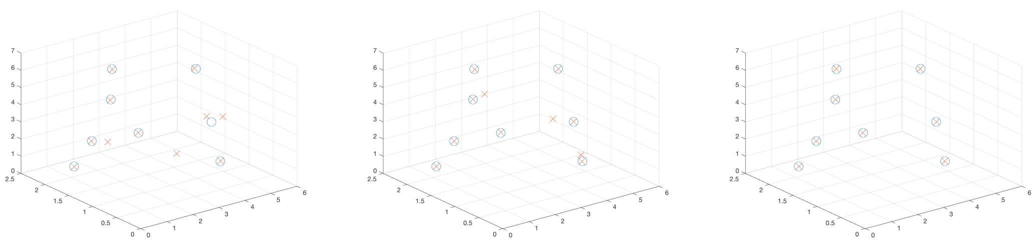
Figure 4.5: (a) The recovery error for moments of order $N = 2$ and different values of the noise level τ averaged over twenty test sets. (b) The induced super-resolution error for values of the super-resolution factor $\frac{L}{N}$ ranging from 1 to 16 for a noiselevel of $\tau = 0.1$ averaged over the test measures.



(a) $\tau = 0.5$ (b) $\tau = 0.4$ (c) $\tau = 0.3$



(d) $\tau = 0.2$ (e) $\tau = 0.1$ (f) $\tau = 0.05$



(g) $\tau = 0.03$ (h) $\tau = 0.02$ (i) $\tau = 0.01$

Figure 4.6: Recovered support points for different values of τ. The representation is done in Euler angles. The true support of the measure $\mu^\star$ are marked with blue circles and the recovered points with orange crosses. One observes, that although the recovered points approach the support of the sought measure as the noiselevel decreases, the noise on the measurements induces artificial support points close to the true support.

We would like to mention, that Algorithm 1 do not use any a priori discretization. Nevertheless, the complexity of the semi-definite program grows dramatically in N, the order of the given moments. More concrete, the inspection of Algorithm 1 shows that the number of variables of the semi-definite program grows like $\mathcal{O}(N^6)$. A very important question is, how to adapt Algorithm 1 to make it applicable for higher order moments. One possibility is the use of customized solvers for the semi-definite program, which we leave for future research.

Instead, we consider a different approach, that relies on an a priori discretization, the solution of a finite-dimensional problem and a clustering step in the next section.

4.2 Discretization of the Optimization Problem

Due to the high complexity of the semi-definite program for higher order moments, we propose to use a discretization of the primal problems (RP) and (RP$_\tau$). This means, we choose a sequence of discrete sets $\mathcal{G}_n \subset SO(3)$ and solve for the discretized problems

$$\min_{\operatorname{supp}(\mu)\subset\mathcal{G}_n} \|\mu\|_{TV}, \quad \text{subject to} \quad \mathcal{S}_N^*\mu = \mathcal{S}_N^*\mu^\star, \tag{RP$_n$}$$

respectively

$$\min_{\operatorname{supp}(\mu)\subset\mathcal{G}_n} \frac{1}{2}\|\mathcal{S}_N^*(\mu-\mu^\star-\eta)\|^2_{L^2(SO(3))} + \tau\|\mu\|_{TV}. \tag{RP$_{\tau,n}$}$$

In this section, we consider the convergence behavior of the solutions of the discretized problems. The convergence depends on the filling distance of the chosen discretization $\mathcal{G}_n$, given by

$$h(\mathcal{G}_n) = \sup_{x\in SO(3)} \inf_{y\in\mathcal{G}_n} \omega(x^{-1}y).$$

By doing this, we follow ideas from [Tang et al., 2013], where the convergence for continuously parametrized dictionaries has been examined.

Theorem 4.7. *Assume $\mu^\star = \sum_i c_i\delta_{x_i}$ is the unique solution of* (RP) *and the sequence $\mathcal{G}_n$ of discretizations is chosen, such that*

$$h(\mathcal{G}_n) \to 0.$$

Then each sequence of solutions μ_n of (RP$_n$) *converges to $\mu^\star$ and the solutions $\mu_{\tau,n}$ of* (RP$_{\tau,n}$) *converges to the unique solution μ_τ of* (RP$_\tau$) *in the weak*-topology.*

Moreover, there exist $\varepsilon > 0$, such that

$$\mu_n(B_\varepsilon(x_i)) \to c_i, \quad |\mu_n|(B_\varepsilon(x_i)) \to |c_i|,$$

and

$$|\mu_n|\left((\cup_i B_\varepsilon(x_i))^c\right) \to 0. \tag{4.17}$$

Proof. The proof is quite similar to the proof of [Tang et al., 2013, Thm. 2]. We show, that each sequence of solutions is bounded and thus, due to the sequentially Banach-Alaoglu Theorem, admits a weak* convergent subsequence, which converges to a solution of the continuous problems (RP) respectively (RP$_\tau$). The uniqueness of the solution to the continuous problems than guarantees the convergence of the whole sequence by an subsequence-subsequence argument.

For the boundedness, consider the convex dual problems

$$\max_{f\in\Pi_N} \operatorname{Re}(\langle f, \mathcal{S}_N^*\mu^\star\rangle), \quad \text{s.t.} \quad |f(x)| \le 1, x \in \mathcal{G}_n, \tag{dRP$_n$}$$

respectively

$$\max_{f\in\Pi_N} \mathrm{Re}\langle \mathcal{S}_N^*(\mu^\star+\eta), f\rangle - \tau\|f\|^2_{L^2(SO(3))}, \quad \text{s.t.} \quad |f(x)|\le 1, x\in\mathcal{G}_n. \tag{dRP$_{\tau,n}$}$$

For abbreviation, we write

$$\begin{aligned} Q(f) &= \mathrm{Re}(\langle f, \mathcal{S}_N^*\mu^\star\rangle), \\ Q_\tau(f) &= \mathrm{Re}\langle \mathcal{S}_N^*(\mu^\star+\eta), f\rangle - \tau\|f\|^2_{L^2(SO(3))}. \end{aligned}$$

We show, that the feasible sets of the dual problems, i.e. the sets of $f\in\Pi_N$ such that $|f(x)|\le 1$ for $x\in\mathcal{G}_n$, are bounded in Π_N and therefore compact. Since $h(\mathcal{G}_n)\to 0$, we have that for large enough n,

$$h(\mathcal{G}_n) < \frac{1}{N}.$$

Applying a Marcinkiewicz-Zygmund inequality, see [Schmid, 2008, Thm. 4.4], yields for all $f\in\Pi_N$

$$\|f\|_\infty \le (1-Nh(\mathcal{G}_n))^{-1}\max_{x\in\mathcal{G}_n}|f(x)|,$$

meaning all feasible sets are uniformly bounded and thus compact. This shows, that each discretized problem has a solution. In the case of the problem (dRP$_{\tau,n}$), this solution is unique, due to the strict convexity of the objective function Q_τ. We denote these minimizer sequences by f_n respectively $f_{\tau,n}$.

The rest of the proof is identical to the proof of [Tang et al., 2013, Thm. 2]. We will sketch it briefly. Due to the uniform boundedness, one can show, that each sequence f_n converges to a solution $f^\star$ of the continuous dual problem

$$\max_{f\in\Pi_N} Q(f), \quad \text{subject to} \quad \|f\|_\infty\le 1, \tag{dRP}$$

and the sequence $f_{\tau,n}$ converges to the unique solution f_τ of

$$\max_{f\in\Pi_N} Q_\tau(f), \quad \text{subject to} \quad \|f\|_\infty\le 1. \tag{dRP$_\tau$}$$

Since strong duality holds for the discretized problems as well as for the continuous ones, one knows that

$$\begin{aligned} Q(f_n) &= \|\mu_n\|_{TV}, \\ Q_\tau(f_{\tau,n}) &= \frac{1}{2}\|\mathcal{S}_N^*(\mu_{\tau,n}-\mu^\star-\eta)\|^2_{L^2(SO(3))} + \tau\|\mu_{\tau,n}\|_{TV} \end{aligned}$$

and

$$\begin{aligned} Q(f^\star) &= \|\mu^\star\|_{TV}, \\ Q_\tau(f_\tau) &= \frac{1}{2}\|\mathcal{S}_N^*(\mu_\tau-\mu^\star-\eta)\|^2_{L^2(SO(3))} + \tau\|\mu_\tau\|_{TV}. \end{aligned}$$

This shows, that the sequences μ_n and $\mu_{\tau,n}$ are bounded. Again, by the Banach-Alaoglu Theorem, one gets the weak* convergence to the minimizers $\mu^\star$ respectively μ_τ.

The convergence of the measure of the epsilon balls follows identically as in [Tang et al., 2013, Cor. 1]. □

As also mentioned in [Tang et al., 2013], heuristically the property (4.17) suggests that for fine enough discretization the support of the solution of the minimization problem clusters around the true support of the sought measure. In the following, we will consider this behavior numerically.

First, we have to chose an appropriate discretization. We consider the following grid, defined in the Euler angle parametrization. For $n \in \mathbb{N}$, set

$$\alpha_{k_1} = \frac{2\pi k_1}{2n}, \quad \beta_{k_2} = \frac{\pi(2k_3+1)}{4n}, \quad \gamma_{k_3} = \frac{2\pi k_3}{2n}, \tag{4.18}$$

for $0 \leq k_1, k_2, k_3 \leq 2n-1$, generating a grid $\mathcal{G}_n$ of $8n^3$ sampling points. As shown in [Kostelec and Rockmore, 2008], with this choice of sampling one can compute the coefficients of any $f \in \Pi_{n-1}$ exactly. More concrete, there are quadrature weights w_{k_2}, $k_2 = 0, \dots, 2n-1$ such that

$$\langle f, D^l_{k,m} \rangle = \frac{1}{(2n)^2} \sum_{k_1,k_2,k_3=0}^{2n-1} w_{k_2} f(x(\alpha_{k_1}, \beta_{k_2}, \gamma_{k_3})) \overline{D^l_{k,m}(x(\alpha_{k_1}, \beta_{k_2}, \gamma_{k_3}))},$$

for all $l \leq n-1$, i.e. the quadrature is exact on Π_{n-1}, see [Kostelec and Rockmore, 2008, Thm. 2.1]. More important to us, using the equally spaced Euler angle grid makes it possible to use fast Fourier summation to calculate the matrix vector products with the matrix $\mathbf{D}_N \in \mathbb{C}^{(2n)^3 \times \dim(\Pi_N)}$, whose entries are given by the evaluation of the Wigner D-functions up to degree N on the sampling grid $\mathcal{G}_n$ given by (4.18), i.e.

$$\mathbf{D}_N = (D^l_{k,m}(x(\alpha_{k_1}, \beta_{k_2}, \gamma_{k_3})))^{k_1,k_2,k_3}_{l,k,m}, \tag{4.19}$$

with $l \leq N \leq n-1$. With this matrix notation, the problem (RP$_n$) reads as

$$\min_{c \in \mathbb{C}^{8n^3}} \|c\|_1, \quad \text{subject to} \quad \mathbf{D}^*_N c = g, \tag{4.20}$$

with $g = (\langle \mu^\star, D^l_{k,m} \rangle)_{l,k,m}$, i.e. it is a finite-dimensional basis pursuit problem involving the matrix $\mathbf{D}^*_N$. In the same way, the Thikonov-type problem (RP$_{\tau,n}$) can be written as

$$\min_{c \in \mathbb{C}^{8n^3}} \frac{1}{2} \|\mathbf{D}^*_N c - g^\eta\|_2^2 + \tau \|c\|_1, \tag{4.21}$$

with $g^\eta = (\langle \mu^\star + \eta, D^l_{k,m} \rangle)_{l,k,m}$.

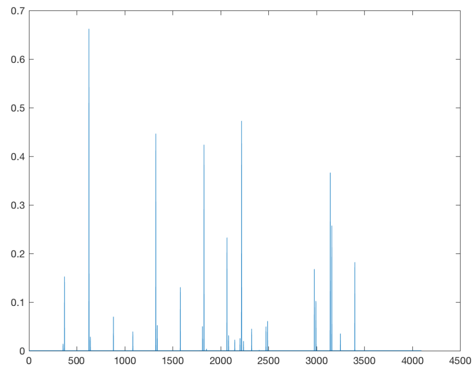
Experiment 4 (Noise-free recovery on finite grids)**.** *In this experiment, we solve the finite-dimensional basis pursuit problem* (4.20)*. For this, consider the measure $\mu^\star$ defined in Experiment 1, given by table 4.1. We compute moments with respect to the Wigner D-functions up to order $N = 2$ and solve the problem on the grids $\mathcal{G}_n$, for $n = 8, 12, 16, 20$ using again CVX calling the open source solver SeDuMi at the highest precision. We plot the absolute value of the solution vector in figure 4.7.*

Since the support points of $\mu^\star$ are not contained in $\mathcal{G}_n$, we cannot hope to exactly recover $\mu^\star$. Even worse, the recovered solution has much more spikes, than the measure $\mu^\star$, actually in this examples, each entry of the solution vector is non-zero. This phenomenon is known as basis mismatch*, see e.g. [Chi et al., 2011].*

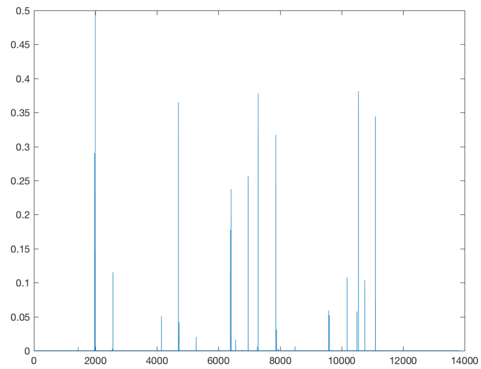
Nevertheless, only few entries are large in absolute value and it makes sense to keep only those, which are above a certain threshold in absolute value. More concrete, if we denote the solution of the minimization by $c^\star \in \mathbb{C}^{(2n)^3}$, we keep only those recovered points in $\mathcal{G}_n$, such that the corresponding coefficient obeys

$$|c^\star_i| > thres,$$

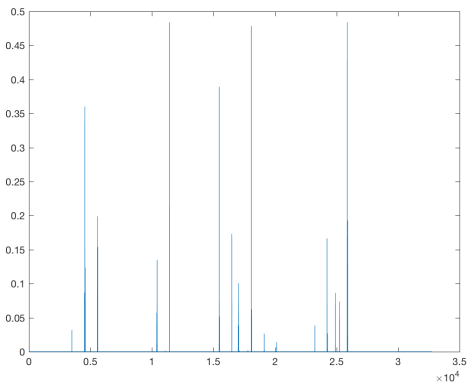
for some $thres > 0$. The result of this process for different values of $thres$ are plotted in figure 4.8. One observes, that the recovered grid points cluster around the true support of the the measure $\mu^\star$. Moreover, the effect of the basis mismatch gets smaller for increasing grid size, such that the thresholding becomes more accurate for denser grids, which is in line with the theoretical convergence statement (4.17)*.*



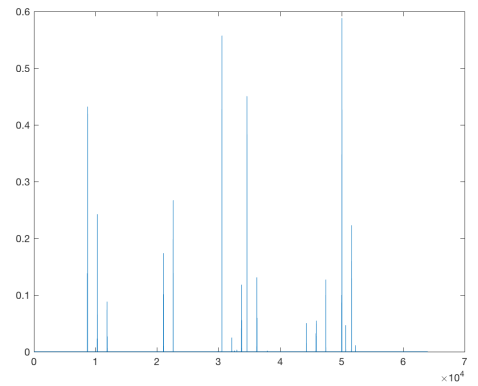
(a) $n = 8$



(b) $n = 12$

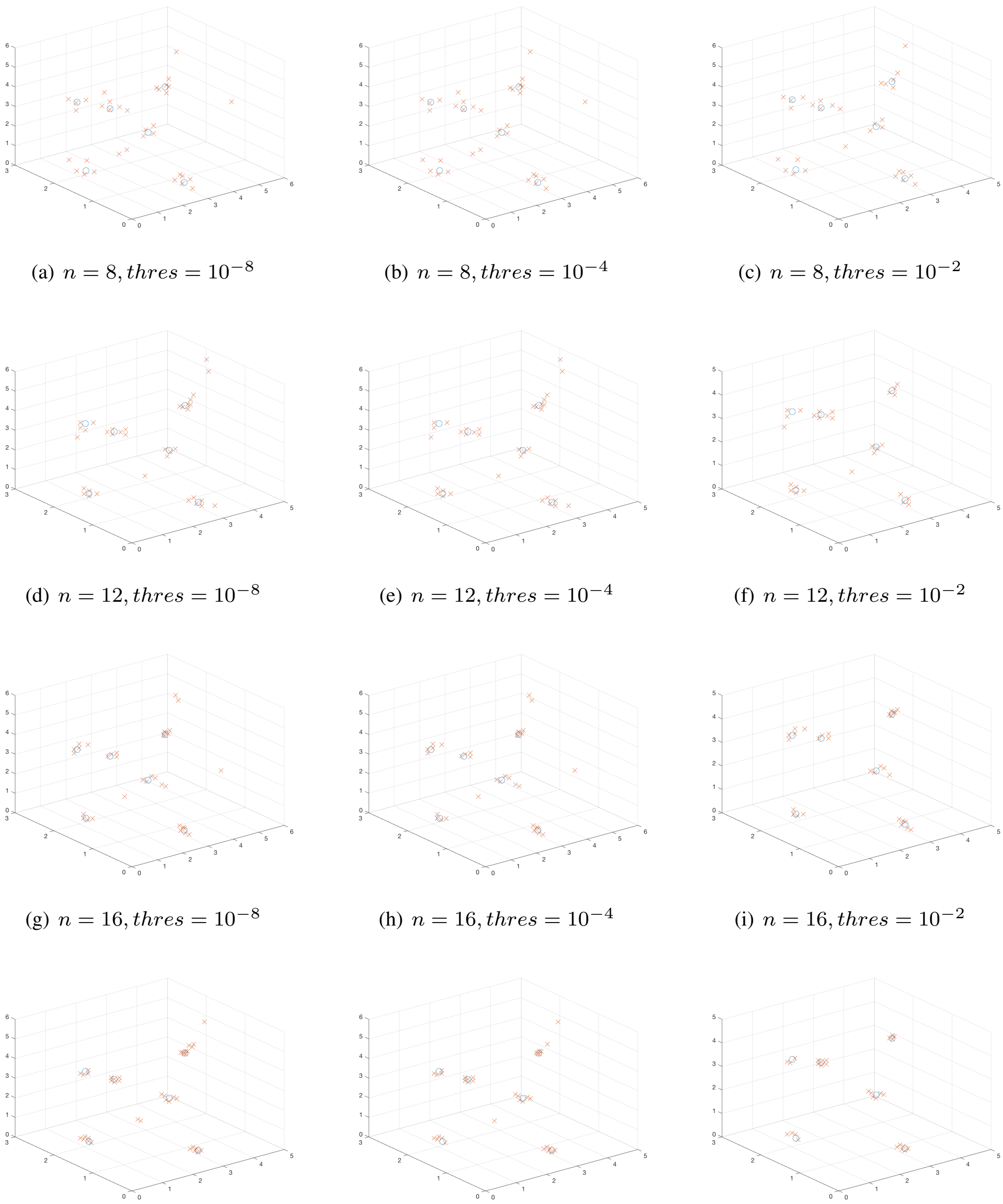


(c) $n = 16$



(d) $n = 20$

Figure 4.7: Absolute value of the $8n^3$ recovered coefficients of the solution of the minimization problem (4.20) for different grids $\mathcal{G}_n$. One observes, that an increase of the grid size yields a better concentration on few coefficients, whose corresponding grid points are close to a true support point of the sought measure.



(a) $n = 8, thres = 10^{-8}$ (b) $n = 8, thres = 10^{-4}$ (c) $n = 8, thres = 10^{-2}$

(d) $n = 12, thres = 10^{-8}$ (e) $n = 12, thres = 10^{-4}$ (f) $n = 12, thres = 10^{-2}$

(g) $n = 16, thres = 10^{-8}$ (h) $n = 16, thres = 10^{-4}$ (i) $n = 16, thres = 10^{-2}$

(j) $n = 20, thres = 10^{-8}$ (k) $n = 20, thres = 10^{-4}$ (l) $n = 20, thres = 10^{-2}$

Figure 4.8: Recovered support points on the grid $\mathcal{G}_n$ for different values of n and values of the threshold $thres$. The representation is done in Euler angles. The true support of the measure $\mu^\star$ are marked with blue circles and the recovered grid points with orange crosses.

The previous experiment suggests, that we can recover clusters around the true support of the sought measure via the discrete recovery process in combination with a thresholding step.

To find the clusters, we use an algorithm known as *mean-shift clustering*, see e.g. [Tuzel et al., 2005], [Comaniciu and Meer, 2002]. We will describe it briefly in the following. We assume, that the recovered points are samples of an unknown distribution, which is localized around the true support points of the sought measure. A popular way to approximate this distribution, is to approximate it using a *kernel-density estimator*, see e.g. [Pelletier, 2005], [Hielscher, 2013] and references therein. More concrete, if we denote the recovered grid points by $\{x_i^\star\}_{i=1}^{n^\star} \subset \mathcal{G}_n$, we set

$$f_\psi(x) = \frac{1}{n^\star} \sum_{i=1}^{n^\star} \psi(\omega(x^{-1}x_i^\star)),$$

where $\psi : \mathbb{R} \to \mathbb{R}_+$ is a suitable profile function. When the profile function is suitably localized, the clustering of the recovered points leads to local maximizer of f_ψ near the the unknown true support points. For the profile function, we will use the function

$$\psi_\varepsilon(t) = \begin{cases} \frac{15}{16}\left(1 - \left(\frac{t}{\varepsilon}\right)^2\right)^2, & |t| \leq \varepsilon, \\ 0, & \text{else}, \end{cases}$$

which is known as *quartic kernel* scaled with parameter $\varepsilon > 0$. Calculating the gradient of f_{ψ_ε} at x yields for $\varepsilon < \sqrt{2}\log(2)$,

$$\begin{aligned} \nabla f_{\psi_\varepsilon}(x) &= \frac{1}{n^\star} \sum_{i=1}^{n^\star} x \sum_{j=1}^{3} e_j(x^{-1}x_i^\star)\mathcal{L}_j \psi_\varepsilon'(\omega(x^{-1}x_i^\star)), \\ &= x\frac{1}{n^\star} \sum_{i=1}^{n^\star} \log(x^{-1}x_i^\star)\psi_\varepsilon'(\omega(x^{-1}x_i^\star)) \in T_x SO(3). \end{aligned}$$

Observe, that this computation is valid due to the compact support of the kernel ψ_ε. The mean shift of x takes a weighted gradient ascent step in the tangent space $T_x SO(3)$ given by

$$m_{\psi_\varepsilon}(x) = x \frac{\sum_{i=1}^{n^\star} \log(x^{-1}x_i^\star)\psi_\varepsilon'(\omega(x^{-1}x_i^\star))}{\sum_{i=1}^{n^\star} \psi_\varepsilon'(\omega(x^{-1}x_i^\star))}$$

and projecting back with the exponential map

$$x \exp_x(m_{\psi_\varepsilon}(x)).$$

We start this procedure at each point, which appears in the solution of the minimization (4.20) and whose corresponding amplitude exceeds the threshold $thres$ in absolute value and is iterated until convergence. For convergence results of this procedure, see e.g. [Comaniciu and Meer, 2002]. For better readability, we summarize the mean-shift procedure in Procedure 2.

After introducing the mean-shift clustering, we formulate our second recovery algorithm based on the finite-dimensional optimization problem (RP$_n$), respectively (RP$_{\tau,n}$). We state it in Algorithm 2. To test the algorithm, we generate randomly twenty points with random complex amplitudes and compute moments of order $N = 4$. The points are generated in a way, such that they are separated by $\frac{4}{N}$. We generate a grid with parameter $n = 20$, see (4.18). As the support points are in general located off the grid, similar to the case of noisy data, we cannot hope to have equality in the side condition of the problem (4.20). Instead, we solve the regularized problem (4.21) with regularization parameter $\tau = 1$. We set the

Procedure 2: Local mean-shift iteration

Input: Point set $\{x_i\}_{i=1}^n \subset SO(3)$.
Parameters: Scaling parameter ε and $0 < tol < 1$.
initialization: $res = 1$;
while $res > tol$ **do**
  for *i=1:n* **do**
    Compute the mean-shift

$$m_\psi(x_i) = \frac{\left(\sum_{x_j} \log(x_i^{-1}x_j)\psi'_\varepsilon\left(\omega(x_i^{-1}x_j)\right)\right)}{\left(\sum_{x_j} \psi'_\varepsilon\left(\omega(x_i^{-1}x_j)\right)\right)};$$

  end
  Update

$$x_i = x_i \exp(m_\psi(x_i)), \quad i = 1, \dots, n;$$

  Set

$$res = \max_i \|m_\psi(x_i)\|_2;$$

end

$$x_i^{\text{mean}} = x_i, \quad i = 1, \dots, n;$$

Output: Set of mean rotations $\{x_i^{\text{mean}}\}_{i=1}^n \subset SO(3)$.

threshold parameter to $thres = 10^{-8}$, the scaling parameter of the kernel to $\varepsilon = 8$ and the tolerance for terminating the mean-shift iteration to $tol = 10^{-10}$. The result of Algorithm 2 are plotted in Figure 4.9. The distance of the true support to the recovered points obeys

$$\max_i \min_j \omega(x_i^{-1}x_j^{\text{rec}}) < 0.078.$$

Remark 4.8 (Influence of the parameters)**.** *We would like to informally discuss the influence of the parameters appearing in Algorithm 2. The most obvious parameter is the grid parameter, since we can only recover grid points in the minimization process. The higher the grid parameter, the closer is the grid to the unknown support points, which becomes important if two support points become close. Since we demand a minimal separation condition for the unknown support points of the form $\frac{c}{N}$, where $N \in \mathbb{N}$ is the order of the given moments and $c > 0$ is a relatively small constant, the grid parameter should be coupled to N. If n is to small in comparison to the order N of given moments, then two unknown points close to each other may not be distinguished by the algorithm. On the other hand, if n is too big, then the matrix $\mathbf{D}_N$, given in* (4.19)*, is 'nearly singular', which may degrade the recovery process at all. However, a good rule for choosing the pre-defined grid depending on the given order of moments is missing.*

Closely related to the 'nearly singularity' of the matrix $\mathbf{D}_N$ is the choice of the regularization parameter τ. For a parameter τ, which is too small, the minimizer of the minimization (4.21) *has weights on nearly all grid points and the clusters that should indicate the unknown support points are not identifiable any more. If τ is chosen too big, it may happen that some support points are completely missed, i.e. no surrounding grid point is assigned a weight by the minimization.*

The thresholding parameter $thres$ should avoid, that all grid points are considered, since typically all points are assigned a weight, but away from a support point, these are usually small as also seen in Experiment 4. We found that a small value like $thres = 10^{-8}$ is sufficient.

Algorithm 2: Recovery on a pre-defined grid

Input: low-frequency approximation $g \in \Pi_N$ of $\mu^\star$
Parameters: grid-parameter $n \in \mathbb{N}$, regularization parameter τ, threshold $thres > 0$, scaling parameter ε, $tol > 0$
initialization: $\mathcal{X} = \emptyset$
do

Solve for $c^\star \in \mathbb{C}^{8n^3}$,

$$\min_{c \in \mathbb{C}^{8n^3}} \frac{1}{2}\|\mathbf{D}_N^* c - g^\eta\|_2^2 + \tau\|c\|_1,$$

where $\mathbf{D}_N$ is given in (4.19), using an interior point method;

Choose those $x_i \in \mathcal{G}_n$, such that

$$|c_i| > thres,$$

and set $\mathcal{X} = \{x_i\}_i$;

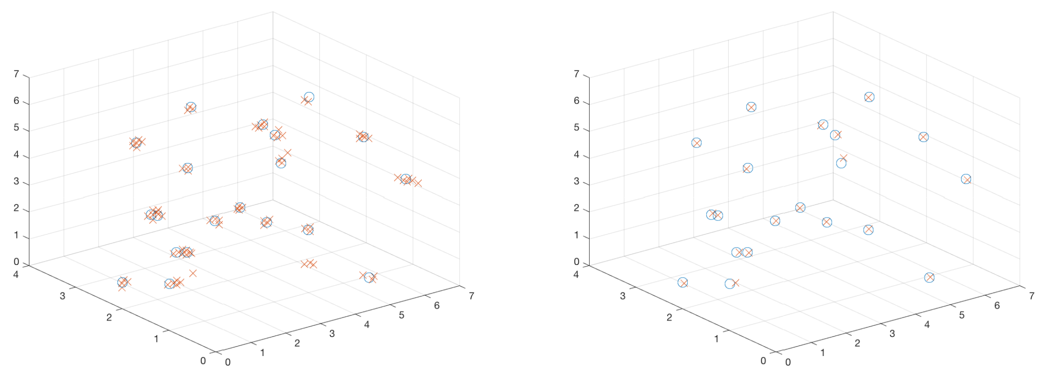
Apply Procedure 2 with parameters ε and tol to the set $\mathcal{X} = \{x_i\}_i$ and generate

$$\mathcal{X} = \{x_i^{\text{mean}}\}_i;$$

Set $\nu = \sum_{x_i \in \mathcal{X}} c_i \delta_{x_i}$, with c_i such that

$$\nu = \operatorname{argmin} \|S_N^* \nu - g\|_2;$$

end;
Output: $\nu \in \mathcal{M}(SO(3))$



(a) Result of minimization with regularization parameter $\tau = 1$ and thresholding with $thres = 10^{-8}$.

(b) Application of mean-shift clustering with scaling parameter $\varepsilon = 8$ and tolerance $tol = 10^{-10}$.

Figure 4.9: Application of Algorithm 2 to recover a measure $\mu^\star$ with 20 support points from moments of order $N = 4$. The true support of the measure $\mu^\star$ are marked with blue circles and the recovered grid points with orange crosses. The left picture shows the result of the minimization and thresholding step, whereas the right picture shows the result of the mean-shift iteration.

Regarding the scaling parameter ε of the mean-shift clustering, we would like to mention, that a parameter chosen to high results in generating too many clusters, i.e. two grid points actually associated to one unknown support points are assigned to two different clusters producing an artificial support point. In the case the scaling parameter is chosen to small, two different clusters belonging to two different support points may be merged resulting in completely missed support points. We observe, that choosing ε proportional to the order N of given moments with a small proportional constant works well.

We found, that the choice of the tolerance tol for terminating the mean-shift iteration is rather uncritical. The value $tol = 10^{-10}$ is typically achieved within three to ten iterations.

Summarizing, the dominant problem is the choice of the grid parameter $n \in \mathbb{N}$ in combination with the regularization parameter $\tau > 0$.

Notes and References. *Relaxing non-negativity constraints for trigonometric polynomials with sum of squares representations is a wide used tool in signal processing, see [Dumitrescu, 2006], [Dumitrescu, 2007]. This relaxation leads to finite-dimensional semi-definite programs, which can be solved using standard convex optimization algorithms. Nevertheless, the number of variables of these programs grows exponentially in the degree of the involved polynomial. More concrete, in our case the number of variables grows like $\mathcal{O}(N^6)$, when N is the order of given moments. Recently, several authors proposed to use customized interior point algorithms to solve semi-definite programs stemming from sum of squares relaxations of non-negativity constraints of trigonometric polynomials, see [Roh and Vandenberghe, 2006] and [Roh et al., 2007]. An interesting question for future research is the adaption of those proposed methods to solve the optimization problems* $(\mathrm{dRP}_{\mathrm{rel}})$ *and* $(\mathrm{dRP}_{\tau,\mathrm{rel}})$ *appearing in Algorithm 1.*

A different line of future research is due to the a priori discretized optimization problems (RP_n) *resp.* $(\mathrm{RP}_{\tau,n})$ *involved in Algorithm 2. Since the number of variables is $8n^3$, where $n \in \mathbb{N}$ is the grid parameter, it may be valuable to use specialized interior point methods or first order methods in combination with fast algorithm for the matrix vector product with the matrix $\mathbf{D}_N$ resp. $\mathbf{D}_N^*$, see [Kostelec and Rockmore, 2008]. One could also consider a local refinement of the grid near the points recovered after the mean-shift clustering and restart the minimization process involving the new discretization. We leave these considerations for future research.*

Chapter 5

Super-Resolution on the Sphere

In this chapter, we show the applicability of the approach of the previous chapters, i.e. using Hermite interpolation to construct a dual certificate, to the super-resolution problem on the two dimensional Euclidean sphere.

Beforehand, we discuss the limitation of extending this approach to more general settings. Indeed, one could consider the recovery of point measures on a smooth compact Riemannian manifold $\mathcal{M}$ from low order moments with respect to the eigenfunctions $\{\varphi_k\}_k$ of the Laplace-Beltrami operator on the manifold, which includes the described setting of the rotation group and the sphere.

In order to construct a dual certificate, one can formulate the Hermite interpolation problem again. As seen in Chapter 2 and 3, it was crucial to control quantities of the form

$$X_i^x X_j^y \Phi_N(x,y), \quad x,y \in \mathcal{M},$$

where X_i^x, X_j^y are vector fields and $\Phi_N : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ is an interpolation kernel of the form

$$\Phi_N(x,y) = \sum_{k=0}^{N} a_k \varphi_k(x) \overline{\varphi_k(y)}.$$

Although, there are asymptotic results of the form

$$|X_i^x X_j^y \Phi_N(x,y)| \leq \frac{C}{N^s d(x,y)^s}, \tag{5.1}$$

where $d(x,y)$ is the geodesic distance, see [Filbir and Mhaskar, 2010], the constant $C > 0$ is in general not accessible. In addition, the behavior near the diagonal is in general not clear. The consequence is, that, although one can show the invertibility of the interpolation matrix for a minimal separation distance *big enough*, it is not possible to show the bound

$$|q(x)| < 1,$$

where q is the solution of the interpolation problem and x is near an interpolation point, using only asymptotic estimates of the form (5.1) with unknown constants.

Thus, the approach using Hermite interpolation to construct a dual certificate seems limited to those cases, where one can construct interpolation kernels obeying asymptotic localization estimates of the form (5.1) with explicit constants and can be controlled well near the diagonal, by using e.g. Taylor expansions.

In this chapter, we consider the recovery of point measures on the two dimensional sphere from low order moments with respect to spherical harmonics. This setting has been considered at first in [Bendory et al., 2015a]. Nevertheless, there are severe gaps in the proofs due to the choice of interpolation kernel, that do not allow for good control of its pointwise behavior. We aim to close these gaps in the following chapter.

At first, we start by describing the setting of super-resolution on the sphere in Section 5.1, including the differentiable structure on the sphere, the involved basis function, called *Spherical harmonics*, as well as the recovery problem from moments with respect to these basis functions. Afterwards, in Section 5.2, we derive the needed localization estimates for a specific interpolation kernel. Finally, we construct the dual certificate in Section 5.3. In short, the structure of this chapter is based on the structure of the previous chapters, dedicated to the recovery on the rotation group.

5.1 Analysis on the Sphere

In this section, we briefly summarize the analytical tools on the sphere, which are needed in the following, and formulate the super-resolution problem on the sphere.

The two-dimensional sphere is given by

$$\mathbb{S}^2 = \{x \in \mathbb{R}^3 : \|x\|_2 = 1\}.$$

It is a two-dimensional smooth manifold and a metric, which is compatible with its topology, is the great-circle distance

$$d(x, y) = \arccos(\langle x, y\rangle), \quad x, y \in \mathbb{S}^2.$$

The tangent space at a point $x \in \mathbb{S}^2$ is

$$T_x\mathbb{S}^2 = \{y \in \mathbb{R}^3 : \langle x, y\rangle = 0\}.$$

In the following we shortly describe the differential structure on the sphere. We will use two different local coordinates on $\mathbb{S}^2$. First, for $v \in T_x\mathbb{S}^2$, there is a unique geodesic $\gamma_{x,v}$ such that $\gamma_{x,v}(0) = x$ and $\gamma'_{x,v}(0) = v$. It has the form

$$\gamma_{x,v}(t) = \cos(\|v\|_2 t)x + \sin(\|v\|_2 t)\frac{v}{\|v\|_2}.$$

The exponential map at $s \in \mathbb{S}^2$ is

$$\exp_x(v) = \gamma_{x,v}(1).$$

We fix an orthonormal basis η_1^x, η_2^x in $T_x\mathbb{S}^2$, such that $\eta_2^x = \eta_1^x \times x$ and thus $\eta_1^x = x \times \eta_2^x$. This is always possible, although we cannot choose the local bases in a continuous way, as there is no continuous nowhere vanishing vector field on the sphere, due to the *Hairy ball theorem.* This is in contrast to the case of the rotation group, where the tangent space is basically a translation of the tangent space at the identity. One way to obtain η_i^x is to chose a $z \in \mathbb{S}^2$ and an orthonormal basis η_1^z, η_2^z of $T_z\mathbb{S}^2$ and set

$$\eta_i^x = \begin{cases} \mathrm{e}^{d(x,z)\cdot\left[\frac{z\times x}{\sin(d,z)}\right]}\eta_i^z, & x \neq -z, \\ -\eta_i^z, & x = -z, \end{cases}$$

where for a vector $v \in \mathbb{R}^3$

$$[v] = \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix}$$

is the corresponding skew-symmetric matrix in the algebra $\mathfrak{so}(3)$. In other words, we rotate the local coordinate system at z, which is continuous for all points but the antipodal point $-z$.

Combining the coordinates due to this basis with the exponential map yields the *normal coordinates* centered at $x \in \mathbb{S}^2$, i.e. we parametrize a neighborhood of x by

$$\varphi(v_1, v_2) = \cos(\|v\|_2)x + \sin(\|v\|_2)\frac{v_1\eta_1^x + v_2\eta_x^2}{\|v\|_2},$$

and the inverse parametrization for z in a neighborhood of x is given by

$$v_i(z) = \frac{d(z,x)}{\sin(d(x,z))}\langle z, \eta_i^x\rangle, \quad i = 1, 2.$$

The vectors

$$\frac{\partial}{\partial v_1}\varphi(v_1(z), v_2(z)), \quad \frac{\partial}{\partial v_2}\varphi(v_1(z), v_2(z))$$

form an basis of $T_z\mathbb{S}^2$. One can schow, that in the center of the normal coordinates

$$\frac{\partial^2}{\partial v_i^2}\varphi(v_1(x), v_2(x)) = -x, \quad \frac{\partial^2}{\partial v_i v_j}\varphi(v_1(x), v_2(x)) = 0.$$

In this local coordinates centered at $x \in \mathbb{S}^2$, the gradient of a differentiable function f at x has the representation

$$\nabla f(x) = \begin{pmatrix} X_1 f(x) \\ X_2 f(x) \end{pmatrix},$$

with

$$X_i f(x) = \frac{\partial}{\partial v_i}(f \circ \varphi)(v_1(x), v_2(x)) = \lim_{t\to 0} t^{-1}(f(\gamma_{x,\eta_i^x}(t)) - f(x)). \tag{5.2}$$

In the center x of the normal coordinates, the Hessian of a two times differentiable function f has the representation

$$\begin{pmatrix} X_1X_1f(x) & X_1X_2f(x) \\ X_2X_1f(x) & X_2X_2f(x) \end{pmatrix}.$$

This is only true in the center of the normal coordinates, since the Christoffel symbols vanish. For different points, we would have to compute the Christoffel symbols with respect to the normal coordinates, which becomes quite complicated.

Alternatively, we introduce a second set of coordinates, such that the computation of the Christoffel symbols is much simpler. For a point $z \in \mathbb{S}^2$, we parametrize the set $B_\pi(0) \setminus \{0\}$ by

$$v(r, \theta) = r(\cos(\theta)\eta_1^z + \sin(\theta)\eta_2^z)$$

for $(r, \theta) \in (0, \pi) \times [0, 2\pi)$. Combining this with the exponential map, i.e.

$$\varphi^{\text{pol}}(r, \theta) = \exp_z(v(r, \theta)),$$

yields the *polar coordinates* centered at $z \in \mathbb{S}^2$, which parametrize $\mathbb{S}^2 \setminus \{z, -z\}$. For $z = (0, 0, 1)^T$, these are the usual spherical coordinates on the sphere, given by

$$\begin{pmatrix} \cos(r)\sin(\theta) \\ \sin(r)\sin(\theta) \\ \cos(r) \end{pmatrix}.$$

For $\mathbb{S}^2 \setminus \{z, -z\}$, the inverse parametrization, is given by

$$r(x) = \arccos(\langle x, z\rangle) = d(x, z),$$
$$\theta(x) = \arctan_2 \left(\langle x, \eta_2^z\rangle, \langle x, \eta_1^z\rangle\right),$$

where $\arctan_2(y, x)$ denotes the $\arctan$ of $\frac{y}{x}$ with respect to the different branches of the tangent function, which means

$$\cos(\theta(x)) = \frac{\langle x, \eta_1^z\rangle}{\sin(d(x, z))},$$
$$\sin(\theta(x)) = \frac{\langle x, \eta_2^z\rangle}{\sin(d(x, z))}.$$

For each $x \neq z, -z$, the vectors

$$\gamma_1^x = \frac{\partial \varphi^{\text{pol}}}{\partial r}(r(x), \theta(x)), \quad \gamma_2^x = \frac{1}{\sin(r(x))}\frac{\partial \varphi^{\text{pol}}}{\partial \theta}(r(x), \theta(x)) \tag{5.3}$$

form an orthonormal basis of the tangent space $T_x\mathbb{S}^2$. Notably, we have $\gamma_2^x = \gamma_1^x \times x$ and $\gamma_1^x = x \times \gamma_2^x$.

In these coordinates, the Riemannian metric takes the form

$$g(r, \theta) = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2(r) \end{pmatrix},$$

and the Christoffel symbols in this coordinates are therefore given by

$$\Gamma^r(r, \theta) = \begin{pmatrix} 0 & 0 \\ 0 & -\sin(r)\cos(r) \end{pmatrix}, \quad \Gamma^\theta(r, \theta) = \begin{pmatrix} 0 & \cot(r) \\ \cot(r) & 0 \end{pmatrix}.$$

For a two times differentiable function f and $x \in \mathbb{S}^2 \setminus \{z, -z\}$, the Hessian with respect to the polar coordinates centered at z, i.e. with respect to the basis (5.3), is represented by the matrix

$$Hf = \begin{pmatrix} \frac{\partial^2 f \circ \varphi^{\text{pol}}}{\partial r^2} & \frac{1}{\sin(r)}\frac{\partial^2 f \circ \varphi^{\text{pol}}}{\partial r \partial \theta} \\ \frac{1}{\sin(r)}\frac{\partial^2 f \circ \varphi^{\text{pol}}}{\partial \theta \partial r} & \frac{1}{\sin^2(r)}\frac{\partial^2 f \circ \varphi^{\text{pol}}}{\partial \theta^2} \end{pmatrix} - \frac{1}{\sin^2(r)}\frac{\partial f \circ \varphi^{\text{pol}}}{\partial r}\Gamma^r - \frac{1}{\sin(r)}\frac{\partial f \circ \varphi^{\text{pol}}}{\partial \theta}\Gamma^\theta.$$

Next, we describe the involved basis functions known as *spherical harmonics*. For a detailed overview, see [Atkinson and Han, 2012]. The space $L^2(\mathbb{S}^2)$ is given by all functions f such that

$$\|f\|_2 := \left(\int_{\mathbb{S}^2} |f(x)|^2 \mathrm{d}\Omega(x)\right)^{\frac{1}{2}} < \infty,$$

where Ω is the Riemannian volume form on $\mathbb{S}^2$ for the metric g. In spherical coordinates, this can be written as

$$\|f\|_2 = \left(\int_0^{2\pi}\int_0^{\pi} |f(r, \theta)|^2 \sin(r)\mathrm{d}r\mathrm{d}\theta\right)^{\frac{1}{2}}.$$

It is well known, that the space $L^2(\mathbb{S}^2)$ can be decomposed into an orthogonal sum

$$L^2(\mathbb{S}^2) = \mathrm{cl}_{\|\cdot\|_{L^2}} \bigoplus_{l=0}^{\infty} H_l,$$

where H_l is the eigenspace to the eigenvalue $\lambda_l = -l(l+1)$ of the Laplace-Beltrami operator on $\mathbb{S}^2$. We have $\dim(H_l) = 2l+1$. Set

$$\Pi_N(\mathbb{S}^2) = \bigcup_{l=0}^{N} H_l,$$

then for $f \in \Pi_{N_1}(\mathbb{S}^2)$, $g \in \Pi_{N_2}(\mathbb{S}^2)$

$$f \cdot g \in \Pi_{N_1+N_2}(\mathbb{S}^2).$$

We choose the orthonormal basis that separates in spherical coordinates. They are given in spherical coordinates by

$$Y_m^l(x(r,\theta)) = N_{lm} P_l^m(\cos(r)) \mathrm{e}^{im\theta}, \quad -l \le m \le l,$$

where

$$N_{lm} = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{2l+1}{2} \frac{(l-m)!}{(l+m)!}}$$

is the normalization constant and

$$P_l^m(t) = \frac{(-1)^m}{2^l l!} (1-t^2)^{\frac{m}{2}} \frac{\mathrm{d}^{l+m}}{\mathrm{d}t^{l+m}} (t^2-1)^l$$

are called *associated Legendre polynomials*. With this, we have that the system

$$\{Y_m^l : l \in \mathbb{N}, -l \le m \le l\}$$

is an orthogonal basis of $L^2(\mathbb{S}^2)$. Moreover, the following addition theorem is true

$$P_l(\langle x, y\rangle) := P_l^0(\langle x, y\rangle) = \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_m^l(x) \overline{Y_m^l(y)}.$$

Thus, the projection $\mathcal{P}_N : L^2(\mathbb{S}^2) \to C(\mathbb{S}^2)$, onto the space $\Pi_N(\mathbb{S}^2)$ can be written as

$$\mathcal{P}_N f(x) = \int_{\mathbb{S}^2} f(y) D_N(x,y) \mathrm{d}\Omega(y),$$

with

$$D_N(x,y) = \sum_{l=0}^{N} \frac{2l+1}{4\pi} P_l(\langle x, y\rangle).$$

With this preparation, we are now able to describe the problem of super-resolution on the sphere. Given a discrete measure

$$\mu^\star = \sum_{i=1}^{M} c_i \delta_{x_i}$$

consisting of M support points x_i with amplitudes c_i, the super-resolution problem on the sphere is to recover the unknown parameters $\mathcal{X} = \{x_i\}, c = (c_i)$ from the low frequency information

$$\mathcal{P}_N^* \mu^\star(x) = \int_{\mathbb{S}^2} D_N(x,y) \mathrm{d}\mu^\star(y).$$

Again, we discuss the recovery of $\mu^\star$ via the minimization

$$\min_{\mu \in \mathcal{M}(\mathbb{S}^2, \mathbb{R})} \|\mu\|_{TV}, \quad \text{subject to} \quad \mathcal{P}_N^* \mu = \mathcal{P}_N^* \mu^\star. \tag{SP}$$

As seen in Chapter 1, see Theorem 1.7, to ensure that $\mu^\star$ is the unique minimizer of (SP), one has to construct for each sign sequence u_i a function $q \in \Pi_N(\mathbb{S}^2)$, such that

$$\begin{aligned} q(x_i) &= u_i, \quad x_i \in \mathcal{X}, \\ |q(x)| &< 1, \quad x \in \mathbb{S}^2 \setminus \mathcal{X}. \end{aligned}$$

In complete analogy to the previous discussion regarding the recovery on the rotation group, see (1.16), we formulate the Hermite interpolation problem

$$\begin{aligned} q(x_i) &= u_i, \\ X_1 q(x_i) &= X_2 q(x_i) = 0, \end{aligned}$$

for $x_i \in \mathcal{X}$, and $X_n q$ is defined in (5.2). In order to tackle the interpolation problem, we choose a kernel $J_N : \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{R}$, such that $J_N(\cdot, y), X_n^y J_N(\cdot, y) \in \Pi_N(\mathbb{S}^2)$ for all $y \in \mathbb{S}^2$, and solve for the coefficient vector in the linear system

$$K\alpha := \begin{pmatrix} J_N & X_1^x J_N & X_2^x J_N \\ X_1^y J_N & X_1^x X_1^y J_N & X_2^x X_1^y J_N \\ X_2^y J_N & X_1^x X_2^y J_N & X_2^x X_2^y J_N \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} u \\ 0 \\ 0 \end{pmatrix}.$$

Thus, the interpolant is of the form

$$q(x) = \sum_{i=1}^{M} a_{0,i} J_N(x, x_i) + a_{1,i} X_1^y J_N(x, x_i) + a_{2,i} X_2^y J_N(x, x_i) \in \Pi_N(\mathbb{S}^2).$$

To control the interpolant properly, we need localization results for the kernel J_N and its derivatives. In the next section, we therefore choose a specific kernel, that is of the form

$$J_N(x, y) = \tilde{J}_N(d(x, y))$$

and show analog localization estimates similar to those in Chapter 2.

5.2 Localization Estimates

The aim of this section is to show the necessary localization results for the construction of the dual certificate. We start by discussing the choice of the interpolation kernel. First, the interpolation kernel must have a polynomial expansion. To be more concrete, we choose a kernel that has an expansion of the form

$$J_N(x, y) = \sum_{l=0}^{N} \hat{w}_l \sum_{m=-l}^{l} Y_m^l(x) \overline{Y_m^l(y)} = \sum_{l=0}^{N} \frac{2l+1}{4\pi} \hat{w}_l P_l(\langle x, y \rangle),$$

which means J_N is a zonal function. Thus, by construction we have $J_N(\cdot, y), X_n^y J_N(\cdot, y) \in \Pi_N(\mathbb{S}^2)$ for all $y \in \mathbb{S}^2$. In addition, the kernel has a representation of the form

$$J_N(x, y) = \tilde{J}_N(d(x, y)),$$

where $\tilde{J}_N$ is a trigonometric polynomial. We therefore derive estimates of the interpolation kernel from the corresponding bounds on the trigonometric function $\tilde{J}_N$.

As interpolation kernel, we chose the specific kernel, given by

$$J_N(x, y) = \tilde{J}_N(d(x, y)) = \frac{1}{\lfloor N/2 \rfloor + 1} \frac{\sin^4((\lfloor N/2 \rfloor + 1) d(x, y)/2)}{\sin^4(d(x, y)/2)}, \tag{5.4}$$

i.e. the classical trigonometric Jackson kernel evaluated at the distance of $x, y \in \mathbb{S}^2$.

Lemma 5.1. *The Jackson kernel $J_N(x,y)$ has an expansion of the form*

$$J_N(x,y) = \sum_{l=0}^{N} \frac{2l+1}{4\pi} \hat{w}_l P_l(\langle x,y\rangle)$$

with positive Legendre coefficients

$$\hat{w}_l = 2\pi \int_{-1}^{1} P_l(t)\tilde{J}_N(\arccos(t))\mathrm{d}t,$$

and thus $J_N(\cdot,y), X_n^y J_N(\cdot,y) \in \Pi_N(\mathbb{S}^2)$ for all $y \in \mathbb{S}^2$.

Proof. Set for $t \in [-1,1]$

$$\tilde{F}_n(\arccos(t)) = \frac{1}{(n+1)^2} \frac{\sin^2((n+1)\arccos(t)/2)}{\sin^2(\arccos(t)/2)},$$

i.e. the Fejér kernel evaluated at $\arccos(t)$. Then

$$\tilde{J}_N(\arccos(t)) = \tilde{F}_n^2(\arccos(t)), \quad n = \left\lfloor \frac{N}{2} \right\rfloor.$$

In [Keiner et al., 2007, Lemma 7], it was shown, that

$$\tilde{F}_n(\arccos(t)) = \sum_{l=1}^{M} \frac{2l+1}{4\pi} \hat{v}_l P_l(t),$$

which shows $\tilde{F}_n(d(x,\cdot)), \tilde{F}_n(d(\cdot,y)) \in \Pi_n(\mathbb{S}^2)$ for all $x,y \in \mathbb{S}^2$ and therefore $\tilde{J}_N(d(x,\cdot)), \tilde{J}_N(d(\cdot,y)) \in \Pi_N(\mathbb{S}^2)$. The positivity of the coefficients $\hat{w}_l$ follows from the positivity of the linearization coefficients of a product of two Legendre polynomials

$$P_n(t)P_m(t) = \sum_{l=0}^{\min(m,n)} \frac{2m+2n-4l+1}{2m+2n-2l+1} \frac{A(m-l)A(l)A(n-l)}{A(n+m-l)} P_{m+n-2l}(t),$$

where

$$A(m) = \frac{1\cdot 3\cdot 5\cdot \dots \cdot (2m-1)}{m!},$$

see e.g. [Adams, 1878]. □

Before we state the necessary localization estimates for the Jackson kernel, we would like to mention a different behavior for these kernels, than for the kernels we discussed on the rotation group. It turns out, that the geodesic distance on the sphere behaves differently with respect to the boundedness of the derivatives. As seen in Chapter 2, the geodesic distance $\omega(x^{-1}y)$ is not differentiable at $x=y$ and whenever $\omega(x^{-1}y)=\pi$, i.e. y is an element of the *cut-locus* of x. Nevertheless, the derivatives only have a true pole at $x=y$, which can be handled using the pointwise-wise estimates of Lemma 2.5, which are summand-wise estimates of the trigonometric representations of the kernels. On the sphere, the derivatives of the geodesic distance $d(x,y)$ have true poles at $x=y$ and $x=-y$. Whereas the first case is again covered by estimates similar to those of Lemma 2.5, the second case can not be covered by summand-wise estimation, since the sign of $\cos(kd(x,y))$ and $\sin(kd(x,y))$ alternates with k in a neighborhood of $x=-y$. Consequently, we have do deal with the singularity at $x=-y$ induced by the derivatives of the geodesic distance in a different way. Especially, we have to bound the trigonometric expressions

$$\begin{aligned} G_1(\omega) &= \frac{\tilde{J}_N'(\omega)}{\sin(\omega)}, \quad G_2(\omega) = \tilde{J}_N''(\omega) - \frac{\tilde{J}_N'(\omega)\cos(\omega)}{\sin(\omega)}, \\ G_3(\omega) &= \frac{\tilde{J}_N''(\omega)\sin(\omega) - \tilde{J}_N'(\omega)\cos(\omega)}{\sin^2(\omega)}, \end{aligned} \tag{5.5}$$

that appear in the spherical derivatives. Knowing only asymptotic estimates for $\tilde{J}_N'$ and $\tilde{J}_N''$ is not sufficient, we have to consider the differences in closed form, which prohibits the use of the B-spline kernels, that we considered for the rotation group. To achieve this, we use the closed form expression (5.4) of the Jackson kernel.

Lemma 5.2. *For $|\omega| \neq 0$, we have for $n = \lfloor N/2 \rfloor$*

$$\begin{aligned} |\tilde{J}_N(\omega)| &\leq \frac{\pi^4}{(n+1)^4|\omega|^4}, \quad |\tilde{J}_N'(\omega)| \leq \frac{3\cdot\pi^4}{(n+1)^3|\omega|^4}, \quad |\tilde{J}_N''(\omega)| \leq \frac{12.5\cdot\pi^4}{(n+1)^2|\omega|^4} \\ |G_1(\omega)| &\leq \frac{2\cdot\pi^4}{(n+1)^2|\omega|^4}, \quad |G_2(\omega)| \leq \frac{14.5\cdot\pi^4}{(n+1)^2|\omega|^4}, \quad |G_3(\omega)| \leq \frac{8\cdot\pi^4}{(n+1)|\omega|^4}, \\ |J_N'''(\omega)| &\leq \frac{68\cdot\pi^4}{(n+1)|\omega|^4}, \end{aligned}$$

where G_2, G_2, G_3 are defined in (5.5)*. For $|\omega| \leq \frac{\pi}{4(n+1)}$, we have*

$$\begin{aligned} \left|\tilde{J}_N''(0) - \cos(\omega)G_1(\omega)\right| &\leq \frac{\tilde{J}_N^{(4)}(0)}{2}|\omega|^2, \quad |\tilde{J}_N''(0) - \tilde{J}_N''(\omega)| \leq \frac{\tilde{J}_N^{(4)}(0)}{2}|\omega|^2, \\ |G_2(\omega)| &\leq \frac{\tilde{J}_N^{(4)}(0)}{2}|\omega|^2, \quad |G_3(\omega)| \leq 0.52\cdot\tilde{J}_N^{(4)}(0)|\omega|, \\ \left|\frac{\tilde{J}_N'(\omega) - \cos(\omega)\sin(\omega)\tilde{J}_N''(\omega)}{\sin^2(\omega)}\right| &\leq 0.52\cdot\tilde{J}_N^{(4)}(0)|\omega|, \quad |\tilde{J}_N'''(\omega)| \leq \tilde{J}_N^{(4)}(0)|\omega|. \end{aligned}$$

Moreover,

$$\begin{aligned} \tilde{J}_N(0) &= 1, \quad \tilde{J}_N'(0) = \tilde{J}_N'''(0) = 0, \quad \tilde{J}_N''(0) = -\frac{n(n+2)}{3}, \\ \tilde{J}_N^{(4)}(0) &= \frac{1}{30}n(n+2)(9n(n+2)-2). \end{aligned}$$

Proof. We will use the property

$$|\sin(\omega/2)| \geq \frac{|\omega|}{\pi}, \quad \text{for } \omega \in [-\pi, \pi],$$

for the asymptotic estimates, and component-wise estimates of the expressions for the case $|\omega| \leq \frac{\pi}{4(n+1)}$. The estimate for the kernel itself follows immediately. For the first derivative, we have

$$\tilde{J}_N'(\omega) = 2\tilde{F}_n(\omega)\tilde{F}_n'(\omega)$$

and

$$\tilde{F}_n'(\omega) = \frac{1}{2(n+1)^2\sin^2(\omega/2)}\left((n+1)\sin((n+1)\omega) - \frac{2\cos(\omega/2)\sin^2((n+1)\omega/2)}{\sin(\omega/2)}\right),$$

$$= \frac{1}{2(n+1)^2 \sin^2(\omega/2)} \left((n+1)\sin((n+1)\omega) - 2\cos(\omega/2)\sin((n+1)\omega/2)U_n(\cos(\omega/2)) \right),$$

where U_n denotes the n-th order Chebychev polynomial of the second kind. Since $\|U_n\|_\infty = n$, we get

$$|\tilde{F}_n'(\omega)| \leq \frac{1.5}{(n+1)\sin^2(\omega/2)}.$$

Since the Fejér kernel can be written as

$$\tilde{F}_n(\omega) = \frac{1}{(n+1)} \left(1 + 2\sum_{k=1}^{n} \left(1 - \frac{k}{n+1}\right) \cos(k\omega) \right),$$

we get $\tilde{F}_n'(\pi) = 0$ and therefore $\tilde{J}_N'(\pi) = 0$. Moreover, we have

$$\begin{aligned} \frac{\tilde{F}'(\omega)}{\sin(\omega)} &= \frac{1}{2(n+1)^2 \sin^2(\omega/2)} \left((n+1)\frac{\sin((n+1)\omega)}{\sin(\omega)} - \frac{2\cos(\omega/2)\sin^2((n+1)\omega/2)}{\sin(\omega)\sin(\omega/2)} \right), \\ &= \frac{1}{2(n+1)^2 \sin^2(\omega/2)} \left((n+1)U_n(\cos(\omega)) - (n+1)^2 \tilde{F}_n(\omega) \right). \end{aligned}$$

Since $\|\tilde{F}_n\|_\infty = 1$, we get

$$\left| \frac{\tilde{F}_n'(\omega)}{\sin(\omega)} \right| \leq \frac{1}{\sin^2(\omega/2)},$$

and

$$\lim_{\omega \to 0} \frac{\tilde{F}_n'(\omega)}{\sin(\omega)} = \tilde{F}_n''(0).$$

For $|\omega| \leq \frac{\pi}{4(n+1)}$, as seen in (2.25), we have,

$$k^2 \cos(k\omega) - \frac{k\sin(k\omega)\cos(\omega)}{\sin(\omega)} \leq 0,$$

which yields

$$\begin{aligned} \left| k^2 - \frac{k\sin(k\omega)\cos(\omega)}{\sin(\omega)} \right| = k^2 - \frac{k\sin(k\omega)\cos(\omega)}{\sin(d(x,z))} &\leq k^2(1 - \cos(k\omega))), \\ &\leq k^4 \frac{\omega^2}{2}. \end{aligned}$$

Since,

$$\tilde{J}_N'(\omega) = \frac{1}{(n+1)^2} \left(-2\sum_{k=1}^{2n} c_k k \sin(k\omega) \right),$$

with positive Fourier coefficients c_k, component-wise estimation shows

$$\left| \tilde{J}_N''(0) - \frac{\cos(\omega)\tilde{J}_N'(\omega)}{\sin(\omega)} \right| \leq \frac{\tilde{J}^{(4)}(0)}{2}\omega^2.$$

In the same way, one derives

$$|\tilde{J}_N''(0) - \tilde{J}_N''(\omega)| \leq \frac{\tilde{J}^{(4)}(0)}{2}\omega^2.$$

We compute

$$G_2(\omega) = \tilde{J}_N''(\omega) - \frac{\tilde{J}_N'(\omega)\cos(\omega)}{\sin(\omega)} = 2\left(\left(\tilde{F}_n'(\omega)\right)^2 + \tilde{F}_n(x)\left(\tilde{F}_n''(\omega) - \frac{\tilde{F}_n'(\omega)\cos(\omega)}{\sin(\omega)}\right)\right).$$

Since

$$\begin{aligned}\tilde{F}_n''(\omega) = \frac{1}{2(n+1)^2\sin^2(\omega/2)}\Big(&(n+1)^2((2+\cos(\omega))\tilde{F}_n(\omega) + \cos((n+1)\omega))\\ &- 2(n+1)(1+\cos(\omega))U_n(\cos(\omega))\Big),\end{aligned}$$

we get

$$\begin{aligned}\tilde{F}_n''(\omega) - \frac{\tilde{F}_n'(\omega)\cos(\omega)}{\sin(\omega)} &= \frac{1}{2(n+1)^2\sin^2(\omega/2)}\Big((n+1)^2(2(1+\cos(\omega))\tilde{F}_n(\omega)\\ &\quad + \cos((n+1)\omega)) - (n+1)(2+3\cos(\omega))U_n(\cos(\omega))\Big),\\ &= \frac{1}{2(n+1)^2\sin^2(\omega/2)}\Big((n+1)^2(2(1+\cos(\omega))\left(\tilde{F}_n(\omega) - \frac{U_n(\cos(\omega))}{n+1}\right)\\ &\quad + (n+1)^2\cos((n+1)\omega) - (n+1)\cos(\omega)U_n(\cos(\omega))\Big)\end{aligned}$$

and inserting yields

$$\left|\tilde{J}_N''(\omega) - \frac{\tilde{J}_N'(\omega)\cos(\omega)}{\sin(\omega)}\right| \le \frac{14.5\pi^4}{(n+1)^2|\omega|^4},$$

and similarly

$$\left|\tilde{J}_N''(\omega)\right| \le \frac{12.5\pi^4}{(n+1)^2|\omega|^4}.$$

For $|\omega| \le \frac{\pi}{4(n+1)}$, the estimate follows again by estimating component-wise in the trigonometric representation, yielding

$$\left|\tilde{J}_N''(\omega) - \frac{\tilde{J}_N'(\omega)\cos(\omega)}{\sin(\omega)}\right| \le \frac{\tilde{J}^{(4)}(0)}{2}\omega^2.$$

Furthermore, we have

$$\tilde{J}''(0) = 2\tilde{F}''(0) = -\frac{4}{(n+1)}\sum_{k=1}^{n}\left(k^2 - \frac{k^3}{(n+1)}\right) = -\frac{n(n+2)}{3}.$$

Next, we consider the expression

$$G_3(\omega) = \frac{\tilde{J}_N''(\omega)\sin(\omega) - \tilde{J}_N'(\omega)\cos(\omega)}{\sin^2(\omega)}.$$

Observe, that for $\omega = \pi$ the expression is zero. For $\omega \neq \pi$, we get

$$\frac{\tilde{J}_N''(\omega)\sin(\omega) - \tilde{J}_N'(\omega)\cos(\omega)}{\sin^2(\omega)} = 2\left(\left(\frac{\tilde{F}_n'(\omega)}{\sin(\omega)}\right)\tilde{F}_n'(\omega) + \tilde{F}_n(x)\left(\frac{\tilde{F}_n''(\omega)\sin(\omega) - \tilde{F}_n'(\omega)\cos(\omega)}{\sin^2(\omega)}\right)\right),$$

and

$$\tilde{F}_n(\omega)\left(\frac{\tilde{F}_n''(\omega)\sin(\omega)-\tilde{F}_n'(\omega)\cos(\omega)}{\sin^2(\omega)}\right)=\frac{\sin^2((n+1)/2)}{2(n+1)^4\sin^4(\omega/2)}\Big(2(n+1)^2\frac{(1+\cos(\omega))}{\sin(\omega)}\Big(\tilde{F}_n(\omega)-\frac{U_n(\cos(\omega))}{n+1}\Big)+(n+1)\frac{(n+1)\cos((n+1)\omega)-\cos(\omega)U_n(\cos(\omega))}{\sin(\omega)}\Big).$$

For the first summand of the righthand site, we have the bound

$$\begin{aligned}&\left|\frac{\sin^2((n+1)/2)}{(n+1)^2\sin^4(\omega/2)}\frac{(1+\cos(\omega))}{\sin(\omega)}\left(\tilde{F}_n(\omega)-\frac{U_n(\cos(\omega))}{n+1}\right)\right|\\
&=\frac{1}{(n+1)\sin^4(\omega/2)}\left|\frac{\sin^2((n+1)/2)}{(n+1)}\frac{\cos(\omega/2)}{\sin(\omega/2)}\left(\tilde{F}_n(\omega)-\frac{U_n(\cos(\omega))}{n+1}\right)\right|,\\
&=\frac{1}{(n+1)\sin^4(\omega/2)}\left|\frac{\sin((n+1)/2)}{(n+1)\sin(\omega/2)}\right|\left|\sin((n+1)/2)\cos(\omega/2)\left(\tilde{F}_n(\omega)-\frac{U_n(\cos(\omega))}{n+1}\right)\right|,\\
&=\frac{\sqrt{\tilde{F}_n(\omega)}}{(n+1)\sin^4(\omega/2)}\left|\sin((n+1)/2)\cos(\omega/2)\left(\tilde{F}_n(\omega)-\frac{U_n(\cos(\omega))}{n+1}\right)\right|\le\frac{2}{(n+1)\sin^4(\omega/2)},\end{aligned}$$

For the second summand, observe that, due to the derivative representation of the Chebyshev polynomials, we have

$$\frac{(n+1)\cos((n+1)\omega)-\cos(\omega)U_n(\cos(\omega)}{\sin(\omega)}=-\sin(\omega)U_n'(\cos(\omega)).$$

Using the Bernstein inequality for algebraic polynomials, i.e.

$$|P_n'(t)|\le\frac{n}{\sqrt{1-x^2}}\|P_n\|_\infty,\quad -1<x<1,$$

for a polynomial of degree n, see e.g. [Bernstein, 1912], we derive

$$\left|\frac{(n+1)\cos((n+1)\omega)-\cos(\omega)U_n(\cos(\omega)}{\sin(\omega)}\right|\le n\|U_n\|_\infty\le(n+1)^2.$$

Hence, we can estimate

$$\left|\frac{\sin^2((n+1)/2)}{2(n+1)^4\sin^4(\omega/2)}(n+1)\frac{(n+1)\cos((n+1)\omega)-\cos(\omega)U_n(\cos(\omega))}{\sin(\omega)}\right|\le\frac{0.5}{(n+1)\sin^4(\omega/2)}.$$

Together, this yields

$$\left|\tilde{F}_n(\omega)\left(\frac{\tilde{F}_n''(\omega)\sin(\omega)-\tilde{F}_n'(\omega)\cos(\omega)}{\sin^2(\omega)}\right)\right|\le\frac{2.5}{(n+1)\sin^4(\omega/2)}$$

and consequently

$$\left|\frac{\tilde{J}_N''(\omega)\sin(\omega)-\tilde{J}_N'(\omega)\cos(\omega)}{\sin^2(\omega)}\right|\le\frac{8}{(n+1)\sin^4(\omega/2)}.$$

Again, for $|\omega| \leq \frac{\pi}{4(n+1)}$, we estimate the expression component-wise. Observe, that

$$\left|\frac{k^2\cos(k\omega)\sin(\omega) - k\sin(k\omega)\cos(\omega)}{\sin^2(\omega)}\right| \leq \frac{1}{1+\cos(\omega)}k^3\sin(k\omega), \\ \leq \frac{k^4|\omega|}{1+\cos(\omega)},$$

see (2.23). For this reason, we have for $|\omega| \leq \frac{\pi}{4(n+1)}$

$$\left|\frac{\tilde{J}_N''(\omega)\sin(\omega) - \tilde{J}_N'(\omega)\cos(\omega)}{\sin^2(\omega)}\right| \leq \frac{\tilde{J}_N^{(4)}(0)|\omega|}{1+\cos(\omega)} \leq \frac{\tilde{J}_N^{(4)}(0)|\omega|}{1+\cos(\pi/(4(n+1)))}, \\ \leq \frac{\tilde{J}_N^{(4)}(0)|\omega|}{1+\cos(\pi/8)} \leq 0.52 \cdot \tilde{J}_N^{(4)}(0)|\omega|.$$

Very similar to (2.23), one shows

$$\left|\frac{k^2\cos(k\omega)\cos(\omega)\sin(\omega) - k\sin(k\omega)}{\sin^2(\omega)}\right| \leq \frac{1}{1+\cos(\omega)}k^3\sin(k\omega), \\ \leq \frac{k^4|\omega|}{1+\cos(\omega)},$$

and thus

$$\left|\frac{\tilde{J}_N'(\omega) - \cos(\omega)\sin(\omega)\tilde{J}_N''(\omega}{\sin^2(\omega)}\right| \leq 0.52 \cdot \tilde{J}_N^{(4)}(0)|\omega|.$$

Similarly, we can calculate the third derivative

$$\tilde{J}_N'''(\omega) = 2(3\tilde{F}_n'(\omega)\tilde{F}_n''(\omega) + \tilde{F}_n(\omega)\tilde{F}_n'''(\omega)).$$

Since,

$$\tilde{F}_n(x)\tilde{F}_n'''(x) = \frac{1}{(n+1)\sin^4(x/2)}\Big(- \frac{1}{2(n+1)}\cos(x/2)\sin((n+1)x/2)U_n(\cos(x/2)) \cdot \\ ((5+\cos(x))\tilde{F}_n(x) + \cos((n+1)x)) + \frac{1}{4}\sin((n+1)x \cdot \\ (\tilde{F}_n(x)(10+\cos(x)) + \cos((n+1)x) - 1)\Big)$$

and

$$\tilde{F}_n'(x)\tilde{F}_n''(x) = \frac{1}{3(n+1)\sin^4(x/2)}\Big(- \frac{3}{2(n+1)}\cos(x/2)\sin((n+1)x/2)U_n(\cos(x/2)) \cdot \\ ((2+\cos(x))\tilde{F}_n(x) + 5\cos((n+1)x) + 4) + \frac{3}{4}\sin((n+1)x) \cdot \\ (\tilde{F}_n(x)(4+3\cos(x)) + \cos((n+1)x))\Big),$$

we estimate using the triangle inequality

$$\left|\tilde{J}_N'''(\omega)\right| \leq \frac{68 \cdot \pi^4}{(n+1)|\omega|^4}.$$

With $\tilde{J}_N'''(\pi) = 0$, we get for $|\omega| \leq \frac{\pi}{4(n+1)}$

$$\left|\tilde{J}_N'''(\omega)\right| \leq \tilde{J}_N^{(4)}(0)|\omega|.$$

The fourth derivative can be written as

$$\tilde{J}_N^{(4)}(\omega) = 2\left(3\left(\tilde{F}_n''(\omega)\right)^2 + 4\tilde{F}_n'(\omega)\tilde{F}_n'''(\omega) + \tilde{F}_n(\omega)\tilde{F}_n^{(4)}(\omega)\right)$$

and therefore

$$\tilde{J}_N^{(4)}(0) = 2\left(3\left(\tilde{F}_n''(0)\right)^2 + \tilde{F}_n^{(4)}(0)\right).$$

Because of

$$\tilde{F}_n^{(4)}(0) = \frac{2}{n+1}\sum_{k=1}^{n}\left(k^4 - \frac{k^5}{n+1}\right) = \frac{n(n+2)}{30}(2n(n+2)-1),$$

we get

$$\tilde{J}_N^{(4)}(0) = \frac{1}{30}n(n+2)(9n(n+2)-2).$$

□

Now, we are able to state the localization estimates for the spherical derivatives of the Jackson kernel using the bounds for the trigonometric expressions derived in the previous Lemma. More precise, the bounds on the derivatives in normal coordinates are stated in Theorem 5.3 and those in polar coordinates are derived in Theorem 5.4. In the proofs, we make use of several identities, see Appendix C, regarding the cross product in $\mathbb{R}^3$, which for abbreviation is not detailed in the proofs.

Theorem 5.3. *The Jackson kernel fulfills for $x \neq y$ and $n = \lfloor N/2 \rfloor$*

$$|J_N(x,y)| \leq \frac{\pi^4}{(n+1)^4 d(x,y)^4}, \quad |X_n^y J_N(x,y)| \leq \frac{3\cdot\pi^4}{(n+1)^3 d(x,y)^4}$$
$$|X_i^x X_n^y J_N(x,y)| \leq \frac{16.5\cdot\pi^4}{(n+1)^2 d(x,y)^4}, \quad |X_i^x X_n^x J_N(x,y)| \leq \frac{16.5\cdot\pi^4}{(n+1)^2 d(x,y)^4},$$
$$|X_j^x X_i^x X_n^y J_N(x,y)| \leq \frac{110\cdot\pi^4}{(n+1)d(x,y)^4}$$

and for $x = y$

$$J_N(x,x) = 1, \quad X_i^x X_i^y J_N(x,x) = -X_i^x X_n^x J_N(x,x) = -\tilde{J}_N''(0),$$
$$X_n^y J_N(x,x) = X_i^x X_n^y J_N(x,x) = X_j^x X_i^x X_n^y J_N(x,x) = 0.$$

Proof. The first estimate follows directly from Lemma 5.2. For the second estimate, we first calculate the derivative to get

$$X_n^y J(x,y) = \sum_{l=0}^{N} \frac{2l+1}{4\pi}\hat{w}_l P_l'(\langle x,y\rangle)\langle x,\eta_n^y\rangle,$$

which immediately yields $X_n^y J(x,x) = X_n^y J(x,-x) = 0$. In the case $x \neq -y, y$, we can calculate

$$X_n^y J_N(x,y) = -\tilde{J}_N'(d(x,y))\frac{\langle x,\eta_n^y\rangle}{\sin(d(x,y))},$$

$$
\begin{aligned}
&= \pm \tilde{J}_N'(d(x,y)) \frac{\langle x, \eta_i^y \times y \rangle}{\sin(d(x,y))}, \\
&= \mp \tilde{J}_N'(d(x,y)) \frac{\langle \eta_i^y, x \times y \rangle}{\sin(d(x,y))}, \\
&= \mp \tilde{J}_N'(d(x,y)) \langle \eta_i^y, n_{x,y} \rangle,
\end{aligned}
$$

where $n_{x,y}$ denotes the unique unit vector perpendicular to x and y, which yields together with Lemma 5.2 the second estimate. For the next estimate, we compute the derivatives in the same way to derive

$$
\begin{aligned}
X_i^x X_n^y J_N(x,y) &= \sum_{l=0}^{N} \frac{2l+1}{4\pi} \hat{w}_l (P_l''(\langle x,y \rangle) \langle x, \eta_n^y \rangle \langle y, \eta_i^x \rangle + P_l'(\langle x,y \rangle) \langle \eta_i^x, \eta_n^y \rangle), \\
X_i^x X_n^x J_N(x,y) &= \sum_{l=0}^{N} \frac{2l+1}{4\pi} \hat{w}_l (P_l''(\langle x,y \rangle) \langle \eta_n^x, y \rangle \langle y, \eta_i^x \rangle - \delta_{in} P_l'(\langle x,y \rangle) \langle x,y \rangle).
\end{aligned}
$$

For $i \neq n$, this shows $X_i^x X_n^y J_N(x,x) = X_i^x X_n^y J_N(x,-x) = X_i^x X_n^x J_N(x,x) = X_i^x X_n^x J_N(x,-x) = 0$. In the case $i = n$, we have

$$
\begin{aligned}
-X_i^x X_n^x J_N(x,x) = X_i^x X_i^y J_N(x,x) &= \sum_{l=0}^{N} \frac{2l+1}{4\pi} \hat{w}_l P_l'(1), \\
&= \lim_{t \to 1} - \frac{\tilde{J}_N'(\arccos(t))}{\sqrt{1-t^2}} = \lim_{\omega \to 0} - \frac{\tilde{J}_N'(\omega)}{\sin(\omega)}, \\
&= -\tilde{J}_N''(0).
\end{aligned} \tag{5.6}
$$

In the same way, one shows

$$
\begin{aligned}
-X_i^x X_i^x J_N(x,-x) = X_i^x X_i^y J_N(x,-x) &= -\sum_{l=0}^{N} \frac{2l+1}{4\pi} \hat{w}_l P_l'(-1), \\
&= \lim_{\omega \to \pi} \frac{\tilde{J}_N'(\omega)}{\sin(\omega)} = -\tilde{J}_N''(\pi).
\end{aligned}
$$

For $x \neq y, -y$, the derivatives have the form

$$
\begin{aligned}
X_i^x X_n^y J_N(x,y) &= \tilde{J}_N''(d(x,y)) \frac{\langle x, \eta_n^y \rangle \langle \eta_i^x, y \rangle}{\sin^2(d(x,y))} \\
&\quad - \tilde{J}_N'(d(x,y)) \left(\frac{\langle x, \eta_n^y \rangle \langle \eta_i^x, y \rangle \cos(d(x,y))}{\sin^3(d(x,y))} + \frac{\langle \eta_i^x, \eta_n^y \rangle}{\sin(d(x,y))} \right), \\
&= \langle \eta_i^y, n_{x,y} \rangle \langle \eta_n^x, n_{x,y} \rangle G_2(\omega) - \langle \eta_i^x, \eta_n^y \rangle G_1(\omega),
\end{aligned}
$$

and

$$
X_i^x X_n^x J_N(x,y) = \langle \eta_i^x, n_{x,y} \rangle \langle \eta_n^x, n_{x,y} \rangle G_2(\omega) + \delta_{in} \cos(d(x,y)) G_1(\omega),
$$

where G_1, G_2 are defined in (5.5) and again $n_{x,y}$ denotes the unique unit vector perpendicular to both x and y. This yields

$$
|X_i^x X_n^x J_N(x,y)|, |X_i^x X_n^y J_N(x,y)| \leq |G_2(d(x,y))| + |G_1(\omega)|
$$

$$\leq \frac{16.5 \cdot \pi^4}{(n+1)^2 d(x,y)^4},$$

which are the estimates for the second derivatives. For the third derivatives, observe that

$$X_j^x X_i^x X_n^y J_N(x,y) = \sum_{l=0}^{N} \frac{2l+1}{4\pi} \hat{w}_l \Big(P_l'''(\langle x,y\rangle)\langle \eta_j^x, y\rangle\langle \eta_i^x, y\rangle\langle \eta_j^y, x\rangle + P_l''(\langle x,y\rangle)\langle \eta_i^x, y\rangle\langle \eta_n^y, \eta_j^x\rangle$$
$$P_l''(\langle x,y\rangle)\langle \eta_i^x, \eta_n^y\rangle\langle y, \eta_i^x\rangle$$
$$- \delta_{ij}(P_l''(\langle x,y\rangle) + P_l''(\langle x,y\rangle))\langle x,y\rangle\langle x, \eta_n^y\rangle \Big),$$

which shows

$$X_j^x X_i^x X_n^y J_N(x,x) = X_j^x X_i^x X_n^y J_N(x,-x) = 0.$$

Again, for $x \neq y, -y$ we compute

$$X_j^x X_i^x X_n^y J_N(x,y) = \frac{\langle \eta_j^x, \eta_n^y\rangle\langle \eta_i^x, y\rangle - \delta_{ij}\langle x, \eta_n^y\rangle\langle x,y\rangle}{\sin(d(x,y))} G_2(\omega)$$
$$- \frac{\langle x, \eta_n^y\rangle\langle \eta_i^x, y\rangle\langle \eta_j^x, y\rangle\langle x,y\rangle}{\sin^3(d(x,y))}(\tilde{J}_N'''(d(x,y)) + \tilde{J}_N'(d(x,y)) - 3\cos(d(x,y))G_2(\omega))$$
$$+ \frac{\langle \eta_i^x, \eta_n^y\rangle\langle \eta_j^x, y\rangle}{\sin(d(x,y))} G_3(\omega) + \delta_{i,j}\cos(d(x,y))G_1(\omega).$$

Hence, using the estimates of Lemma 5.2, we have the bound

$$|X_j^x X_i^x X_n^y J_N(x,y)| \leq \frac{110 \cdot \pi^4}{(n+1)d(x,y)^4}. \tag{5.7}$$

□

Theorem 5.4. *For $x, y, z \in \mathbb{S}^2$, pairwise different and $x \neq -z$, with $n = \lfloor N/2 \rfloor$, the entries of the Hessian of $J_N(\cdot, y)$, $X_N^y J_N(\cdot, y)$ in polar coordinates centered at z obey*

$$|(HJ_N(x,y))_{ii}| \leq \frac{16.5 \cdot \pi^4}{(n+1)^2 d(x,y)^4}, \quad |(HJ_N(x,y))_{ij}| \leq \frac{14.5 \cdot \pi^4}{(n+1)^2 d(x,y)^4}$$
$$|(HX_n^y J_N(x,y))_{ii}| \leq \frac{117 \cdot \pi^4}{(n+1)d(x,y)^4}, \quad |(HX_n^y J_N(x,y))_{ij}| \leq \frac{109 \cdot \pi^4}{(n+1)d(x,y)^4}.$$

In the case $y = z$ and $d(x,z) \leq \frac{\delta}{(n+1)}$, with $\delta \leq \pi/4$, we get

$$|\tilde{J}_N''(0) - (HJ_N(x,z))_{ii}| \leq \frac{3}{20}\delta^2(n+1)^2, \quad |(HJ_N(x,z))_{ij}| \leq \frac{3}{20}\delta^2(n+1)^2$$
$$|(HX_n^y J_N(x,z))_{11}| \leq \frac{3}{10}\delta(n+1)^3, \quad |(HX_n^y J_N(x,z))_{22}| \leq \frac{1}{5}\delta(n+1)^3,$$
$$|(HX_n^y J_N(x,z))_{ij}| \leq \frac{1}{5}\delta(n+1)^3.$$

Proof. Remember, in polar coordinates centered at $z \in \mathbb{S}$, we have the local parametrization

$$\varphi^{\text{pol}}(r,\theta) = \cos(r)z + \sin(r)(\cos(\theta)\eta_1^z + \sin(\theta)\eta_2^z),$$

where η_1^z, η_2^z is an orthonormal basis of $T_z\mathbb{S}^2$, such that $\eta_2^z = \eta_1^z \times z$. The implicit inverse parametrization for $x \neq z, -z$ is given by

$$\begin{aligned} r(x) &= d(x,z), \\ \cos(\theta(x)) &= \frac{\langle x, \eta_1^z\rangle}{\sin(d(x,z))} = -\langle \eta_2^z, n_{z,x}\rangle, \\ \sin(\theta(x)) &= \frac{\langle x, \eta_2^z\rangle}{\sin(d(x,z))} = \langle \eta_1^z, n_{z,x}\rangle. \end{aligned}$$

We first calculate the partial derivatives of the function $f_\xi(r,\theta) = \langle \varphi^{\mathrm{pol}}(r,\theta), \xi\rangle$, for $\xi \in \mathbb{S}^2$, given by

$$\begin{aligned} \partial_r f_\xi(r,\theta) &= -\sin(r)\langle z,\xi\rangle + \cos(r)(\cos(\theta)\langle \gamma_1^z,\xi\rangle + \sin(\theta)\langle \gamma_2^z,\xi\rangle), \\ \partial_\theta f_\xi(r,\theta) &= \sin(r)(\cos(\theta)\langle \gamma_2^z,\xi\rangle - \sin(\theta)\langle \gamma_1^z,\xi\rangle), \\ \partial^2_{rr} f_\xi(r,\theta) &= -\cos(r)\langle z,\xi\rangle - \sin(r)(\cos(\theta)\langle \gamma_1^z,\xi\rangle + \sin(\theta)\langle \gamma_2^z,\xi\rangle), \\ \partial^2_{\theta,\theta} f_\xi(r,\theta) &= -\sin(r)(\cos(\theta)\langle \gamma_1^z,\xi\rangle + \sin(\theta)\langle \gamma_2^z,\xi\rangle), \\ \partial^2_{r\theta} f_\xi(r,\theta) &= \partial^2_{\theta,r} f_\xi(r,\theta) = \cos(r)(\cos(\theta)\langle \gamma_2^z,\xi\rangle - \sin(\theta)\langle \gamma_1^z,\xi\rangle). \end{aligned}$$

Inserting the inverse parametrizations a lengthy calculation shows

$$\begin{aligned} \partial_r f_\xi(r(x),\theta(x)) &= \sin(d(x,\xi))\langle n_{x,\xi}, n_{z,x}\rangle, \\ \partial_\theta f_\xi(r(x),\theta(x)) &= \sin(d(x,z))\sin(d(x,\xi))\langle x, n_{x,\xi}\times n_{z,x}\rangle, \\ \partial^2_{rr} f_\xi(r(x),\theta(x)) &= -\cos(d(x,\xi)), \\ \partial^2_{\theta,\theta} f_\xi(r(x),\theta(x)) &= -\sin(d(x,z))\cos(d(x,z))\sin(d(x,\xi))\langle n_{x,\xi}, n_{z,x}\rangle \\ &\quad - \sin^2(d(x,z))\cos(d(x,\xi))), \\ \partial^2_{r\theta} f_\xi(r(x),\theta(x)) &= \cos(d(x,z))\sin(d(x,\xi))\langle x, n_{x,\xi}\times n_{z,x}\rangle, \end{aligned}$$

where

$$n_{x,\xi} = \frac{x\times\xi}{\|x\times\xi\|_2}.$$

After this preparation, we proceed by calculating the full derivatives. We start with the kernel $J_N(\cdot, y)$ and first assume $x \neq -y$. We calculate

$$\begin{aligned} (HJ_N(x,y))_{11} &= (\partial_r f_y(r(x),\theta(x)))^2 \left(\frac{\tilde{J}_N''(\arccos(f(r(x),\theta(x))))}{(1-f(r(x),\theta(x)))^2} - \frac{\tilde{J}_N'(\arccos(f(r(x),\theta(x))))f(r(x),\theta(x))}{(1-f(r(x),\theta(x))^2)^{3/2}}\right) \\ &\quad - \partial^2_{rr} f_y(r(x),\theta(x))\frac{\tilde{J}_N'(\arccos(f(r(x),\theta(x))))}{(1-f(r(x),\theta(x))^2)^{1/2}}, \\ &= (\langle n_{x,y}, n_{z,x}\rangle)^2 G_2(\omega) + \cos(d(x,y))G_1(\omega), \end{aligned}$$

where G_1, G_2 are defined in (5.5). Thus, using Lemma 5.2 yields

$$|(HJ_N(x,y))_{11}| \le \frac{16.5\cdot\pi^4}{(n+1)^2 d(x,y)^4}. \tag{5.8}$$

If $x = -y$, we again use the polynomial representation of J_N, given by Lemma 5.1, to derive

$$(HJ_N(x,-x))_{11} = \sum_{l=0}^{N} \frac{2l+1}{4\pi}\hat{w}_l(P_l''(\langle x,-x\rangle)\,(\partial_r f(r(x),\theta(x)))^2 + P_l'(\langle x,-x\rangle)\partial^2_{rr} f(r(x),\theta(x)),$$

$$
\begin{aligned}
&= \sum_{l=0}^{N} \frac{2l+1}{4\pi} \hat{w}_l (P_l''(-1)(\langle x \times (-x), n_{z,x}\rangle)^2 - P_l'(-1)\cos(d(x,-x))), \\
&= \sum_{l=0}^{N} \frac{2l+1}{4\pi} \hat{w}_l P_l'(-1).
\end{aligned}
$$

With the same argument as in (5.6), we therefore get

$$(HJ_N(x,-x))_{11} = \tilde{J}_N''(\pi).$$

In the case $y = z$ and $d(x,z) \leq \frac{\pi}{4(n+1)}$, we have

$$(HJ_N(x,z))_{11} = \tilde{J}_N''(d(x,z)),$$

and Lemma 5.2 yields

$$|\tilde{J}_N''(0) - (HJ_N(x,z))_{11}| \leq \frac{\tilde{J}^{(4)}(0)}{2} d(x,z)^2 \leq \frac{3}{20} d(x,z)^2 (n+1)^4.$$

For the second diagonal entry, we compute

$$
\begin{aligned}
(HJ_N(x,y))_{22} &= \langle x, n_{x,y} \times n_{z,x}\rangle^2 \left(\tilde{J}_N''(d(x,y)) - \frac{\tilde{J}_N'(d(x,y))\cos(d(x,y))}{\sin(d(x,y))} \right) \\
&\quad + \cos(d(x,y)) \frac{\tilde{J}_N'(d(x,y))}{\sin(d(x,y))}, \\
&= \langle x, n_{x,y} \times n_{z,x}\rangle^2 G_2(\omega) + \cos(d(x,y)) G_1(\omega),
\end{aligned}
$$

which means

$$|(HJ_N(x,y))_{22}| \leq \frac{16.5\pi^4}{(n+1)^2 d(x,y)^4},$$

and

$$(HJ_N(x,-x))_{22} = 0.$$

In the case $y = z$, we get

$$(HJ_N(x,z))_{22} = \cos(d(x,z)) \frac{\tilde{J}_N'(d(x,z))}{\sin(d(x,z)}.$$

and Lemma 5.2 shows

$$
\begin{aligned}
|\tilde{J}_N''(0) - (HJ_N(x,z))_{22}| &\leq \frac{\tilde{J}^{(4)}(0)}{2} d(x,z)^2, \\
&\leq \frac{3}{20} d(x,z)^2 (n+1)^4.
\end{aligned}
$$

For the off-diagonal entries, we get

$$(HJ_N(x,y))_{ij} = \langle n_{x,y}, n_{z,x}\rangle \langle x, n_{x,y} \times n_{z,x}\rangle G_2(\omega),$$

and analogous estimates to the bounds above, yield

$$|(HJ_N(x,y))_{ij}| \leq \frac{14.5\pi^4}{(n+1)^2 d(x,y)^4},$$

$$(HJ_N(x,-x))_{ij} = 0,$$
$$|(HJ_N(x,z))_{ij}| \leq \frac{3}{20} d(x,z)^2(n+1)^4.$$

For the kernel $X_n^y J_N(\cdot, y)$, the estimates are derived in the same way. Remember,

$$X_n^y J_N(x,y) = -\frac{\tilde{J}'_N(d(x,y))}{\sin(d(x,y))}\langle x, \eta_n^y\rangle,$$

where $\eta_n^y \in T_y\mathbb{S}^2$. Then, with $\xi = \eta_n^y$ and the abbreviations in (5.5)

$$\partial_r X_n^y J_N(x,y) = \frac{\langle x,\xi\rangle \partial_r f_y(x)}{\sin^2(d(x,y)} G_2(d(x,y)) - \partial_r f_\xi(x) G_1(d(x,y)),$$
$$\partial_\theta X_n^y J_N(x,y) = \frac{\langle x,\xi\rangle \partial_\theta f_y(x)}{\sin^2(d(x,y)} G_1(d(x,y)) - \partial_\theta f_\xi(x) G_1(d(x,y)).$$

In polar coordinates centered at z, we have for $y \neq -x$

$$\partial_r H(d(x,y) = -\frac{\partial_r f_y(x)}{\sin(d(x,y)} G_3(d(x,y))$$
$$\partial_\theta H(d(x,y)) = -\frac{\partial_\theta f_y(x)}{\sin(d(x,y))} G_3(d(x,y))$$
$$\partial_r G(d(x,y) = -\frac{\partial_r f_y(x)}{\sin^3(d(x,y))}\left(J'''_N(d(x,y)) + \tilde{J}'_N(d(x,y)) - 3\cos(d(x,y))G_3(d(x,y))\right)$$
$$\partial_\theta G(d(x,y)) = -\frac{\partial_\theta f_y(x)}{\sin^3(d(x,y))}\left(J'''_N(d(x,y)) + \tilde{J}'_N(d(x,y)) - 3\cos(d(x,y))G_3(d(x,y))\right).$$

We therefore compute the first diagonal entry as

$$\begin{aligned}(HX_n^y J_n(x,y))_{11} &= \partial^2_{rr} X_n^y J_N(x,y),\\ &= \frac{\partial_r f_\xi(x)\partial_r f_y(x)}{\sin(d(x,y))} G_3(d(x,y)) + \frac{\langle x,\xi\rangle \partial^2_{rr} f_y(x)}{\sin(d(x,y))} G_3(d(x,y))\\ &\quad + \frac{\langle x,\xi\rangle \partial_r f_y(x)}{\sin(d(x,y))}\partial_r G_3(d(x,y)) - \partial^2_{rr} f_\xi(x) G_1(d(x,y)) - \partial_r f_\xi(x)\partial_r G_1(d(x,y)).\end{aligned}$$

Inserting the precomputed derivatives yields

$$\begin{aligned}(HX_n^y J_n(x,y))_{11} &= 2\langle n_{x,\xi}, n_{z,x}\rangle\langle n_{z,x}, n_{x,y}\rangle \sin(d(x,\xi))G_3(d(x,y))\\ &\quad - \frac{\langle x,\xi\rangle\cos(d(x,y))}{\sin(d(x,y))} G_3(d(x,y)) + \langle x,\xi\rangle G_1(d(x,y))\\ &\quad - \frac{\langle x,\xi\rangle\langle n_{z,x}, n_{x,y}\rangle^2}{\sin(d(x,y))}\left(\tilde{J}'''_N(d(x,y)) + \tilde{J}'_N(d(x,y))\right)\\ &\quad + \frac{3\langle x,\xi\rangle\langle n_{z,x}, n_{x,y}\rangle^2\cos(d(x,y))}{\sin(d(x,y))} G_3(d(x,y)).\end{aligned}$$

Since,

$$\langle x,\xi\rangle = \langle x, \eta_n^y\rangle = \pm\langle \eta_i^y, x\times y\rangle = \pm\langle \eta_i^y, n_{x,y}\rangle \sin(d(x,y))$$

we know that

$$\left|\frac{\langle x,\xi\rangle}{\sin(d(x,y))}\right| \leq 1,$$

and we can estimate using Lemma 5.2 to derive

$$|(HX_n^y J_n(x,y))_{11}| \leq \frac{117 \cdot \pi^4}{(n+1)d(x,y)^4}.$$

In the case $y = -x$, we again use the polynomial representation of J_N to get

$$(HX_n^y J_n(x,-x))_{11} = 0.$$

In the case $y = z$ and $d(x,z) \leq \frac{\pi}{4(n+1)}$, we have

$$(HX_n^y J_n(x,z))_{11} = -\frac{\langle x, \xi \rangle}{\sin(d(x,z))} \tilde{J}_N'''(d(x,z)),$$

and again the use of Lemma 5.2 results in,

$$|(HX_n^y J_n(x,z))_{11}| \leq \frac{1}{30} n(n+2)(9n(n+2)-2)d(x,z) \leq \frac{3}{10}\delta(n+1)^3$$

For the second diagonal entry, we get

$$\begin{aligned}(HX_n^y J_n(x,y))_{22} = {} & 2\langle x, n_{x,\xi} \times n_{z,x}\rangle \langle x, n_{x,y} \times n_{z,x}\rangle \sin(d(x,\xi)) G_3(d(x,y)) \\ & - \frac{\langle x, \xi \rangle \cos(d(x,y))}{\sin(d(x,y))} G_3(d(x,y)) + \langle x, \xi \rangle G_1(d(x,y)) \\ & - \frac{\langle x, \xi \rangle \langle x, n_{x,y} \times n_{z,x}\rangle^2}{\sin(d(x,y))} \left(\tilde{J}_N'''(d(x,y)) + \tilde{J}_N'(d(x,y)) \right) \\ & + \frac{3\langle x, \xi \rangle \langle x, n_{x,y} \times n_{z,x}\rangle^2 \cos(d(x,y))}{\sin(d(x,y))} G_3(d(x,y)),\end{aligned}$$

which shows

$$|(HX_n^y J_n(x,y))_{22}| \leq \frac{117 \cdot \pi^4}{(n+1)d(x,y)^4}.$$

Moreover, we have $(HX_n^y J_n(x,-x))_{22} = 0$ and using Lemma 5.2 gives

$$\begin{aligned}|HX_n^y J_n(x,z)| &= \left| \frac{\langle x, \xi \rangle}{\sin(d(x,z))} \left(\frac{\tilde{J}_N'(d(x,z)) - \cos(d(x,z))\sin(d(x,z)))\tilde{J}_N''(d(x,z))}{\sin^2(d(x,z))} \right) \right|, \\ &\leq 0.52 \cdot \frac{1}{30} n(n+2)(9n(n+2)-2)d(x,z) \leq \frac{\delta}{5}(n+1)^3.\end{aligned}$$

Lastly, the off-diagonal entries are computed as

$$\begin{aligned}(HX_n^y J_n(x,y))_{ij} = {} & \langle x, \langle n_{x,y}, n_{z,x}\rangle (n_{x,\xi} \times n_{z,x})\rangle \sin(d(x,\xi)) G_3(d(x,y)) \\ & + \langle x, \langle n_{x,\xi}, n_{z,x}\rangle (n_{x,y} \times n_{z,x})\rangle \sin(d(x,\xi)) G_3(d(x,y)) \\ & - \frac{\langle x, \xi \rangle \langle x, n_{x,y} \times n_{z,x}\rangle \langle n_{x,y}, n_{z,x}\rangle}{\sin(d(x,y))} \left(\tilde{J}_N'''(d(x,y)) + \tilde{J}_N'(d(x,y)) \right) \\ & + \frac{3\langle x, \xi \rangle \langle x, n_{x,y} \times n_{z,x}\rangle \langle n_{x,y}, n_{z,x}\rangle \cos(d(x,y))}{\sin(d(x,y))} G_3(d(x,y)),\end{aligned}$$

which results in

$$|(HX_n^y J_n(x,y))_{ij}| \leq \frac{109 \cdot \pi^4}{(n+1)d(x,y)^4},$$

$(HX_n^y J_n(x,-x))_{ij} = 0$ and

$$|(HX_n^y J_n(x,z))_{ij}| = |\langle x, n_{x,\xi} \times n_{z,x} \rangle \sin(d(x,\xi)) G_3(d(x,y))|,$$
$$\leq \frac{\delta}{5}(n+1)^3.$$

□

We finish this section bounding sums of the pointwise expression of the previous theorems. For a discrete set $\mathcal{X} \subset \mathbb{S}^2$, we assume that its minimal separation is bounded from below in the following way

$$\rho(\mathcal{X}) = \min_{x_i, x_j \in \mathcal{X}, x_i \neq x_j} d(x_i, x_j) \geq \frac{\nu}{n+1}.$$

This is the analog of Lemma 2.6 for the case of the rotation group and involves classical ringing arguments on the sphere.

Lemma 5.5. *Let* $x_j \in \mathcal{X}$*, where* $\mathcal{X} \subset \mathbb{S}^2$ *is a discrete set, which obeys a separation condition of* $\rho(\mathcal{X}) \geq \frac{\nu}{n+1}$ *with* $\nu > 0$*. Let* $x \in \mathbb{S}^2$ *such that* $d(x, x_j) \leq \varepsilon \frac{\nu}{n+1}$*, for* $0 \leq \varepsilon \leq 1/2$*. Suppose* $f : \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{C}$ *obeys,*

$$|f(x,y)| \leq \frac{c_f}{((n+1) \cdot d(x,y))^s}, \tag{5.9}$$

for $x \neq y$ *and* $s \geq 3$*, then*

$$\sum_{x_i \in \mathcal{X} \setminus x_j} |f(x, x_i)| \leq \frac{c_f a_\varepsilon}{\nu^s}, \tag{5.10}$$

where $a_\varepsilon = \zeta(s-1) \cdot \min\{9 \cdot (1-\varepsilon)^{-s} + 25, 25 \cdot (1-\varepsilon)^{-s}\}$*. Here* ζ *denotes the* Riemannian Zeta function.

Proof. The proof is very similar to that of Lemma 2.6. We use a ringing argument on the sphere. More concrete, for $x \in \mathbb{S}^2$, with $d(x, x_j) \leq \varepsilon \frac{\nu}{n+1}$ for some $x_j \in \mathcal{X}$, we define the ring about x by

$$\mathcal{S}_m := \left\{ y \in \mathbb{S}^2 : \frac{\nu m}{n+1} \leq d(x,y) < \frac{\nu(m+1)}{n+1} \right\},$$

for $m \in \mathbb{N}$. As shown in [Keiner et al., 2007, Lemma 5], we can estimate the number of elements in the intersection of $\mathcal{S}_m$ with the set $\mathcal{X} \setminus \{x_j\}$ for $m \geq 1$ by

$$\text{card}(\mathcal{X} \setminus \{x_j\} \cap \mathcal{S}_m) \leq 25m.$$

Thus, it remains to estimate the number of elements in $\mathcal{X} \setminus \{x_j\} \cap \mathcal{S}_0$. We will use the same argument as in [Keiner et al., 2007, Lemma 5]. Observe, that $B_{\frac{\nu}{2(n+1)}}(x_i) \cap B_{\frac{\nu}{2(n+1)}}(x_n) = \emptyset$ for $x_i, x_n \in \mathcal{X} \setminus \{x_j\} \cap \mathcal{S}_0$ and

$$\bigcup_{x_i \in \mathcal{X} \setminus \{x_j\} \cap \mathcal{S}_0} B_{\frac{\nu}{2(n+1)}}(x_i) \subseteq B_{\frac{3\nu}{2(n+1)}}(x).$$

Since $\varepsilon \leq 1/2$ and the Riemannian volume form is rotation invariant, we can bound the number of elements by

$$\text{card}(\mathcal{X} \setminus \{x_j\} \cap \mathcal{S}_0) \leq \frac{\Omega\left(B_{\frac{3\nu}{2(n+1)}}(e_3)\right)}{\Omega\left(B_{\frac{\nu}{2(n+1)}}(e_3)\right)},$$

where $e_3 = (0,0,1)^T$ is the north pole on the sphere. In polar coordinates around e_3, we consequently have the bound

$$\begin{aligned}\operatorname{card}(\mathcal{X}\setminus\{x_j\}\cap\mathcal{S}_0) &\le \frac{\Omega\left(B_{\frac{3\nu}{2(n+1)}}(e_3)\right)}{\Omega\left(B_{\frac{\nu}{2(n+1)}}(e_3)\right)} = \frac{\int_0^{\frac{3\nu}{2(n+1)}}\sin(r)\mathrm{d}r}{\int_0^{\frac{\nu}{2(n+1)}}\sin(r)\mathrm{d}r},\\ &= \frac{1-\cos\left(\frac{3\nu}{2(n+1)}\right)}{1-\cos\left(\frac{\nu}{2(n+1)}\right)} = \left(1+2\cos\left(\frac{\nu}{2(n+1)}\right)\right)^2 \le 9.\end{aligned}$$

Because $d(x,x_j)\le\varepsilon\frac{\nu}{n+1}$, we have $d(x,x_i)\ge\frac{(1-\varepsilon)\nu}{n+1}$ for $x_i\in\mathcal{X}\setminus\{x_j\}\cap\mathcal{S}_0$. Using this and the locality assumption (2.3), we can therefore estimate

$$\begin{aligned}\sum_{x_i\in\mathcal{X}\setminus x_j}|f(x,x_i)| &\le \sum_{x_i\in(\mathcal{X}\setminus x_j)\cap\mathcal{S}_0}\frac{c_f}{((n+1)\cdot d(x,x_i))^s}+\sum_{m=1}^{\infty}\sum_{x_i\in(\mathcal{X}\setminus x_j)\cap\mathcal{S}_m}\frac{c_f}{((n+1)\cdot d(x,x_i))^s},\\ &\le \frac{9c_f(1-\varepsilon)^{-s}}{\nu^s}+25c_f\sum_{m=1}^{\infty}\frac{m}{(m\nu)^s},\\ &\le \frac{9c_f(1-\varepsilon)^{-s}}{\nu^s}+\frac{25c_f}{\nu^s}\sum_{m=1}^{\infty}\frac{1}{m^{s-1}},\\ &\le \frac{(9(1-\varepsilon)^{-s}+25)c_f\zeta(s-1)}{\nu^s},\end{aligned}$$

where the last inequality follows by the definition of the Zeta function. On the other hand, we can define the rings around x_j again by

$$\tilde{\mathcal{S}}_m := \{y\in\mathbb{S}^2 : \frac{(1-\varepsilon)\nu m}{n+1}\le d(x_j,y)\le\frac{(1-\varepsilon)\nu(m+1)}{n+1}\}.$$

Since $d(x,x_j)\le\varepsilon\frac{\nu}{n+1}$, we have $d(x,x_j)\le\varepsilon d(x_i,x_j)$ for $x_i\in(\mathcal{X}\setminus x_j)\cap\tilde{\mathcal{S}}_m$ and therefore $d(x,x_i)\ge d(x_i,x_j)-d(x,x_j)\ge\frac{(1-\varepsilon)\nu m}{n+1}$. Using this and the locality assumption (2.3), we can estimate for $s\ge 3$

$$\begin{aligned}\sum_{x_i\in\mathcal{X}\setminus x_j}|f(x,x_i)| &\le \sum_{m=1}^{\infty}\sum_{x_i\in(\mathcal{X}\setminus x_j)\cap\tilde{\mathcal{S}}_m}\frac{c_f}{((n+1)d(x,x_i))^s},\\ &\le 25c_f\sum_{m=1}^{\infty}\frac{m}{(1-\varepsilon)^s(m\nu)^s},\\ &\le \frac{25c_f}{(1-\varepsilon)^s\nu^s}\sum_{m=1}^{\infty}\frac{1}{m^{s-1}} = \frac{25c_f\zeta(s-1)}{(1-\varepsilon)^s\nu^s}.\end{aligned}$$

□

Having established the necessary localization estimates of the Jackson kernel and its derivatives in this section, we are now able to construct and validate a dual certificate using the Hermite interpolation in the next section.

5.3 Dual Certificate on the Sphere

In this section, we construct a dual certificate as the solution of the Hermite type interpolation problem. Remember, we would like to solve the interpolation problem

$$q(x_i) = u_i,$$
$$X_1 q(x_i) = X_2 q(x_i) = 0,$$

for $x_i \in \mathcal{X}$, where q needs to be an element of $\Pi_N(\mathbb{S}^2)$. The interpolant, we construct, is of the form

$$q(x) = \sum_i a_{0,i} J_N(x, x_i) + a_{1,i} X_1^y J_N(x, x_i) + a_{2,i} X_2^y J_N(x, x_i) \in \Pi_N(\mathbb{S}^2).$$

Thus, the coefficients should satisfy

$$K\alpha := \begin{pmatrix} J_N & X_1^x J_N & X_2^x J_N \\ X_1^y J_N & X_1^x X_1^y J_N & X_2^x X_1^y J_N \\ X_2^y J_N & X_1^x X_2^y J_N & X_2^x X_2^y J_N \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} u \\ 0 \\ 0 \end{pmatrix}, \tag{5.11}$$

where again the matrix K consists of blocks of the form

$$X_i^x X_j^y J_N = (X_i^x X_j^y J_N(x_k, x_m))_{k,m}.$$

For abbreviation, we write

$$J_{ij} = X_i^x X_j^y J_N.$$

We will assume, that the interpolation points $\mathcal{X} = \{x_m\}_m$ obey a minimal separation distance of the form

$$\rho(\mathcal{X}) = \min_{x_k \neq x_m} d(x_k, x_m) \geq \frac{\nu}{n+1}, \tag{5.12}$$

where again $n = \lfloor N/2 \rfloor$.

In order to show the existence of q, we need to show the invertibility of the matrix K. Even more, we need to partially compute its inverse to derive bounds on the coefficients α. The following Lemma is the counterpart of Lemma 3.1 in the case of the rotation group and bounds the entries of the interpolation matrix.

Lemma 5.6. *Suppose the points satisfy the separation condition* (5.12). *Then the entries of the interpolation matrix are bounded in the following way*

$$\|I - J_{00}\|_\infty \leq \frac{C_0}{\nu^4},$$
$$\|J_{0i}\|_\infty, \|J_{i0}\|_\infty \leq \frac{C_1(n+1)}{\nu^4}, \quad \|J_{ij}\|_\infty \leq \frac{C_2(n+1)^2}{\nu^4}, \quad \text{for } i \neq j, i, j \neq 0,$$
$$\left\| -\tilde{J}_N''(0) I - J_{ii} \right\|_\infty \leq \frac{C_2(n+1)^2}{\nu^4},$$

with

$$C_0 = 25 \cdot \zeta(3) \cdot \pi^4, \quad C_1 = 75 \cdot \zeta(3) \cdot \pi^4, \quad C_2 = 412.5 \cdot \zeta(3) \cdot \pi^4.$$

If $n \geq 9$ *and* $\nu^4 > \frac{3}{0.99} \cdot C_2$, *then*

$$\|J_{00}^{-1}\|_\infty \leq \frac{1}{1 - \frac{C_0}{\nu^4}}, \quad \|J_{ii}^{-1}\|_\infty \leq \frac{3}{0.99(n+1)^2 \left(1 - \frac{3C_2}{0.99\nu^4}\right)}.$$

Proof. The bounds follow directly from Theorem 5.3 and Lemma 5.5, as well as from the invertibility of a matrix A, if

$$\|I - A\|_\infty < 1,$$

and the bound on the norm of the inverse given by

$$\|A^{-1}\|_\infty \leq \frac{1}{1 - \|I - A\|_\infty}.$$

For the bound on J_{ii}^{-1}, one uses in addition, that for $n \geq 9$, $|\tilde{J}_N''(0)| = \frac{n(n+2)}{3} \geq \frac{0.99\cdot(n+1)^2}{3}$. □

With this preparation, we state the main Theorem of this section, which gives a condition on the separation of the interpolation points to guarantee the invertibility of the interpolation matrix K and gives bounds on the coefficients α.

Theorem 5.7. *Suppose the separation condition* (5.12) *is satisfied for some $\nu \geq 0$, such that there is a constant $b \geq 3$, with*

$$\nu^4 \geq \frac{3}{0.99} \cdot b \cdot C_2, \tag{5.13}$$

where the constant C_2 is given in Lemma 5.6. Then the matrix K is invertible and the coefficients of the linear system (5.11) *satisfy*

$$\|\alpha_0\|_\infty \leq 1 + \frac{1}{45(b-2)-1},$$
$$\|\alpha_j\|_\infty \leq \frac{(n+1)^{-1}}{4.5(b-2)-0.1}, \quad j = 1, 2.$$

Moreover, we have the lower bound

$$|\alpha_{0,i}| \geq 1 - \frac{1}{45(b-2)-1}.$$

Proof. The proof is in the same line as that of Theorem 3.2. We partition the matrix K into blocks of the form

$$K = \begin{pmatrix} K_0 & \tilde{K}_1 \\ K_1 & K_2 \end{pmatrix},$$

with blocks given by

$$\begin{aligned} K_0 &= J_{00} = J_N, \\ K_1 &= \begin{bmatrix} J_{01} & J_{02} \end{bmatrix}^T = \begin{bmatrix} X_1^y J_N & X_2^y J_N \end{bmatrix}^T, \\ \tilde{K}_1 &= \begin{bmatrix} J_{10} & J_{20} \end{bmatrix} = \begin{bmatrix} X_1^x J_N & X_2^x J_N \end{bmatrix}, \\ K_2 &= \begin{bmatrix} J_{11} & J_{21} \\ J_{12} & J_{22} \end{bmatrix} = \begin{bmatrix} X_1^x X_1^y J_N & X_2^x X_1^y J_N \\ X_1^x X_2^y J_N & X_2^x X_2^y J_N \end{bmatrix} \end{aligned}$$

and use an iterative block inversion. Again, we set for abbreviation

$$a_1 := \frac{3 \cdot C_2}{0.99\nu^4}.$$

It represents the quotient of the off-diagonal upper bound and the on-diagonal lower bound. The assumption of the theorem then reads as

$$a_1 \leq \frac{1}{b},$$

with $b \geq 3$. The same arguments, as those in the proof of Theorem 3.2 show

$$\left\| I - \frac{K_2/J_{22}}{\tilde{J}_N''(0)} \right\|_\infty \leq \frac{1}{(b-1)},$$

and therefore

$$\| \left(K_2/J_{22}\right)^{-1} \|_\infty \leq \frac{3(b-1)}{0.99(n+1)^2(b-2)},$$

which shows the invertibility of the matrix K_2. With the abbreviation $T = K_2/J_{22}$, we have the representation

$$K_2^{-1} = \begin{pmatrix} T^{-1} & -T^{-1}J_{21}\left(J_{22}\right)^{-1} \\ -\left(J_{22}\right)^{-1}J_{12}T^{-1} & \left(J_{22}\right)^{-1} + \left(J_{22}\right)^{-1}J_{12}T^{-1}J_{21}\left(J_{22}\right)^{-1} \end{pmatrix}. \tag{5.14}$$

In the next step, one can show, that

$$\|I - K/K_2\|_\infty \leq \frac{1}{45(b-2)},$$

which yields

$$\| \left(K/K_2\right)^{-1} \|_\infty \leq \frac{1}{1 - \frac{1}{45(b-2)}}$$

and the invertibility of K. Thus, using the representation (5.14) of the inverse of K_2 and the abbreviation $S = K/K_2$, one can calculate the solution of the interpolation problem as

$$\begin{aligned} \alpha_0 &= S^{-1}u, \\ \alpha_1 &= -T^{-1}(J_{01} - J_{21}(J_{22})^{-1}J_{02})\alpha_0, \\ \alpha_2 &= -J_{22}^{-1}(J_{12}\alpha_1 + J_{02}\alpha_0). \end{aligned}$$

This yields the bounds

$$\begin{aligned} \|\alpha_0\|_\infty &\leq 1 + \frac{1}{45(b-2)-1}, \\ \|\alpha_1\|_\infty &\leq \frac{(n+1)^{-1}}{4.5(b-2)-0.1}, \\ \|\alpha_2\|_\infty &\leq \frac{(n+1)^{-1}}{4.5(b-2)-0.1}. \end{aligned}$$

Moreover, we have the estimate

$$\begin{aligned} |\alpha_{0,i}| &= \left| \left(\left(I - \left(I - \left(K/K_2\right)^{-1} \right) \right) u \right)_i \right|, \\ &\geq \left| 1 - \left| \left(\left(I - \left(K/K_2\right)^{-1} \right) u \right)_i \right| \right|, \\ &\geq \left| 1 - \left| \left(I - K/K_2\right) \left(K/K_2\right)^{-1} \right| \right|. \end{aligned}$$

Since $b \geq 3$, we have

$$\left|(I - K/K_2)\,(K/K_2)^{-1}\right| \leq \|I - K/K_2\|_\infty \|\,(K/K_2)^{-1}\,\|_\infty, \\ \leq \frac{1}{45(b-2)-1} < 1,$$

and therefore

$$|\alpha_{0,i}| \geq 1 - \frac{1}{45(b-2)-1}.$$

□

Corollary 5.8. *Suppose the interpolation points $\mathcal{X} = \{x_i\}$ obey a separation distance of*

$$\rho(\mathcal{X}) \geq \frac{20\pi}{N}, \tag{5.15}$$

for $N \geq 20$. Then the interpolation problem (5.11) *admits a unique solution, such that*

$$\|\alpha_0\|_\infty \leq 1 + \frac{1}{45 \cdot 6 - 1} \leq 1 + 3.8 \cdot 10^{-3}, \\ \|\alpha_j\|_\infty \leq \frac{(n+1)^{-1}}{4.5 \cdot 6 - 0.1} \leq \frac{3.8 \cdot 10^{-2}}{(n+1)}, \quad j = 1, 2,$$

and we have the lower bound

$$|\alpha_{0,i}| \geq 1 - \frac{1}{45 \cdot 6 - 1} \geq 1 - 3.8 \cdot 10^{-3}.$$

Proof. For $n = \lfloor N/2 \rfloor$, we have $(n+1) \geq N/2$ and therefore

$$\rho(\mathcal{X}) \geq \frac{10\pi}{(n+1)}.$$

One shows, that with $\nu = 10\pi$, we have

$$\nu^4 \geq \frac{3}{0.99} \cdot b \cdot C_2,$$

with $b = 8$ and Theorem 5.7 yields the assertion. □

Using the derived bounds on the coefficients of the interpolant, we proceed by showing the upper bound in absolute value of the interpolating function q of Corollary 5.8, i.e.

$$|q(x)| < 1,$$

whenever x is not an interpolation point. Again, we split the argument into two parts. Points, which are close to an interpolation point, are covered by Lemma 5.9, where the argument involves convexity arguments via the definiteness of the Hessian. The bound for points that are well separated from any interpolation point is the content of Lemma 5.10.

Lemma 5.9. *Suppose the interpolation points $\mathcal{X}$ obey a separation condition of the form* (5.15) *and $x \in \mathbb{S}^2$ satisfies $0 < d(x, x_i) \leq \frac{0.5}{(n+1)}$, for an interpolation point $x_i \in \mathcal{X}$. Then the interpolating function q of Corollary 5.8 fulfills*

$$|q(x)| < 1.$$

Proof. Remember, the interpolant is of the form

$$q(x) = \sum_j a_{0,j} J_N(x, x_j) + a_{1,i} X_1^y J_N(x, x_j) + a_{2,j} X_2^y J_N(x, x_j).$$

Without loss of generality, we assume that $u_i = 1$ and we have to show that the Hessian is negative definite. The case $u_i = -1$ is completely analog with changing signs. At the interpolation point x_i in normal coordinates at x_i, the Hessian has the form

$$\tilde{H} q(x_i) = \begin{pmatrix} X_1 X_1 q(x_i) & X_1 X_2 q(x_i) \\ X_2 X_1 q(x_i) & X_2 X_2 q(x_i) \end{pmatrix}.$$

We have, using the bounds of Theorem 5.3, Lemma 5.5 and Corollary 5.8

$$\begin{aligned} X_n X_n q(x_i) &= \sum_j a_{0,j} X_n^x X_n^x J_N(x_i, x_j) + a_{1,j} X_n^x X_n^x X_1^y J_N(x_i, x_j) + a_{2,j} X_n^x X_n^x X_2^y J_N(x_i, x_j), \\ &= a_{0,i} X_n^x X_n^x J_N(x_i, x_i) + \sum_{j \neq i} \Big(a_{0,j} X_n^x X_n^x J_N(x_i, x_j) + a_{1,i} X_n^x X_n^x X_1^y J_N(x_i, x_j) \\ &\quad + a_{2,j} X_n^x X_n^x X_2^y J_N(x_i, x_j) \Big), \\ &\leq -\left(1 - \frac{1}{45 \cdot 6 - 1}\right) \frac{n(n+2)}{3} + \left(1 + \frac{1}{45 \cdot 6 - 1}\right) \frac{25\zeta(3) \cdot 16.5}{10^4} (n+1)^2 \\ &\quad + \frac{2}{4.5 \cdot 6 - 0.1} \frac{25\zeta(3) \cdot 110}{10^4} (n+1)^2. \end{aligned}$$

Since $N \geq 20$, we have $(n+1) \geq 10$ and therefore

$$\frac{n(n+2)}{3} \geq \frac{n(n+2)}{(n+1)^2} \cdot \frac{(n+1)^2}{3} \geq 0.99 \cdot \frac{(n+1)^2}{3}.$$

This yields

$$X_n X_n q(x_i) \leq -0.25 \cdot (n+1)^2.$$

Similarly, we can bound the off-diagonal entries by

$$|X_k X_n q(x_i)| \leq 0.075 \cdot (n+1)^2.$$

Combining these to bounds, we have

$$\operatorname{trace}(\tilde{H} q(x_i)) \leq -0.5 \cdot (n+1)^2, \quad \det(\tilde{H} q(x_i)) \geq 0.05 \cdot (n+1)^2,$$

which means that the Hessian at x_i is strictly negative definite and x_i is an isolated local maximal point of q and $q(x_i) = 1$ is a local maximum.

For $x \neq x_i$, $d(x, x_i) \leq \frac{0.5}{n+1}$ we argue in the same way, but instead of the bounds of Theorem 5.3 we use the bounds derived in Theorem 5.4. Thus, for the diagonal entries we get

$$(Hq(x))_{ii} = \sum_j a_{0,j} (H J_N(x, x_j))_{ii} + a_{1,i} (H X_1^y J_N(x, x_j))_{ii} + a_{2,j} (H X_2^y J_N(x, x_j))_{ii},$$

$$
\begin{aligned}
&\le a_{0,i}\tilde{J}_N''(0) + \|\alpha_0\|_\infty \left(|\tilde{J}_N''(0) - (HJ_N(x,x_i))_{ii}| + \sum_{j\neq i} |(HJ_N(x,x_j))_{ii}| \right) \\
&\quad + \|\alpha_1\|_\infty (|(HX_1^y J_N(x,x_j))_{ii}| + \sum_{j\neq i} |(HX_1^y J_N(x,x_j))_{ii}|) \\
&\quad + \|\alpha_2\|_\infty (|(HX_2^y J_N(x,x_i))_{ii}| + \sum_{j\neq i} |(HX_2^y J_N(x,x_j))_{ii}|), \\
&\le -0.19\cdot(n+1)^2
\end{aligned}
$$

and in the same way for the off-diagonal entries

$$|(Hq(x))_{ij}| \le 0.13(n+1)^2.$$

This shows

$$\operatorname{trace}(Hq(x_i)) \le -0.3\cdot(n+1)^2, \quad \det(Hq(x_i)) \ge 0.01\cdot(n+1)^2.$$

Hence, the function q is strictly concave on $B_{0.5/(n+1)}(x_i)\setminus\{x_i\}$, which shows $q(x) < 1$. Moreover, the Taylor expansion of the cosine and the sine function shows

$$
\begin{aligned}
J_N(x,x_i) &\ge 1 + \frac{\tilde{J}_N''(0)}{2} d(x,x_i)^2 \ge 1 - \frac{(n+1)^2}{6} d(x,x_i)^2, \\
|X_n^y J_N(x,x_i)| &\le |\tilde{J}_N'(d(x,x_i))| \le |\tilde{J}_N''(0)| d(x,x_i) \le \frac{(n+1)^2}{3} d(x,x_i),
\end{aligned}
$$

meaning for $d(x,x_i) \le \frac{0.5}{(n+1)}$

$$
\begin{aligned}
q(x) &\ge \alpha_{0,i} J(x,x_i) - \|\alpha_1\|_\infty |X_1^y J_N(x,x_i)| + \|\alpha_2\|_\infty |X_2^y J_N(x,x_i)| \\
&\quad + \sum_{x_j\neq x_i} \|\alpha_0\|_\infty |J_N(x,x_j)| + \|\alpha_1\|_\infty |X_1^y J_N(x,x_j)| + \|\alpha_2\|_\infty |X_2^y J_N(x,x_j)|, \\
&\ge 0.93.
\end{aligned}
$$

Combining this shows $0.93 \le q(x) < 1$. In the case $q(x_i) = -1$, the analog arguments with changing signs show, that x_i is an isolated local minimal point, q is strictly convex on $B_{0.5/(n+1)}(x_i)$ and $-1 < q(x) \le 0.93$. □

Lemma 5.10. *Under the assumptions of Lemma 5.9, we have that for all $x \in \mathbb{S}^2$ with $d(x,x_m) \ge \frac{0.5}{(n+1)}$ for all $x_m \in \mathcal{X}$ the interpolating function q of Corollary 5.8 fulfills*

$$|q(x)| < 1.$$

Proof. We split the proof into three cases. The first case corresponds to those $x \in \mathbb{S}^2$, such that $\frac{0.5}{(n+1)} \le d(x,x_m) \le \frac{1.1\cdot\pi}{(n+1)}$, the second to $\frac{1.1\cdot\pi}{(n+1)} \le d(x,x_m) \le \frac{5\cdot\pi}{(n+1)}$ and the last to those $x \in \mathbb{S}^2$, such that $d(x,x_j) > \frac{5\cdot\pi}{(n+1)}$ for all interpolation points x_i. We have to bound the summands in

$$
\begin{aligned}
|q(x)| &= \sum_j a_{0,j} J_N(x,x_j) + a_{1,i} X_1^y J_N(x,x_j) + a_{2,j} X_2^y J_N(x,x_j)., \\
&\le \|\alpha_0\|_\infty |J_N(x,x_m)| + \|\alpha_1\|_\infty |X_1^y J_N(x,x_m)| + \|\alpha_2\|_\infty |X_2^y J_N(x,x_m)| \\
&\quad + \sum_{x_j\neq x_m} \|\alpha_0\|_\infty |J_N(x,x_j)| + \|\alpha_1\|_\infty |X_1^y J_N(x,x_j)| + \|\alpha_2\|_\infty |X_2^y J_N(x,x_j)|.
\end{aligned}
$$

In the first case, we have due to the Taylor expansion of the cosine function and the positivity of the Jackson kernel

$$|J_N(x, x_m)| = \tilde{J}_N(d(x, x_m)) \leq 1 - \frac{|\tilde{J}_N''(0)|}{2} d(x, x_m)^2 + \frac{|\tilde{J}_N^{(4)}(0)|}{24} d(x, x_m)^4.$$

Since $N \geq 20$, we have $(n+1) \geq 10$ and consequently

$$|\tilde{J}_N''(0)| = \frac{n(n+2)}{3} \geq \frac{n(n+2)}{(n+1)^2} \cdot \frac{(n+1)^2}{3} \geq 0.99 \cdot \frac{(n+1)^2}{3}.$$

In addition, we have

$$|\tilde{J}_N^{(4)}(0)| = \frac{1}{30} n(n+2)(9n(n+2) - 2) \leq \frac{3}{10}(n+1)^4.$$

Thus, for $t \in [0.5, 1.1\pi]$, we have the bound

$$\tilde{J}_N(t/(n+1)) \leq 1 - \frac{0.99}{6} t^2 + \frac{1}{80} t^4.$$

The polynomial on the righthand side is monotonic decreasing for $t \in [0.5, t_0]$, for $t_0 = \sqrt{\frac{20 \cdot 0.99}{3}}$ and monotonic increasing for $t = [t_0, 1.1\pi]$. Similarly, we have due to $|\sin(k\omega)| \leq k\omega$,

$$|X_n^y J_N(x, x_m)| \leq |\tilde{J}_N'(d(x, x_m))| \leq |\tilde{J}_N''(0)| d(x, x_m) \leq \frac{(n+1)^2}{3} d(x, x_m).$$

Accordingly, for $\frac{0.5}{(n+1)} \leq d(x, x_m) \leq \frac{1.1\pi}{(n+1)}$, we can estimate using the bounds above and the estimates of Theorem 5.3, Lemma 5.5 and Corollary 5.8

$$|q(x)| \leq \left(1 + \frac{1}{45 \cdot 6 - 1}\right)\left(1 - \frac{0.99}{6} t^2 + \frac{1}{80} t^4\right) + \frac{2}{4.5 \cdot 6 - 0.1} \cdot \frac{t}{3}$$
$$\left(1 + \frac{1}{45 \cdot 6 - 1}\right) \frac{a_{t/10\pi}}{10^4} + \frac{2}{4.5 \cdot 6 - 0.1} \frac{3a_{t/10\pi}}{10^4},$$

where $t = d(x, x_m)(n+1)$. This results in the bounds

$$|q(x)| \leq \begin{cases} 0.993, & \frac{0.5}{(n+1)} \leq d(x, x_m) \leq \frac{1}{(n+1)}, \\ 0.92, & \frac{1}{(n+1)} \leq d(x, x_m) \leq \frac{t_0}{(n+1)}, \\ 0.91, & \frac{t_0}{(n+1)} \leq d(x, x_m) \leq \frac{1.1\pi}{(n+1)}, \end{cases}$$

which complete the first case. For the second case, i.e. $\frac{1.1 \cdot \pi}{(n+1)} \leq d(x, x_m) \leq \frac{5 \cdot \pi}{(n+1)}$, we use the bounds of Theorem 5.3 to derive

$$|J(x, x_m)| \leq \frac{\pi^4}{(n+1)^4 d(x, x_m)^4} \leq \frac{\pi^4}{t^4},$$
$$|X_n^y J(x, x_m)| \leq \frac{3\pi^4}{(n+1)^4 d(x, x_m)^4} \leq \frac{3\pi^4}{t^4},$$

with $t = d(x, x_m)(n+1)$. For this reason, we have the estimate

$$|q(x)| \leq \left(1 + \frac{1}{45 \cdot 6 - 1}\right) \frac{\pi^4}{t^4} + \frac{2}{4.5 \cdot 6 - 0.1} \cdot \frac{3\pi^4}{t^4}$$
$$\left(1 + \frac{1}{45 \cdot 6 - 1}\right) \frac{a_{t/10\pi}}{10^4} + \frac{2}{4.5 \cdot 6 - 0.1} \frac{3a_{t/10\pi}}{10^4},$$

which shows for $\frac{1.1 \cdot \pi}{(n+1)} \leq d(x, x_m) \leq \frac{5 \cdot \pi}{(n+1)}$

$$|q(x)| \leq 0.87.$$

Lastly, in the case $d(x, x_j) > \frac{5\pi}{n+1}$ for all interpolation points x_j, the set $\mathcal{X} \cup x$ obeys a separation distance of $\frac{5\pi}{n+1}$ and we again use the bounds of Theorem 5.3, Lemma 5.5 and Corollary 5.8 to estimate

$$\begin{aligned} |q(x)| &\leq \|a_0\|_\infty \sum_j |J(x, x_j)| + \|a_1\|_\infty \sum_j |X_1^y J(x, x_j)| + \|a_2\|_\infty \sum_j |X_2^y J(x, x_j)|, \\ &\leq \left(1 + \frac{1}{45 \cdot 6 - 1}\right) \frac{25 \cdot \zeta(3)}{5^4} + \frac{2}{4.5 \cdot 6 - 0.1} \cdot \frac{75 \cdot \zeta(3)}{5^4} \leq 0.06. \end{aligned}$$

□

Combining Corollary 5.8, Lemma 5.9 and Lemma 5.10, shows the existence of a dual certificate. We summarize this result in the following theorem.

Theorem 5.11. *Suppose the points $\mathcal{X} = \{x_1, \dots, x_M\}$ obey a separation distance of $\rho(\mathcal{X}) \geq \frac{20\pi}{N}$ for $N \geq 20$. Then for each sign combination $u_i \in \{-1, 1\}$, there is a $q \in \Pi_N(\mathbb{S}^2)$ such that*

$$\begin{aligned} q(x_i) &= u_i, \quad \text{for } x_i \in \mathcal{X}, \\ |q(x)| &< 1, \quad \text{for } x \in \mathbb{S}^2 \setminus \mathcal{X}. \end{aligned}$$

The existence of a dual certificate immediately yields the recovery of the sought measure via the minimization of the total variation.

Corollary 5.12. *Suppose the support of the signed measure $\mu^\star$ obeys the separation condition*

$$\min_{x \neq y} d(x, y) \geq \frac{20\pi}{N}, \quad x, y \in \operatorname{supp}(\mu^\star),$$

for $N \geq 20$. Then $\mu^\star$ is the unique real solution of the minimization problem

$$\min_{\mu \in \mathcal{M}(\mathbb{S}^2, \mathbb{R})} \|\mu\|_{TV}, \quad \text{subject to} \quad \mathcal{P}_N^* \mu = \mathcal{P}_N^* \mu^\star. \tag{SP}$$

Proof. Theorem 5.11 guarantuees the existence of a dual certificate. Hence, by Theorem 1.7, $\mathcal{P}_N^*$ has the null-space property with respect to $\operatorname{supp}(\mu^\star)$ and Theorem 1.6 shows that $\mu^\star$ is the unique real solution. □

Notes and References. *Whereas the problem of super-resolution on the sphere was first considered in [Bendory et al., 2015a], we would like to mention, that to the best of our knowledge the first valid proof for exact recovery in the context of total variation minimization is given in this thesis. We briefly state the necessary modifications.*

The choice of the Jackson kernel as interpolation kernel has two advantages. First, the closed-form representation (5.4) *of the kernel allows for asymptotic estimates with* explicit constants*, see Theorem*

5.3. Second, the behavior of the derivatives near the diagonal can be controlled efficiently, see Theorem 5.4. Both ingredients are necessary *for showing the bound $|q(x)| < 1$, whenever x is not an interpolation point. This is in contrast to [Bendory et al., 2015a], where the use of an unspecified polynomial kernel prohibits those explicit estimates.*

In addition, the geometry of the sphere, which plays a crucial role in bounding the Hessian, was not considered properly in [Bendory et al., 2015a]. Both gaps have been closed in this chapter.

Numerical considerations regarding the super-resolution problem on the sphere were stated in [Bendory et al., 2015b]. The proposed algorithm uses a semi-definite relaxation of the dual problem using the Bounded Real Lemma and is similar to Algorithm 1 for the case of the rotation group. The authors observe numerically an exact recovery, whenever the support points of the sought measure obey a separation distance of $\frac{2.5 \cdot \pi}{N}$, which suggests that the constant $20 \cdot \pi$ of Theorem 5.11 is not optimal.

Nevertheless, the investigation of the numerical aspects of the super-resolution problem on the sphere, especially regarding applications, has to be considered in more detail. We leave this for future research.

Appendix A

B-spline Filter

In this section, we derive estimates that involve the perfect B-spline of order s. This function is given by

$$g_{s-1}(x) = \frac{(-1)^{s-1}}{(s-2)!} \int_{-1}^{x} \sum_{k=0}^{s-1} (-1)^k \chi_{(\cos(\frac{k+1}{s}\pi),\cos(\frac{k}{s}\pi)]}(t)(x-t)^{s-2}\mathrm{d}t. \tag{A.1}$$

Proposition A.1. *[Bojanov et al., 1993, Sec. 6.1] We have for $s \in \mathbb{N}$ that the function g_{s-1} given in (A.1) is a spline of order $s-1$ with support $[-1,1]$ and $\|g_{s-1}\|_1 = \frac{1}{(s-1)!2^{s-2}}$. Moreover, we have for all $n \leq s-1$ the identity*

$$g_{s-1}^{(n)}(x) = f_{s-1-n}(x),$$

where

$$f_0(x) = (-1)^{s-1}\operatorname{sign}(U_{s-1}(x)), \quad f_k(x) = \int_{-1}^{x} f_{k-1}(t)\mathrm{d}t, \quad k > 0,$$

with the explicit representation for $k > 0$,

$$f_k(x) = \frac{(-1)^{s-1}}{(k-1)!} \int_{-1}^{x} (x-t)^{k-1}\operatorname{sign}(U_{s-1}(t))\mathrm{d}t. \tag{A.2}$$

Here U_{s-1} denotes the Chebychev polynomial of the second kind of order $s-1$.

Lemma A.2. *Denote by $\tilde{g}_{s-1} = g_{s-1}(2(\cdot))$ the scaled perfect B-spline. Then for $m \in \mathbb{N}$, we have*

$$\|z^{2m}\tilde{g}_{s-1}(z)\|_1 = \frac{(2m)! \cdot s}{4^m \cdot m! \cdot 2^{s+2m-1}(s+m)!}.$$

Proof. By Proposition A.1, we have $g_{s-1}(x) = f_{s-1}(x)$ and integration by parts shows

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} x^{2m}\tilde{g}_{s-1}(x)\mathrm{d}x = \frac{1}{2^{2m+1}} \sum_{l=0}^{2m} (-1)^l \frac{(2m)!}{(2m-l)!} f_{s+l}(1).$$

Now we use a specific orthogonality relation for Chebeychev polynomials of the second type, see e.g. [Bojanov et al., 1993]. For each polynomial p of maximal degree $s-2$ it is true, that

$$\int_{-1}^{1} p(t)\operatorname{sign}(U_{s-1}(t))\mathrm{d}t = 0,$$

and for $m \in \mathbb{N}$

$$\int_{-1}^{1} t^{s+2m} \operatorname{sign}(U_{s-1}(t)) \mathrm{d}t = 0,$$

since the functions $t^{s+2m} \operatorname{sign}(U_{s-1}(t))$ are always odd. Moreover we can calculate

$$\begin{aligned}
\int_{-1}^{1} t^{s+2m-1} \operatorname{sign}(U_{s-1}(t)) \mathrm{d}t &= \int_{-1}^{1} t^{s+2m-1} \sum_{k=0}^{s-1} (-1)^k \chi_{(\cos(\frac{k+1}{s}\pi), \cos(\frac{k+1}{s}\pi)]}(t) \mathrm{d}t, \\
&= \frac{1}{s+2} \sum_{k=0}^{s-1} (-1)^k \left(\cos^{s+2}\left(\frac{k}{s}\pi\right) - \cos^{s+2}\left(\frac{k+1}{s}\pi\right) \right), \\
&= \frac{2}{s+2m} \left(1 + \sum_{k=1}^{s-1} (-1)^k \cos^{s+2m}\left(\frac{k}{s}\pi\right) \right).
\end{aligned}$$

We can write

$$\begin{aligned}
\sum_{k=1}^{s-1} (-1)^k \cos^{s+2}\left(\frac{k}{s}\pi\right) &= \frac{1}{2^{s+2}} \sum_{k=1}^{s-1} (-1)^k \left(\mathrm{e}^{i\pi\frac{k}{s}} + \mathrm{e}^{-i\pi\frac{k}{s}} \right)^{s+2}, \\
&= \frac{1}{2^{s+2}} \sum_{k=1}^{s-1} \mathrm{e}^{ik\pi\frac{s}{s}} \sum_{m=0}^{s+2} \binom{s+2}{m} \mathrm{e}^{i\pi\frac{k(2m-s-2)}{s}}, \\
&= \frac{1}{2^{s+2}} \sum_{m=0}^{s+2} \binom{s+2}{m} \sum_{k=1}^{s-1} \left(\mathrm{e}^{2\pi i \frac{m-1}{s}} \right)^k.
\end{aligned}$$

Since

$$\sum_{k=1}^{s-1} \left(\mathrm{e}^{2\pi i \frac{m-1}{s}} \right)^k = \begin{cases} -1 + \frac{1-\mathrm{e}^{2\pi i (m-1)}}{1-\mathrm{e}^{2\pi i \frac{m-1}{s}}} = -1, & m \notin \{1, s+1\}, \\ s-1, & m \in \{1, s+1\} \end{cases}$$

we have

$$\sum_{k=1}^{s-1} (-1)^k \cos^{s+2}\left(\frac{k}{s}\pi\right) = \frac{s}{2^{s+2m-1}} \binom{s+2m}{m} - 1,$$

which yields

$$\int_{-1}^{1} t^{s+2m-1} \operatorname{sign}(U_{s-1}(t)) \mathrm{d}t = \frac{s \cdot (s+2m-1)!}{2^{s+2m-2} \cdot m! \cdot (s+m)!}.$$

Thus, using the explicit representation (A.2) of the f_k yields

$$\begin{aligned}
f_{s+l}(1) &= \frac{1}{(s-1)! 2^{s-2}} \sum_{r=0}^{\frac{l}{2}} \frac{1}{(l-2r)!} \frac{s!}{4^r \cdot r! \cdot (s+r)!}, \quad l \text{ even}, \\
f_{s+l}(1) &= \frac{1}{(s-1)! 2^{s-2}} \sum_{r=0}^{\frac{l-1}{2}} \frac{1}{(l-2r)!} \frac{s!}{4^r \cdot r! \cdot (s+r)!}, \quad l \text{ odd}.
\end{aligned}$$

Using this, we can derive

$$\| z^{2m} \tilde{g}_{s-1}(z) \|_1 = \frac{(2m)! \cdot s}{4^m \cdot m! \cdot 2^{s+2m-1} (s+m)!}.$$

□

Lemma A.3. *For $s \in 2\mathbb{N}, s \geq 6$ we have*

$$|\tilde{g}_{s-1}^{(s-1)}|_V = 2^s s, \quad \|\tilde{g}_{s-1}^{(s-1)}\|_\infty = 2^{s-1},$$
$$|\tilde{g}_{s-1}^{(s-2)}|_V = 2^{s-1}, \quad \|\tilde{g}_{s-1}^{(s-2)}\|_\infty = 2^{s-2}\tan\left(\frac{\pi}{2s}\right),$$
$$|\tilde{g}_{s-1}^{(s-3)}|_V = 2^{s-4}\tan^2\left(\frac{\pi}{2s}\right)s, \quad \|\tilde{g}_{s-1}^{(s-3)}\|_\infty = 2^{s-4}\tan^2\left(\frac{\pi}{2s}\right),$$
$$|\tilde{g}_{s-1}^{(s-4)}|_V \leq 2^{s-4}\tan^2\left(\frac{\pi}{2s}\right), \quad \|\tilde{g}_{s-1}^{(s-4)}\|_\infty = 2^{s-4}\frac{3\sin^2\left(\frac{\pi}{2s}\right)\tan\left(\frac{\pi}{2s}\right)}{2\cos\left(\frac{\pi}{s}\right)-1},$$
$$|\tilde{g}_{s-1}^{(s-5)}|_V \leq 2^{s-4}\frac{3\sin^2\left(\frac{\pi}{2s}\right)\tan\left(\frac{\pi}{2s}\right)}{2\cos\left(\frac{\pi}{s}\right)-1}.$$

In the case $s = 8$ we have in addition

$$\|\tilde{g}_7^{(j)}\|_\infty \leq \frac{4^j}{2^5(6-j)!}, \quad |\tilde{g}_7^{(j-1)}|_V \leq \frac{4^j}{2^4(6-j)!}, \quad j = 1, 2, 3.$$

Proof. First observe that we have for $0 \leq n \leq 3$

$$|\tilde{g}_{s-1}^{(s-1-n)}|_V = 2^{(s-1-n)}|g_{s-1}^{(s-1-n)}|_V, \quad \|\tilde{g}_{s-1}^{(s-1-n)}\|_\infty = 2^{(s-1-n)}\|g_{s-1}^{(s-1-n)}\|_\infty.$$

By Proposition A.1, we know that $g_{s-1}^{(s-1-n)} = f_n$. For $n = 0$, this means that

$$g_{s-1}^{(s-1)} = f_0(x) = (-1)^{s-1}\operatorname{sign}(U_{s-1}(x))$$

and since $U_{s-1}(x)$ has $s-1$ zeros in the interval $(-1, 1)$ and is not equal to zero for $x = 1, -1$, we have that

$$|g_{s-1}^{(s-1)}| = |f_0|_V = 2s, \quad \|g_{s-1}^{(s-1)}\|_\infty = \|f_0\|_\infty = 1.$$

For $n = 1$, the total variation of f_1 is given by

$$|f_1|_V = \sup\left(\sum_i |f_1(x_{i+1}) - f_1(x_i)|\right) = \sup\left(\sum_i \left|\int_{x_i}^{x_{i+1}} f_0(t)\mathrm{d}t\right|\right) \leq 2,$$

where the supremum is taken over all partitions of $[-1, 1]$. Actually, choosing as partition the sequence of zeros of U_s shows that $|g_{s-1}^{(s-2)}|_V = 2$. Furthermore, we have the representation

$$f_1(x) = \sum_{k=0}^{s-1}(-1)^k \chi_{\left(\cos\left(\frac{(s-k)\pi}{s}\right), \cos\left(\frac{(s-k-1)\pi}{s}\right)\right]}(x) \cdot \left(x - \cos\left(\frac{(s-k)\pi}{s}\right) - \tan\left(\frac{\pi}{2s}\right)\sin\left(\frac{\pi k}{s}\right)\right), \tag{A.3}$$

which shows

$$\|g_{s-1}^{(s-2)}\|_\infty = \|f_1\|_\infty = |f_1(0)| = \tan\left(\frac{\pi}{2s}\right).$$

Since f_2 is continuously differentiable with $f_2' = f_1$, we have

$$|f_2|_V = \int_{-1}^{1}|f_2'(t)|\mathrm{d}t = \int_{-1}^{1}|f_1(t)|\mathrm{d}t.$$

Using the representation (A.3) of f_1, a lengthy calculation shows

$$|g_{s-1}^{(s-3)}|_V = |f_2|_V = \frac{s}{2}\tan^2\left(\frac{\pi}{2s}\right).$$

For $n = 2$, we have the representation

$$\begin{aligned} f_2(x) = \sum_{k=0}^{s-1}(-1)^k \chi_{\left(\cos\left(\frac{(s-k)\pi}{s}\right),\cos\left(\frac{(s-k-1)\pi}{s}\right)\right]}(x) \\ \cdot \frac{1}{2}\left(x^2 + 2\left(\cos\left(\frac{k\pi}{s}\right) - \tan\left(\frac{\pi}{2s}\right)\sin\left(\frac{\pi k}{s}\right)\right)x\right. \\ \left. + \frac{1}{2}\left(1+\cos\left(\frac{2\pi k}{s}\right) - \sin\left(\frac{2\pi k}{s}\right)\tan\left(\frac{\pi}{s}\right)\right)\right). \end{aligned} \tag{A.4}$$

Since the local extrema of f_2 are the zeros of f_1, which are given by

$$\tan\left(\frac{\pi}{2s}\right)\sin\left(\frac{\pi k}{s}\right) - \cos\left(\frac{k\pi}{s}\right),$$

we can calculate the absolute extreme values of f_2 on the intervals $\left(\cos\left(\frac{(s-k)\pi}{s}\right), \cos\left(\frac{(s-k-1)\pi}{s}\right)\right]$ for $k = 0, \ldots, s-1$. A lengthy calculation shows that the absolute values at these points are given by

$$\frac{1}{4}\tan^2\left(\frac{\pi}{2s}\right)\left(1 - \frac{\cos\left(\frac{2k+1}{s}\pi\right)}{\cos\left(\frac{\pi}{s}\right)}\right)$$

which becomes maximal for $k = \frac{s}{2} - 1$. Since

$$|f_3|_V = \int_{-1}^{1} |f_3'(t)|\mathrm{d}t = \int_{-1}^{1} |f_2(t)|\mathrm{d}t,$$

we get immediately $|f_3|_V \le \tan^2\left(\frac{\pi}{2s}\right)$. In the same way we can derive a piecewise representation of f_3

$$\begin{aligned} f_3(x) = \sum_{k=0}^{s-1}(-1)^k \chi_{\left(\cos\left(\frac{(s-k)\pi}{s}\right),\cos\left(\frac{(s-k-1)\pi}{s}\right)\right]}(x) \\ \cdot \frac{1}{2}\left(\frac{x^3}{3} + x^2\left(\cos\left(\frac{k\pi}{s}\right) - \sin\left(\frac{\pi k}{s}\right)\tan\left(\frac{\pi}{2s}\right)\right)\right. \\ + x\left(\cos\left(\frac{k\pi}{s}\right)^2 - \frac{1}{2}\sin\left(\frac{2\pi k}{s}\right)\tan\left(\frac{\pi}{s}\right)\right) \\ \left. + \frac{1}{3}\cos\left(\frac{k\pi}{s}\right)^3 - \frac{1}{4}\sin\left(\frac{k\pi}{s}\right)\tan\left(\frac{\pi}{2s}\right) - \frac{1}{12}\sin\left(\frac{3k\pi}{s}\right)\tan\left(\frac{3\pi}{2s}\right)\right). \end{aligned}$$

and plug in the zeros of f_2, which are given by

$$\tan\left(\frac{\pi}{2s}\right)\sin\left(\frac{\pi k}{s}\right) - \cos\left(\frac{k\pi}{s}\right) + \tan\left(\frac{\pi}{2s}\right)\sqrt{\frac{\sin\left(\frac{\pi k}{s}\right)\sin\left(\frac{\pi(k+1)}{s}\right)}{\cos\left(\frac{\pi}{s}\right)}}, \quad k = 0, \ldots, \frac{s}{2} - 1,$$

$$\tan\left(\frac{\pi}{2s}\right)\sin\left(\frac{\pi k}{s}\right)-\cos\left(\frac{k\pi}{s}\right)-\tan\left(\frac{\pi}{2s}\right)\sqrt{\frac{\sin\left(\frac{\pi k}{s}\right)\sin\left(\frac{\pi(k+1)}{s}\right)}{\cos\left(\frac{\pi}{s}\right)}},\quad k=\frac{s}{2},\dots,s.$$

A lengthy calculation for this shows

$$\|g_{s-1}^{(s-4)}\|_\infty=\|f_3\|_\infty=|f(0)|=\frac{1}{24}\left(\tan\left(\frac{3\pi}{2s}\right)-3\tan\left(\frac{\pi}{2s}\right)\right)=\frac{3\sin^2\left(\frac{\pi}{2s}\right)\tan\left(\frac{\pi}{2s}\right)}{2\cos\left(\frac{\pi}{s}\right)-1},$$

and again $|g_{s-1}^{(s-5)}|_V \le 2\|f_3\|_\infty$.

For the case $s=8$, we use the bound

$$\|g_{s-1}^{(j)}\|_\infty \le \frac{2^{j+1}}{2^{s-2}(s-j-2)!},$$

see [Schumaker, 2007, Thm. 4.36] with a different normalization of the spline and again $|g_{s-1}^{(j-1)}|_V \le 2\|g_{s-1}^{(j)}\|_\infty$. □

Appendix B

Measure Theory

Here, we give a quick reminder on the measure theoretic statements, we use in the thesis. These can be found, e.g. in [Rudin, 1987].

Let X be a locally compact Hausdorff space and $\mathcal{B}(X)$ the corresponding Borel σ-algebra. A mapping

$$\mu : \mathcal{B}(X) \to \mathbb{C}, \quad \text{resp.} \quad \mu : \mathcal{B}(X) \to \mathbb{R},$$

is called *complex* resp. *signed Borel measure*, if it is σ-additive, i.e. for pairwise disjoint $E_1, E_2, \dots \in \mathcal{B}(X)$, we have

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i).$$

The *variation* of μ on $E \in \mathcal{B}(X)$ is given by,

$$|\mu|(E) = \sup\left\{\sum_i |\mu(E_i)|\right\},$$

where the supremum is taken over all partitions of E. The mapping $|\mu|$ defines a finite positive Borel measure.

Let μ be a complex or signed Borel measure and ν be a finite positive Borel measure. The measure μ is called *absolutely continuous* with respect to ν, denoted by $\mu \ll \nu$, if $\mu(E) = 0$ for all $E \in \mathcal{B}(X)$, such that $\nu(E) = 0$. The two complex or signed Borel measures μ, ν are called *mutually singular*, denoted by $\mu \perp \nu$, if there are two disjoint sets $A, B \subset X$, such that $A \cup B = X$ and $\mu(E) = 0$ for all $E \in \mathcal{B}(X) \cap A$, while $\nu(E) = 0$ for all $E \in \mathcal{B}(X) \cap B$.

Theorem B.1 (Lebesgue decomposition)**.** *Let μ be a complex resp. signed Borel measure and ν be a finite positive Borel measure. Then μ has a unique decomposition of the form*

$$\mu = \mu_\nu + \mu_{\nu^\perp},$$

such that $\mu_\nu, \mu_{\nu^\perp}$ are complex resp. signed measures, with the property $\mu_\nu \ll \nu$ and $\mu_{\nu^\perp} \perp \nu$. If μ is positive, so are μ_ν and $\mu_{\nu^\perp}$.

Theorem B.2 (Polar decomposition)**.** *Let μ be a complex or signed measure. Then there is a measurable function h, such that $|h(x)| = 1$ for all $x \in X$ and*

$$\mu = h \cdot |\mu|.$$

Using the polar decomposition, we can define integration with respect to a complex measure μ, by

$$\int_A f(x)\mathrm{d}\mu(x) = \int_A f(x)h(x)\mathrm{d}|\mu|(x),$$

for all $|\mu|$-measurable sets. We set

$$\langle \mu, f\rangle := \int_X f(x)\mathrm{d}\mu(x),$$

whenever this is well-defined.

A complex resp. signed Borel measure is called *regular*, if for all $E \in \mathcal{B}(X)$,

$$\begin{aligned} |\mu|(E) &= \sup\{|\mu|(C) : C \subset E, C \text{ compact}\}, \\ &= \inf\{|\mu|(O) : E \subset O, O \text{ open}\}. \end{aligned}$$

Definition B.3. *The space of all regular complex resp. signed measures is denoted by $\mathcal{M}(X,\mathbb{C})$ resp. $\mathcal{M}(X,\mathbb{R})$.*

The *total variation of a measure*, given by

$$\|\mu\|_{TV} := |\mu|(X),$$

defines a norm on $\mathcal{M}(X,\mathbb{C})$ resp. $\mathcal{M}(X,\mathbb{R})$ and $(\mathcal{M}(X,\mathbb{C}), \|\cdot\|_{TV})$ and $(\mathcal{M}(X,\mathbb{R}), \|\cdot\|_{TV})$ are Banach spaces. If assertion hold for both spaces, we write $\mathcal{M}(X,\mathbb{K})$ or simply $\mathcal{M}(X)$.

The space of complex resp. real-valued continuous functions is denoted by $C(X,\mathbb{C})$ resp. $C(X,\mathbb{R})$. The subspace of function that vanish at infinity is given by

$$C_0(X,\mathbb{C}) := \{f \in C(X,\mathbb{C}) : \{x \in X : \|f(x)\| \geq \varepsilon\} \text{ is compact f.a } \varepsilon > 0\}$$

and in the same way for real-valued functions. Equipped with the supremum norm

$$\|f\|_\infty = \sup_{x\in X} |f(x)|,$$

they are Banach spaces.

Theorem B.4 (Representation Theorem of Riesz)**.** *For a locally compact Hausdorff space X, we have that $(C_0(X,\mathbb{C}))' \cong \mathcal{M}(X,\mathbb{C})$ and $(C_0(X,\mathbb{R}))' \cong \mathcal{M}(X,\mathbb{R})$. Moreover, the positive functionals on $C_0(X,\mathbb{R})$, denoted by $(C_0(X,\mathbb{R}))'_+$ can be identified with the positive Borel measures, denoted by $\mathcal{M}_+(X,\mathbb{R})$. Moreover, we have*

$$\|\mu\|_{TV} = \sup_{f\in C(X,\mathbb{C}), \|f\|_\infty \leq 1} |\langle \mu, g\rangle| = \sup_{f\in C(X,\mathbb{C}), \|f\|_\infty \leq 1} \operatorname{Re}(\langle \mu, f\rangle).$$

Let $\mu \in \mathcal{M}(X,\mathbb{R})$ be non-negative, then the L^p-spaces are defined in the usual way for $1 \leq p < \infty$

$$\begin{aligned} \mathcal{L}^p(X,\mu) &= \left\{ f : X \to \mathbb{C}, \text{ measurable } : \int_X |f(x)|^p \mathrm{d}\mu(x) < \infty \right\}, \\ \mathcal{N} &= \{f \in \mathcal{L}^p(X,\mu), f = 0\mu - \text{a.e}\}, \end{aligned}$$

$$L^p(X,\mu) = \mathcal{L}^p(X,\mu) \setminus \mathcal{N}.$$

With $\|f\|_p = \left(\int_X |f(x)|^p \mathrm{d}\mu(x)\right)^{1/p}$, these spaces are Banach spaces. For $p = 2$, it is a Hilbert space with the inner product

$$\langle f, g \rangle_{L^2} = \int_X f(x)\overline{g(x)} \mathrm{d}\mu(x), \quad f, g \in L^2(X,\mu).$$

Appendix C

Linear Algebra

In this section, we state several results, that are useful to us. This includes results regarding Schur complements, which can be found e.g. in [Zhang, 2006], and assertions regarding the cross product of vectors in $\mathbb{R}^3$.

Given a block matrix $X \in \mathbb{C}^{d \times d}$ of the form

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where D is invertible, the *Schur complement* of D in X is given by

$$X/D = A - BD^{-1}C.$$

Lemma C.1. *The matrix X is invertible, if D and X/D are invertible. In the case that*

$$X = \begin{pmatrix} A & B \\ B^H & C \end{pmatrix},$$

with $C \succ 0$, the matrix X is positive semi-definite, if and only if, $A \succcurlyeq 0$ and $X/C \succcurlyeq 0$.

Lemma C.2 (Quotient rule)**.** *Let X, D, E be nonsingular square matrices such that*

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad D = \begin{pmatrix} E & F \\ G & H \end{pmatrix}.$$

Then

$$X/D = (X/H)/(D/H).$$

The induced operator norm for the l^∞-norm on $\mathbb{C}^d$ is given by

$$\|X\|_\infty = \max_i \sum_j |x_{i,j}|.$$

Lemma C.3 (Neumann series)**.** *Let $X \in \mathbb{C}^{d\times d}$. If $\|I - X\|_\infty < 1$, then X is invertible with*

$$X^{-1} = \sum_{k=0}^{\infty} (I - X)^k.$$

Moreover, we have $\|X^{-1}\|_\infty < \frac{1}{1-\|I-X\|_\infty}$.

In the following we state some identities regarding the cross product of vectors in $\mathbb{R}^3$. Let $a, b \in \mathbb{R}^3$, then the cross product is given by the vector

$$a \times b = |a||b| \sin(\theta) n_{a,b},$$

where θ is the angle between a and b and $n_{a,b}$ is the unit vector perpendicular to a and b such that $a, b, n_{a,b}$ follow the right hand rule. Its components are given by

$$a \times b = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

Lemma C.4 (Identities cross product)**.** *Let $a, b, c, d \in \mathbb{R}^3$, then*

(i) $a \times b = -(b \times a)$,

(ii) $a \cdot (b \times c) = b \cdot (c \times a) = c \cdot (a \times b)$,

(iii) $a \times (b \times c) = b(a \cdot c) - c(a \cdot b)$,

(iv) $(a \times b) \times (a \times c) = (a \cdot (b \times c))a$,

(v) $(a \times b) \cdot (c \times d) = (a \cdot c)(b \cdot d) - (a \cdot d)(b \cdot c)$.

Appendix D

Convex Analysis

A short reminder on convex analysis, which can be found e.g. in [Peypouquet, 2015], which is used for the optimization problem in the measure space $\mathcal{M}(X)$, see also [Bredies and Pikkarainen, 2013].

Let (B, τ) be a Hausdorff space and $J : B \to \mathbb{R} \cup \{\infty\}$ be an extended real-valued function. We give a quick reminder on the existence and uniqueness of minimizers of the minimization problem

$$\min_{x \in B} J(x).$$

The function f is called *proper* if its *effective domain*, given by

$$\operatorname{dom}(J) := \{x \in B : J(x) < \infty\},$$

is not empty. A simple example is given by the indicator function of a subset $C \subset B$, i.e.

$$\chi_C(x) = \begin{cases} 0, & x \in B, \\ \infty, & x \notin C. \end{cases}$$

A function f is called *sequentially lower semi-continuous* at $x_0 \in \operatorname{dom}(J)$ if

$$\liminf_{x \to x_0} J(x) \geq J(x_0)$$

and *sequentially lower semi-continuous* if it is *sequentially lower semi-continuous* for all points in $\operatorname{dom}(J)$. For abbreviation we write *s.l.s.c*. Given two s.l.s.c functions J, G, the sum

$$J + \alpha G$$

is s.l.s.c for $\alpha \geq 0$. The property of sequentially lower semi-continuity can be employed to show existence of minimizers, which is known as the *direct method in the calculus of variations*.

Lemma D.1. *Suppose B is the dual of a separable normed space and $J : B \to \mathbb{R} \cup \{\infty\}$ is proper and s.l.s.c with respect to the weak*-topology on B. Then J has a minimizer.*

Proof. Choose a minimizing sequence $\{x_n\}_{n \in \mathbb{N}}$, i.e.

$$\lim_{n \to \infty} J(x_n) = \inf_{x \in B} J(x).$$

Using the *sequential Banach-Alaoglu Theorem*, i.e. bounded sets are sequentially compact with respect to the weak*-topology, there is a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ with $x_{n_k} \overset{w^\star}{\longrightarrow} \overline{x}$. By the sequential lower semi-continuity of J we have

$$\inf_{x\in B} J(x) = \lim_{k\to\infty} J(x_{n_k}) \geq \liminf_{k\to\infty} J(x_{n_k}) \geq J(\overline{x}) \geq \inf_{x\in B} J(x).$$

□

A set $C \subset B$ is *convex* if

$$tx + (1-t)y \in C,$$

for all $x, y \in C$ and $t \in (0,1)$. An extended real-valued function J is called convex, if

$$J(tx + (1-t)y) \leq tJ(x) + (1-t)J(y), \tag{D.1}$$

for all $x, y \in X$ and $t \in [0,1]$. If the inequality is strict, whenever $x \neq y$ and $t \in (0,1)$, J is called *strictly convex*. It is easy to check that the indicator function χ_C is convex, if and only if the set C is convex. Again, for J, G convex, we have the convexity of

$$J + \alpha G$$

for $\alpha \geq 0$, meaning the proper, convex and lower semi-continuous functions are forming a convex cone. The *subdifferential* of a proper convex function J is given by

$$\partial J(x) = \{x^* \in B^* : J(y) \geq J(x) + \langle x^*, y - x\rangle, \text{ for all } y \in B\}.$$

Theorem D.2 (Fermat's Rule). *Let $J : B \to \mathbb{R}\cup\{\infty\}$ be proper and convex, then x is a global minimizer of J if, and only if, $0 \in \partial J(x)$.*

Theorem D.3 (Moreau-Rockafellar Theorem). *Let $J, G : B \to \mathbb{R}\cup\{\infty\}$ be proper, convex, and lower-semicontinuous. For each $x \in B$, we have*

$$\partial J(x) + \partial G(x) \subseteq \partial(J+G)(x).$$

Equality holds for every $x \in B$ iff is continuous at some $x_0 \in \mathrm{dom}(G)$.

In the following, we concentrate on the case $B = \mathcal{M}(X, \mathbb{C})$, where (X, g) is a compact d-dimensional smooth Riemannian manifold. Since the Laplace-Beltrami operator is compact in the L^2-topology, the eigenspaces H_l are dense in $L^2(X) = L^2(X, \nu)$, where ν is the Riemannian volume measure. This means,

$$L^2(X) = \mathrm{cl}_{\|\cdot\|_2} \bigcup_{l=1}^{\infty} H_l,$$

where H_l is the eigenspace to the $l-$th eigenvalue. The frequency information is now carried in the ascending spaces

$$\Pi_N(X) = \mathrm{span} \bigcup_{l=1}^{N} H_l, \quad N = 1, 2 \ldots.$$

In this setting, the L^2-projection operator onto the space $\Pi_N(X)$ for a fixed N can be written in the following way. Choose for each $l = 1, 2, \ldots$ an orthonormal basis $\{\varphi_{l,k}\}_{k=1}^{\dim(H_l)}$ of H_l, and set

$$K_N(x,y) = \sum_{l=1}^{N} \sum_{k=1}^{\dim(H_l)} \varphi_{l,k}(x)\overline{\varphi_{l,k}(y)}.$$

Then the projection operator $S_N : L^2(X) \to C(X)$ onto the space $\Pi_N(X)$ can be written as

$$S_N f(x) = \int_X f(y) K_N(x,y) \mathrm{d}\nu(y).$$

For a measure $\mu^\star \in \mathcal{M}(X,\mathbb{K})$, the available information are given by $\mathcal{S}_N^* \mu^\star$ for some $N \in \mathbb{N}$. We state the necessary *strong duality* result in the following theorem.

Theorem D.4 (Strong duality). *Let X be a compact d-dimensional smooth Riemannian manifold and $\mu^\star \in \mathcal{M}(X,\mathbb{K})$. Then for all $N \in \mathbb{N}$, $\tau > 0$, $\varepsilon > 0$ and $\eta \in \mathcal{M}(X,\mathbb{K})$, such that $\|\mathcal{S}_N^* \eta\|_{L^2} \leq \varepsilon$, we have the following duality result. The values*

$$\begin{aligned}
p^\star &= \inf_{\mu \in \mathcal{M}(X,\mathbb{K})} \{\|\mu\|_{TV} : \mathcal{S}_N^* \mu = \mathcal{S}_N^* \mu^\star\}, \\
d^\star &= \sup_{f \in \Pi_N(X)} \{\mathrm{Re}\langle f, \mathcal{S}_N^* \mu^\star \rangle_{L^2} : \|f\|_\infty \leq 1\}, \\
p^\star_\tau &= \inf_{\mu \in \mathcal{M}(X,\mathbb{K})} \frac{1}{2} \|\mathcal{S}_N^*(\mu - \mu^\star - \eta)\|^2_{L^2} + \tau \|\mu\|_{TV}, \\
d^\star_\tau &= \sup_{f \in \Pi_N(X)} \{\mathrm{Re}\langle \mathcal{S}_N^*(\mu^\star + \eta), f\rangle_{L^2} - \tau \|f\|^2_{L^2} : \|f\|_\infty \leq 1\},
\end{aligned}$$

are finite and $p^\star = d^\star$ and $p^\star_\tau = d^\star_\tau$.

Proof. The proof follows from standard Fenchel-duality results, which can be found e.g. in [Borwein and Zhu, 2005], and the fact that

$$\|\mu\|_{TV} = \sup_{f \in C(X)} \{\mathrm{Re}(\langle \mu, f \rangle) : \|f\|_\infty \leq 1\}.$$

□

In the following, we briefly state some notions on Riemannian manifolds, which can be found e.g. in [Udriste, 1994]. Let (X, g) be a compact d-dimensional smooth Riemannian manifold. For a point $x \in X$ and a neighborhood U of x with a chart $\varphi : U \to \mathbb{R}^d$, we consider all curves $\gamma : (-1,1) \to X$ such that $\varphi \circ \gamma : (-1,1) \to \mathbb{R}^d$ is differentiable and we call

$$\dot\gamma(0) = (\varphi \circ \gamma)'(0) \in \mathbb{R}^d$$

a tangent vector at $x \in X$. Calling two curves γ_1, γ_2 equivalent if $\dot\gamma_1(0) = \dot\gamma_2(0)$, the set of equivalence classes is called the *tangent space* $T_x X$ at x, which is a d-dimensional vector space. At each point $x \in X$ the Riemannian metric g defines an inner product on $T_x X$,

$$g_x : T_x X \times T_x X \to \mathbb{R},$$

that varies smoothly in x. A smooth curve $\gamma_0 : [a,b] \to \mathbb{R}$ is called a *geodesic*, if it has minimal length for all curves joining $\gamma_0(a)$ and $\gamma_0(b)$, i.e.

$$\gamma_0 = \mathrm{argmin}_\gamma \int_a^b \sqrt{g_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))} \mathrm{d}t.$$

The value

$$d(x,y) = \min \int_b^a \sqrt{g_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))} \mathrm{d}t,$$

where the minimization is taken with respect to all curves $\gamma : [a, b] \to \mathbb{R}$, such that $\gamma(a) = x$ and $\gamma(b) = y$, is called *geodesic distance* of $x, y \in X$. It defines a metric on X, that is compatible with the topology of X, i.e. the geodesic balls

$$B_r(x) = \{y \in X : d(x, y) < r\},$$

generate the topology. For each $x \in X$ and $v \in T_xX$, there is a unique geodesic $\gamma_{x,v}$, such that $\gamma_{x,v}(0) = x$ and $\dot{\gamma}_{x,v}(0) = v$. The mapping

$$\exp_x : T_xX \to X, \quad v \mapsto \gamma_{x,v}(1),$$

is called *exponential map* at $x \in X$. Taking the canonical coordinates $x_1, \ldots, x_d$ of T_xX yields the *Riemannian normal coordinates* centered at $x \in X$. Parametrizing the d-dimensional euclidean space T_xX by d-dimensional spherical coordinates $(r, \theta_1, \ldots, \theta_{d-1})$, yields the *polar coordinates* centered at $x \in X$. In these local coordinates, the Riemannian metric is represented by a symmetric positive definite matrix $(g_{ij}(y)) \in \mathbb{R}^{d\times d}$. Its inverse is denoted by $(g^{ij}(y)) = (g_{ij}(y))^{-1}$. At the center $x \in X$ of Riemannian normal coordinates we have $g_{ij}(x) = \delta_{ij}$. If we denote the normal coordinate mapping centered at $x \in X$ by φ, then the gradient of a smooth function is given locally by

$$(\nabla f)_i = \sum_{j=1}^{d} g^{ij} \partial_j (f \circ \varphi), \quad i = 1, \ldots, d.$$

The Christoffel symbols are given in normal coordinates centered at $x \in X$ for y in a neighborhood of x by

$$\Gamma^i_{kl}(y) = \frac{1}{2} \sum_{m=1}^{d} g^{im}(y) \left(\frac{\partial g_{mk}(y)}{\partial x^l} + \frac{\partial g_{ml}(y)}{\partial x^k} - \frac{\partial g_{kl}(y)}{\partial x^m} \right).$$

The Christoffel symbols $\tilde{\Gamma}^i_{kl}$ for the polar coordinates are defined analogous. If we denote the normal coordinate mapping centered at $x \in X$ by φ and the polar coordinate mapping by $\tilde{\varphi}$, then the local expression of the Hessian of a smooth function is given for $r > 0$, smaller than the injectivity radius of X, by

$$(Hf(y))_{ij} = \partial_{ij}(f \circ \varphi)(\varphi^{-1}(y)) - \sum_{k=1}^{d} \Gamma^k_{ij}(y) \partial_k (f \circ \varphi)(\varphi^{-1}(y)), \quad y \in B_r(x),$$

respectively

$$(\tilde{H}f(y))_{ij} = \partial_{ij}(f \circ \tilde{\varphi})(\tilde{\varphi}^{-1}(y)) - \sum_{k=1}^{d} \tilde{\Gamma}^k_{ij}(y) \partial_k (f \circ \tilde{\varphi})(\tilde{\varphi}^{-1}(y)), \quad y \in B_r(x) \setminus \{x\}.$$

Lemma D.5 (Local minimizer). *Let $r > 0$ be smaller than the injectivity radius of X and $f : X \to \mathbb{R}$ be a smooth function. If $x \in X$ is an isolated local minimizer of f and the matrix $(Hf(y))_{ij}$ respectively $(\tilde{H}f(y))_{ij}$ is positive definite for all $y \in B_r(x) \setminus \{x\}$, then x is an isolated minimizer of f on $B_r(x)$.*

Bibliography

[Adams, 1878] Adams, J. (1878). On the expression of the product of any two Legendre's coefficients by means of a series of Legendre's coefficients. *Proc. R. Soc. Lond.*, 27:63–71.

[Atkinson and Han, 2012] Atkinson, K. and Han, W. (2012). *Spherical Harmonics and Approximations on the Unit Sphere: An Introduction*. Number 2044 in Lecture Notes in Mathematics. Springer.

[Bajaj et al., 2013] Bajaj, C., Bauer, B., Bettadupura, R., and Vollrath, A. (2013). Nonuniform Fourier transforms for rigid-body and multidimensional rotational correlations. *SIAM J. Sci. Comput.*, 35(4).

[Bendory, 2017] Bendory, T. (2017). Robust recovery of positive stream of pulses. *IEEE Transactions on Signal Processing*, 65(8):2114–2122.

[Bendory et al., 2016a] Bendory, T., Bar-Zion, A. D., Adam, D., Dekel, S., and Feuer, A. (2016a). Stable support recovery of stream of pulses with application to ultrasound imaging. *IEEE Trans. Signal Processing*, 64(14):3750–3759.

[Bendory et al., 2014] Bendory, T., Dekel, S., and Feuer, A. (2014). Exact recovery of non-uniform splines from the projection onto spaces of algebraic polynomials. *J. Approx. Theory*, 182:7–17.

[Bendory et al., 2015a] Bendory, T., Dekel, S., and Feuer, A. (2015a). Exact recovery of Dirac ensembles from projection onto spaces of spherical harmonics. *Constr. Approx.*, 42(2):183–207.

[Bendory et al., 2015b] Bendory, T., Dekel, S., and Feuer, A. (2015b). Super-resolution on the sphere using convex optimization. *IEEE Transactions on Signal Processing*, 63(9):2253–2262.

[Bendory et al., 2016b] Bendory, T., Dekel, S., and Feuer, A. (2016b). Robust recovery of stream of pulses using convex optimization. *Journal of Mathematical Analysis and Applications*, 442(2):511–536.

[Bendory and Eldar, 2015] Bendory, T. and Eldar, Y. (2015). Recovery of sparse positive signals on the sphere from low resolution measurements. arXiv:1508.06262.

[Benedetto and Li, 2016] Benedetto, J. J. and Li, W. (2016). Super-resolution by means of Beurling minimal extrapolation. *arXiv preprint arXiv:1601.05761*.

[Bernstein and Fernandez-Granda, 2017] Bernstein, B. and Fernandez-Granda, C. (2017). Deconvolution of point sources: A sampling theorem and robustness guarantees. *arXiv preprint arXiv:1707.00808*.

[Bernstein, 1912] Bernstein, S. (1912). On the best approximation of continuous functions by polynomials of a given degree. *Comm. Soc. Math. Kharkow, Ser*, 2(13):49–194.

[Bertero and Boccacci, 2003] Bertero, M. and Boccacci, P. (2003). Super-resolution in computational imaging. *Micron*, 34(6):265–273.

[Beurling, 1938] Beurling, A. (1938). Sur les integrales de fourier absolument convergentes et leur application a une transformation fonctionelle, ninth scandinavian math. congress, helsingfors, 1938. *Reprinted in Collected Works of Arne Beurling*, 2:39–60.

[Beurling, 1989a] Beurling, A. (1989a). Balayage of Fourier-Stieltjes transforms. *The Collected Works of Arne Beurling*, 2:341–350.

[Beurling, 1989b] Beurling, A. (1989b). Interpolation for an interval in $\mathbb{R}$. *The Collected Works of Arne Beurling*, 2:351–365.

[Bojanov et al., 1993] Bojanov, B., Hakopian, H., and Sahakian, A. (1993). *Spline Functions and Multivariate Interpolations*. Kluwer Academic Publishers.

[Borwein and Zhu, 2005] Borwein, J. M. and Zhu, Q. J. (2005). Variational techniques in convex analysis. *Techniques of Variational Analysis*, pages 111–163.

[Boyer et al., 2017] Boyer, C., De Castro, Y., and Salmon, J. (2017). Adapting to unknown noise level in sparse deconvolution. *Information and Inference: A Journal of the IMA*, page iaw024.

[Bredies and Pikkarainen, 2013] Bredies, K. and Pikkarainen, H. (2013). Inverse problems in spaces of measures. *ESAIM: Control, Optimisation and Calculus of Variations*, 19(1):190–218.

[Bunge, 1982] Bunge, H. (1982). *Texture Analysis in Material Science*. Butterworths.

[Burger, 2016] Burger, M. (2016). *Advances in Mathematical Modeling, Optimization and optimal Control*, chapter Bregman Distances in Inverse Problems and Partial Differential Equations. Springer.

[Burger and Osher, 2004] Burger, M. and Osher, S. (2004). Convergence rates of convex variational regularization. *Inverse Problems*, 20(5).

[Candés and Fernandez-Granda, 2014] Candés, E. and Fernandez-Granda, C. (2014). Towards a mathematical theory of super-resolution. *Comm. on Pure and Appl. Math.*, 67(6):906–956.

[Candés and Fernandez-Granda, 2013] Candés, E. J. and Fernandez-Granda, C. (2013). Super-resolution from noisy data. *Journal of Fourier Analysis and Applications*, 19(6):1229–1254.

[Castrillon-Candas et al., 2005] Castrillon-Candas, J., Siddavanahalli, V., and Bajaj, C. (2005). Nonequispaced Fourier transforms for protein-protein docking. ICES Report 05-44, Univ. Texas.

[Castro and Mijoule, 2015] Castro, Y. D. and Mijoule, G. (2015). Non-uniform spline recovery from small degree polynomial approximation. *Journal of Mathematical Analysis and applications*, 430(2):971–992.

[Chandrasekaran et al., 2012] Chandrasekaran, V., Recht, B., Parrilo, P., and Willsky, A. (2012). The convex geometry of linear inverse problems. *Found. Comp. Math.*, 12(6):805–849.

[Chen et al., 1998] Chen, S., Donoho, D., and Saunders, M. (1998). Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20:33–61.

[Chi et al., 2011] Chi, Y., Scharf, L., Pezeshki, A., and Calderbank, A. (2011). Sensitivity to basis mismatch in compressed sensing. *IEEE Transactions on Signal Processing*, 59(5):2182–2195.

[Chirikjian, 2012] Chirikjian, G. (2012). *Stochastic Models, Information Theory and Lie Groups*, volume 2 of *ANHA*. Brikhäuser.

[Chirikjian and Kyatkin, 2000] Chirikjian, G. and Kyatkin, A. (2000). *Engineering Applications of Noncummutative Harmonic Analysis*. CRC Press.

[Comaniciu and Meer, 2002] Comaniciu, D. and Meer, P. (2002). Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619.

[Cremer and Masters, 2013] Cremer, C. and Masters, B. R. (2013). Resolution enhancement techniques in microscopy. *The European Physical Journal H*, 38(3):281–344.

[de Castro and Gamboa, 2012] de Castro, Y. and Gamboa, F. (2012). Exact reconstruction using Beurling minimal extrapolation. *J. Math. Anal. Appl.*, 395(1):336–354.

[De Castro et al., 2017] De Castro, Y., Gamboa, F., Henrion, D., and Lasserre, J.-B. (2017). Exact solutions to super resolution on semi-algebraic domains in higher dimensions. *IEEE Transactions on Information Theory*, 63(1):621–630.

[Denoyelle et al., 2015a] Denoyelle, Q., Duval, V., and Peyré, G. (2015a). Asymptotic of sparse support recovery for positive measures. In *Journal of Physics: Conference Series*, volume 657, page 012013. IOP Publishing.

[Denoyelle et al., 2015b] Denoyelle, Q., Duval, V., and Peyré, G. (2015b). Support recovery for sparse deconvolution of positive measures. *arXiv preprint arXiv:1506.08264*.

[Deslauriers-Gauthier and Marziliano, 2012] Deslauriers-Gauthier, S. and Marziliano, P. (2012). Spherical finite rate of innovation theory for the recovery of fiber orientations. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 2294–2297. IEEE.

[Deslauriers-Gauthier and Marziliano, 2013] Deslauriers-Gauthier, S. and Marziliano, P. (2013). Sampling signals with a finite rate of innovation on the sphere. *IEEE Transactions on Signal Processing*, 61(18):4552–4561.

[Diederichs and Iske, 2015] Diederichs, B. and Iske, A. (2015). Parameter estimation for bivariate exponential sums. In *Sampling Theory and Applications (SampTA), 2015 International Conference on*, pages 493–497. IEEE.

[Donoho, 1992] Donoho, D. (1992). Superresolution via sparsity constraints. *SIAM J. Math. Anal.*, 23(5):1309–1331.

[Dumitrescu, 2006] Dumitrescu, B. (2006). Trigonometric polynomials positive on frequency domains and applications to 2-D FIR filter design. *IEEE transactions on signal processing*, 54(11):4282–4292.

[Dumitrescu, 2007] Dumitrescu, B. (2007). *Positive Trigonometric Polynomials and Signal Processing Applications*. Springer.

[Duval and Peyré, 2015] Duval, V. and Peyré, G. (2015). Exact support recovery for sparse spikes deconvolution. *Found. Comp. Math.*, 15:1315–1355.

[Duval and Peyré, 2015] Duval, V. and Peyré, G. (2015). Sparse spikes deconvolution on thin grids. *arXiv preprint arXiv:1503.08577*.

[Duval and Peyré, 2016] Duval, V. and Peyré, G. (2016). Sparse spikes super-resolution on thin grids II: the continuous basis pursuit.

[Edmonds, 1957] Edmonds, A. (1957). *Angular Momentum in Quantum Mechanics*. Princeton University Press.

[Fernandez-Granda, 2016] Fernandez-Granda, C. (2016). Super-resolution of point sources via convex programming. *Information and Inference: A Journal of the IMA*, 5(3):251–303.

[Filbir and Mhaskar, 2010] Filbir, F. and Mhaskar, H. (2010). A quadrature formula for diffusion polynomials corresponding to a generalized heat kernel. *Journal of Fourier Analysis and Applications*, 16(5):629–657.

[Filbir et al., 2012] Filbir, F., Mhaskar, H., and Prestin, J. (2012). On the problem of parameter estimation in exponential sums. *Constructive Approximation*, 35(3):323–343.

[Gamboa and Gassiat, 1994] Gamboa, F. and Gassiat, E. (1994). The maximum entropy method on the mean: Applications to linear programming and superresolution. *Math. Programming*, 66(1):103–122.

[Gamboa and Gassiat, 1996] Gamboa, F. and Gassiat, E. (1996). Sets of superresolution and the maximum entropy method on the mean. *SIAM J. Math. Anal.*, 27(4):1129–1152.

[Gräf, 2013a] Gräf, M. (2013a). *Efficient Algorithms for the Computation of Optimal Quadrature Points on Riemannian Manifolds*. PhD thesis, TU Chemnitz.

[Gräf, 2013b] Gräf, M. (2013b). http://homepage.univie.ac.at/manuel.graef/quadrature.php.

[Gräf and Hielscher, 2015] Gräf, M. and Hielscher, R. (2015). Fast global optimization on the torus, the sphere, and the rotation group. *SIAM Journal on Optimization*, 25(1):540–563.

[Gräf and Kunis, 2008] Gräf, M. and Kunis, S. (2008). Stability results for scattered data interpolation on the rotation group. *Elect. Trans. on Num. Ana.*, 31:30–39.

[Grant and Boyd, 2008] Grant, M. and Boyd, S. (2008). Graph implementations for nonsmooth convex programs. In Blondel, V., Boyd, S., and Kimura, H., editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited. `http://stanford.edu/~boyd/graph_dcp.html`.

[Grant and Boyd, 2014] Grant, M. and Boyd, S. (2014). CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`.

[Hielscher, 2013] Hielscher, R. (2013). Kernel density estimation on the rotation group and its application to crystallographic texture analysis. *Journal of Multivariate Analysis*, 119:119–143.

[Hielscher et al., 2008] Hielscher, R., Potts, D., Prestin, J., Schaeben, H., and Schmalz, M. (2008). The Radon transform on $SO(3)$: A Fourier slice theorem and numerical inversion. *Inverse Problems*, 24.

[Hofmann et al., 2007] Hofmann, B., Kaltenbacher, B., and Pöschel, C. (2007). A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Problems*, 23:987–1010.

[Hua and Sarkar, 1990] Hua, Y. and Sarkar, T. K. (1990). Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(5):814–824.

[Huang et al., 2009] Huang, B., Bates, M., and Zhuang, X. (2009). Super-resolution fluorescence microscopy. *Annual review of biochemistry*, 78:993–1016.

[Keiner et al., 2007] Keiner, J., Kunis, S., and Potts, D. (2007). Efficient reconstruction of functions on the sphere from scattered data. *Journal of Fourier Analysis and Applications*, 13(4):435–458.

[Keiner and Vollrath, 2012] Keiner, J. and Vollrath, A. (2012). A new algorithm for the nonequispaced fast Fourier transform on the rotation group. *SIAM J. Sci. Comput.*, 34(5):A2599–A2624.

[Kostelec and Rockmore, 2008] Kostelec, P. and Rockmore, D. (2008). FFTs on the rotation group. *J. Fourier Anal. Appl.*, 14:145–179.

[Kovacs et al., 2003] Kovacs, J., Chaón, P., Cong, Y., Metwally, E., and Wriggers, W. (2003). Fast rotational matching of rigid bodies by fast Fourier transfrom acceleration of five degrees of freedom. *Acta Crystallographica*, Sect. D(59):1371–1376.

[Kunis et al., 2016a] Kunis, S., Möller, H. M., and von der Ohe, U. (2016a). Prony's method on the sphere. *arXiv preprint arXiv:1603.02020*.

[Kunis et al., 2016b] Kunis, S., Peter, T., Römer, T., and von der Ohe, U. (2016b). A multivariate generalization of Prony's method. *Linear Algebra and its Applications*, 490:31–47.

[Kunis and Potts, 2007] Kunis, S. and Potts, D. (2007). Stability results for scattered data interpolation by trigonometric polynomials. *SIAM J. Scientific Comp.*, 29(4):1403–1419.

[Lewis, 1996] Lewis, A. (1996). Superresolution in the Markov moment problem. *Journal of Mathematical Analysis and Applications*, 197:774–780.

[Li and Tang, 2016] Li, Q. and Tang, G. (2016). Approximate support recovery of atomic line spectral estimation: A tale of resolution and precision. In *Signal and Information Processing (GlobalSIP), 2016 IEEE Global Conference on*, pages 153–156. IEEE.

[Li, 2017] Li, W. (2017). Elementary L^∞ error estimates for super-resolution de-noising. arXiv:1702.03021.

[Lindberg, 2012] Lindberg, J. (2012). Mathematical concepts of optical superresolution. *Journal of Optics*, 14(8):083001.

[Mairhuber, 1956] Mairhuber, J. C. (1956). On Haars theorem concerning Chebychev approximation problems having unique solutions. *Proceedings of the American Mathematical Society*, 7(4):609–615.

[Mhaskar and Prestin, 2000] Mhaskar, H. and Prestin, J. (2000). On the detection of singularities of a periodic function. *Adv. in Comp. Math.*, 12:95–131.

[Morgenshtern and Candés, 2014] Morgenshtern, V. I. and Candés, E. J. (2014). Stable super-resolution of positive sources: the discrete setup. *preprint*.

[Nasrollahi and Moeslund, 2014] Nasrollahi, K. and Moeslund, T. B. (2014). Super-resolution: a comprehensive survey. *Machine vision and applications*, 25(6):1423–1468.

[Park et al., 2003] Park, S. C., Park, M. K., and Kang, M. G. (2003). Super-resolution image reconstruction: a technical overview. *IEEE signal processing magazine*, 20(3):21–36.

[Pelletier, 2005] Pelletier, B. (2005). Kernel density estimation on Riemannian manifolds. *Statistics & Probability Letters*, 73(3):297–304.

[Peter, 2013] Peter, T. (2013). *Generalized Prony Method*. PhD thesis, Georg-August Universität Göttingen.

[Peypouquet, 2015] Peypouquet, J. (2015). *Convex Optimization in Normed Spaces*. SpringerBriefs in Optimization. Springer.

[Plonka and Wischerhoff, 2013] Plonka, G. and Wischerhoff, M. (2013). How many Fourier samples are needed for real function reconstruction? *Journal of Applied Mathematics and Computing*, 42(1-2):117–137.

[Potts et al., 2009] Potts, D., Prestin, J., and Vollrath, A. (2009). A fast algorithm for nonequispaced Fourier transforms on the rotation group. *Numerical Algorithms*, 52(355).

[Potts and Tasche, 2010] Potts, D. and Tasche, M. (2010). Parameter estimation for exponential sums by approximate Prony method. *Signal Processing*, 90(5):1631–1642.

[Potts and Tasche, 2011] Potts, D. and Tasche, M. (2011). Nonlinear approximation by sums of nonincreasing exponentials. *Applicable Analysis*, 90(3-4):609–626.

[Potts and Tasche, 2013] Potts, D. and Tasche, M. (2013). Parameter estimation for nonincreasing exponential sums by Prony-like methods. *Linear Algebra and its Applications*, 439(4):1024–1039.

[Rauhut and Foucart, 2013] Rauhut, H. and Foucart, S. (2013). *A Mathematical Introduction to Compressive Sensing*. Birkhäuser.

[Risbo, 1996] Risbo, T. (1996). Fourier transform summation of Legendre series and D-functions. *Journal of Geodesy*, 70:383–396.

[Roh et al., 2007] Roh, T., Dumitrescu, B., and Vandenberghe, L. (2007). Interior-point algorithms for sum-of-squares optimization of multidimensional trigonometric polynomials. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, volume 3, pages III–905. IEEE.

[Roh and Vandenberghe, 2006] Roh, T. and Vandenberghe, L. (2006). Discrete transforms, semidefinite programming, and sum-of-squares representations of nonnegative polynomials. *SIAM Journal on Optimization*, 16(4):939–964.

[Roy and Kailath, 1989] Roy, R. and Kailath, T. (1989). ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on acoustics, speech, and signal processing*, 37(7):984–995.

[Rudin, 1987] Rudin, W. (1987). *Real and Complex Analysis*. Tata McGraw-Hill Education.

[Schaeben and v.d. Boogart, 2003] Schaeben, H. and v.d. Boogart, K. (2003). Spherical harmonics in texture analysis. *Tectophysics*, 370:253–268.

[Schermelleh et al., 2010] Schermelleh, L., Heintzmann, R., and Leonhardt, H. (2010). A guide to super-resolution fluorescence microscopy. *The Journal of cell biology*, 190(2):165–175.

[Scherzer and Walch, 2009] Scherzer, O. and Walch, B. (2009). Sparsity regularization for Radon measures. In *Scale Space and Variational Methods in Computer Vision*, volume 5567 of *Lecture Notes in Computer Sience*, pages 452–463. Springer, Berlin Heidelberg.

[Schmid, 2008] Schmid, D. (2008). Marcinkiewicz-Zygmund inequalities and polynomial approximation from scattered data on $SO(3)$. *Numerical Functional Analysis and Optimization*, 29(7–8):855–882.

[Schmid, 2009] Schmid, D. (2009). *Scattered Data Approximation on the Rotation Group and Generalizations*. PhD thesis, TU München.

[Schmidt, 1986] Schmidt, R. (1986). Multiple emitter location and signal parameter estimation. *IEEE transactions on antennas and propagation*, 34(3):276–280.

[Schumaker, 2007] Schumaker, L. (2007). *Spline Functions: Basic Theory*. Cambridge University Press.

[Simon, 2011] Simon, B. (2011). *Convexity: An Analytic Viewpoint*, volume 187 of *Cambridge Tracts in Mathematics*. Cambridge University Press.

[Stevensson and Edén, 2011] Stevensson, B. and Edén, M. (2011). Interpolation by fast Wigner transforms for rapid calculations of magnetic spectra from powders. *J. Chem. Phys.*, 134.

[Stoica et al., 2005] Stoica, P., Moses, R. L., et al. (2005). *Spectral Analysis of Signals*, volume 452. Pearson Prentice Hall Upper Saddle River, NJ.

[Sturm, 1997] Sturm, J. (1997). *Primal-Dual Interior Point Approach to Semidefinite Programming*. PhD thesis, Tilburg university.

[Sturm, 1999] Sturm, J. (1999). Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653. Version 1.05 available from `http://fewcal.kub.nl/sturm`.

[Tang et al., 2013] Tang, G., Bhaskar, B., and Recht, B. (2013). Sparse recovery over continuous dictionaries-just discretize. In *Asilomar Conference on Signals, Systems and Computers, 2013*.

[Tang et al., 2015] Tang, G., Bhaskar, B., and Recht, B. (2015). Near minimax line spectral estimation. *IEEE Transactions on Information Theory*, 61(1).

[Teutsch and Kellermann, 2006] Teutsch, H. and Kellermann, W. (2006). Acoustic source detection and localization based on wavefield decomposition using circular microphone arrays. *The Journal of the Acoustical Society of America*, 120(5):2724–2736.

[Tuzel et al., 2005] Tuzel, O., Subbarao, R., and Meer, P. (2005). Simultaneous multiple 3D motion estimation via mode finding on Lie groups. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 1, pages 18–25 Vol. 1.

[Udriste, 1994] Udriste, C. (1994). *Convex Functions and Optimization Methods on Riemannian Manifolds*, volume 297. Springer Science & Business Media.

[v.d. Boogart et al., 2007] v.d. Boogart, K., Hielscher, R., Prestin, J., and Schaeben, H. (2007). Kernel-based methods for inversion of the Radon transform on $SO(3)$ and their application to texture analysis. *J. Comp. Appl. Math.*, 199:122–140.

[Vielva et al., 2003] Vielva, P., Martinez-Gonzalez, E., Gallegos, J. E., Toffolatti, L., and Sanz, J. L. (2003). Point source detection using the spherical mexican hat wavelet on simulated all-sky Planck maps. *Monthly Notices of the Royal Astronomical Society*, 344(1):89–104.

[Wischerhoff and Plonka, 2016] Wischerhoff, M. and Plonka, G. (2016). Reconstruction of polygonal shapes from sparse Fourier samples. *Journal of Computational and Applied Mathematics*, 297:117–131.

[Xu et al., 2014] Xu, W., Cai, J.-F., Mishra, K. V., Cho, M., and Kruger, A. (2014). Precise semidefinite programming formulation of atomic norm minimization for recovering d-dimensional (d$\geq$2) off-the-grid frequencies. In *Information Theory and Applications Workshop (ITA), 2014*, pages 1–4. IEEE.

[Young, 2001] Young, R. (2001). *An Introduction to Non-Harmonic Fourier Series, Revised*. Academic Press (New York).

[Zhang, 2006] Zhang, F. (2006). *The Schur Complement and its Applications*, volume 4. Springer Science & Business Media.