KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Engineering Science

Applications and analysis of lattice points: time-stepping and integration over $\mathbb{R}^d$

Yuya Suzuki

Supervisor:
Prof. dr. ir. Dirk Nuyens

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Computer Science

January 2020

Applications and analysis of lattice points: time-stepping and integration over $\mathbb{R}^d$

Yuya SUZUKI

Examination committee:
Prof. dr. ir. Robert Puers, chair
Prof. dr. ir. Dirk Nuyens, supervisor
Prof. dr. ir. Ronald Cools, co-supervisor
Prof. dr. ir. Daan Huybrechs
Prof. dr. ir. Wim Michiels
Prof. dr. Christian Lubich
(Universität Tübingen)
Prof. dr. Daniel Potts
(TU Chemnitz)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Computer Science

January 2020

Preface

In the middle of August 2015, I received the offer for this PhD position after assignments and an interview. It was late at night in Japan and I got too happy to sleep well.

I would like to express my deepest gratitude to my supervisor Dirk Nuyens, for his guidance, patience and inspiration. Through Dirk I learned so many things, not only specific research-related knowledge but also a more general attitude towards new results. During these four years he introduced me to many inspiring researchers, which I appreciate very much.

Ronald Cools, my co-supervisor, I thank for his invaluable support. His advice has always been accurate, clear, and reliable. Because I knew I can ask him for advice, I always felt safe.

I thank all members in my examination committee. I thank Daniel Potts for his accurate remarks on this thesis. I thank Christian Lubich for pointing out an important aspect in this thesis so that I could modify it before finalizing. I also thank for his warm encouragement to my work, which gave me confidence. I thank Wim Michiels for his comments, which increase the readability of the thesis. I thank Daan Huybrechs for examining the thesis in detail and for the valuable comments. I appreciate his assistance for giving me feedback after the preliminary defence. I thank Robert Puers for being the chairperson of my examination committee.

I would like to express my sincere gratitude to all members of the QMC group. I thank Adrian for his friendship during these four years, not only as an everyday lunch mate, but also as a trustful colleague to discuss QMC topics. I thank Gowri for her kindness and thoughtfulness. I am very grateful that I could collaborate with her. I thank Pieterjan for his openness and helpfulness. I am thankful especially for his help when I had problems with Dutch language. I thank Dong for her humor and sharp mind. I thank Zhiyu for being a calm office mate.

I thank my many colleagues in the department, Bart, Bert, Bruno, Daan Camps, Deesh, Dirk Roose, Giovanni, Emil, Francesco, Ilias, Joris, Julian, Karl, Keith, Kobe, Koen, Korneel, Libo, Luca, Marcus, Micol, Neline, Nick, Nico, Niel, Przemek, Roel, Simon, Stefan, Stelios, Vincent and Ward. I also thank the system group and secretaries, especially Kris, Margot and Inge, for their practical assistance.

I would like to thank the QMC community. I thank Peter Kritzer for being friendly and thoughtful, also for giving me the opportunity of a research visit at RICAM. I thank Frances Kuo for her valuable comments and encouragements on my research. I thank Yoshihito Kazashi for talking with me when we meet at different conferences. He taught me many valuable tips as an early stage researcher, and also I enjoyed his sense of humor very much. I also thank Takashi Goda, Kosuke Suzuki and Alec Gilbert for interesting discussions during the conferences.

I thank people who enriched my life in Leuven. I thank Stelios for bouldering together for three years. Thanks to him I could enjoy Belgian beer without changing my figure. I thank Kris from Cafe Noir. If the statement "a mathematician is a machine for turning coffee into theorems" is true, the ingredients of the theorems in this thesis are the coffee beans which he has sold to me.

I thank my family and friends in Japan. My parents always allowed me to pursue what I want. I am very happy to have such parents. My brother is always very helpful. I thank my friends from my master's study for warm friendship, especially Daisuke Sugiura, Kazumasa Aoki and Yo Takahashi. I thank my master's thesis supervisor Junichi Imai for his support when I was planning to get a PhD position abroad.

Finally, but not least, I thank my partner Franziska Steiner. Without her unconditional mental and physical support, I would certainly not have finished this thesis. I thank not only for her support but also for her inspiration. She has been always curious for new things and trying to broaden her mind, which inspired me as a fellow being who explores the unknown.

Abstract

This thesis aims to develop quasi-Monte Carlo methods for new applications, specifically real world applications in practical settings. We propose new numerical methods to solve multidimensional problems and analyse numerical errors of the proposed methods. Particularly, we focus on lattice points which is one important branch of quasi-Monte Carlo methods.

We consider two problems: the time-dependent Schrödinger equation and integration over $\mathbb{R}^d$. For the time-dependent Schrödinger equation, our contributions are the following. We provide a new algorithm which combines the Fourier spectral collocation method on lattice points with higher-order exponential operator splitting. We prove the regularity of the solution in terms of Korobov spaces which is necessary for using lattice points. Then we provide the total error bounds. With explicitly specified regularity conditions, the time stepping error is proved to have higher-order convergence. We conduct numerical experiments in various settings. Comparing with the sparse grid method [26], our method is shown to be superior. Even in higher dimensions and with higher-order schemes, our method shows stable and higher-order convergence numerically. One of the essential tasks here is to compute the Fourier transform and the inverse transform repeatedly in a higher-dimensional space for simulating the time-stepping operator of the time-dependent Schrödinger equation in a stable manner. Our proposed method solves this task efficiently.

The second problem we consider is numerical integration over $\mathbb{R}^d$. For quasi-Monte Carlo methods, integration over the unit cube is well studied, whereas integration over $\mathbb{R}^d$ is still a challenging problem. We aim to obtain higher-order convergence under general measures. We propose a new cubature method where we first truncate the integration domain $\mathbb{R}^d$ to a box and then apply the lattice rule. For analysing the error of this method, we make use of the Bernoulli polynomial method which was originally proposed for modifying a non-periodic function into a periodic function. By carefully defining two kinds of reproducing kernel Hilbert spaces, we can interpret the Bernoulli polynomial method as an

orthogonal projection from a non-periodic space onto a periodic space. This gives us an expression of the integration error in terms of a periodic part and a non-periodic part of the integrand. By combining this with a decay condition on the integrand, we obtain higher-order convergence of the integration error of our method including the truncation error. Our error analysis is general enough to include integrands which are finitely smooth under general measures.

Beknopte samenvatting

Het doel van deze thesis is het ontwikkelen van quasi-Monte Carlomethoden voor nieuwe toepassingen, met name toepassingen in de echte wereld en in praktische situaties. We stellen nieuwe numerieke methoden voor om meerdimensionale problemen op te lossen en analyseren de numerieke fout in de voorgestelde methoden. We concentreren ons in het bijzonder op roosterpunten, een belangrijke tak binnen de quasi-Monte Carlomethoden.

We beschouwen twee problemen: de tijdsafhankelijke Schrödingervergelijking en integratie over $\mathbb{R}^d$. Voor de tijdsafhankelijke Schrödingervergelijking zijn onze bijdragen de volgende. We stellen een nieuw algoritme voor dat de Fourierspectrale collocatiemethode op de roosterpunten combineert met hogere-orde exponentiële operatorsplitsing. We bewijzen de regelmatigheid van de oplossing in termen van Korobovruimtes, die nodig zijn voor het gebruik van de roosterpunten. Vervolgens formuleren we een bovengrens op de totale fout. Met gegeven regelmatigheidsvoorwaarden kan worden bewezen dat de tijdstapfout hogere-orde convergentie vertoont. We voeren verschillende numerieke experimenten uit die aantonen dat onze methode superieur is in vergelijking met met de *sparse grid*-methode [26]. Zelfs in meer dimensies en met hogere-orde schema's vertoont onze methode stabiele en hogere-orde convergentie. Eén van de essentiële taken is het herhaaldelijk berekenen van de Fouriertransformatie en zijn inverse in een meerdimensionale ruimte. Dit is nodig voor het op een stabiele manier simuleren van de tijdstapoperator in de tijdsafhankelijke Schrödingervergelijking. De methode die wordt voorgesteld in deze thesis doet dit op een efficiënte manier.

Het tweede probleem dat we beschouwen is numerieke integratie over $\mathbb{R}^d$. Voor quasi-Monte Carlomethoden is integratie over de eenheidskubus goed bestudeerd, terwijl integratie over $\mathbb{R}^d$ nog steeds een uitdagend probleem is. Het doel is het bekomen van hogere-orde convergentie onder een algemene maat. We stellen een nieuwe kubatuurmethode voor waarbij we eerst het integratiedomein $\mathbb{R}^d$ afbreken in een kubus en vervolgens een

standaard roosterregel toepassen. Voor het analyseren van de fout van deze methode gebruiken we de Bernoulliveeltermmethode, die oorspronkelijk werd voorgesteld voor het wijzigen van een niet-periodieke functie in een periodieke functie. Door het zorgvuldig definiëren van twee soorten Hilbertruimten met voortbrengende kern (Engels: “reproducing kernel Hilbert space”), kunnen we de Bernoulliveeltermmethode interpreteren als een orthogonale projectie van een niet-periodieke ruimte op een periodieke ruimte. Dit levert ons een uitdrukking op voor de integratiefout in termen van een periodiek stuk en een niet-periodiek stuk van de integrand. Door deze interpretatie te combineren met een voorwaarde op de afname van de integrand verkrijgen we de gezochte hogere-orde convergentie van de integratiefout van onze methode, inclusief de afbrekingsfout. Onze foutenanalyse is algemeen genoeg zodat deze ook integranden met eindige zachtheid bevat, onder een algemene maat.

Notations

i	imaginary unit
d	number of dimensions
n	number of points
j	running index for dimensions
i	running index for points
τ	running index for smoothness
$\mathfrak{u}, \mathfrak{w}$	set of indices
$\mathfrak{u} \setminus \mathfrak{w}$	difference set of indices for $\mathfrak{w} \subseteq \mathfrak{u}$
$\Lambda(\boldsymbol{z}, n)$	lattice point set with generating vector $\boldsymbol{z}$ and number of points n
$\mathcal{H}(K)$	reproducing kernel Hilbert space with kernel K
$\langle \cdot, \cdot \rangle_K$	inner product in reproducing kernel Hilbert space with kernel K
$\mathbb{C}$	set of complex numbers
$\mathbb{N}$	set of natural numbers
$\mathbb{R}$	set of real numbers
$\mathbb{Z}$	set of integers
$\mathbb{Z}_+$	set of non-negative integers
$\mathbb{Z}_n$	set of integers modulo n
$\mathbb{T}$	torus ($\mathbb{R}/\mathbb{Z}$)
$a \equiv b \pmod{n}$	a and b are congruent modulo n

$a \bmod n$ binary modulo operator modulo n of a

$\Lambda(\boldsymbol{z}, n)$ rank-1 lattice points with generating vector $\boldsymbol{z}$ and number of points n

$\mathcal{A}(\boldsymbol{z}, n)$ anti-aliasing set corresponding to the rank-1 lattice points $\Lambda(\boldsymbol{z}, n)$

∇^2 Laplacian $\sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2}$

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Quasi-Monte Carlo (QMC) methods are equal weight cubature rules on the unit cube where the sum of the weights is 1. Sometimes QMC rules are also mentioned as deterministic cubature rules in comparison with Monte Carlo methods. Randomized QMC (RQMC) methods, such as randomly shifted lattice rules and scrambled digital nets, are also one kind of QMC methods. Here the randomization is carefully operated to preserve the original structure of the QMC method. The essence of the QMC methods is this original structure; having a small discrepancy or having a small worst case error in some reproducing kernel Hilbert space. QMC methods are mainly used for numerical integration, but, due to their special structure, lattice points are also used for function approximation.

The aim of this thesis is to explore new usage of QMC methods, in particular lattice point sets, and analyse the numerical error of the method. In the following we explain the detailed contribution of each chapter.

1.1 Lattice points combined with exponential operator splitting for the time-dependent Schrödinger equation

In Chapters 3–5, we will consider the time-dependent Schrödinger equation of the form

$$\frac{\partial u}{\partial t}(\boldsymbol{x},t) = Au(\boldsymbol{x},t) + Bu(\boldsymbol{x},t), \quad u(\boldsymbol{x},0) = g(\boldsymbol{x}),$$

where A and B are differential operators, $\boldsymbol{x} \in [0,1)^d$ is a spatial variable, t is time, $g(\boldsymbol{x})$ is the initial condition and $u(\boldsymbol{x})$ is the solution of the PDE. We will always consider periodic boundary conditions. Our aim is to solve this problem by combining pseudospectral Fourier methods on lattice points with exponential operator splitting for time-dependent PDEs.

When one wants to solve multidimensional problems using the pseudospectral method, a natural choice for the spatial discretization is the tensor product of the one-dimensional method. For instance, when the domain is a d-dimensional box and the periodic boundary condition is imposed, regular grids (isotropic equidistant points) are commonly used. Then the interpolating Fourier basis is constructed on the grid and the PDE is solved in terms of the Fourier coefficients. The problem of using regular grids, or using tensor product methods, is that the computational cost is heavily cursed by the dimensionality. This curse of dimensionality is coming from the construction of the numerical algorithms and not necessarily from the problem itself. Therefore it may be possible to avoid that curse of dimensionality by constructing another discretization scheme.

Instead of regular grids, lattice points are proposed to be used for multidimensional problems; see, e.g., [17, 50, 53]. The success of using lattice points is partially due to that flexible structure. For instance, compared with the regular grid, one can choose the number of points more freely. Also with using lattice points for function approximation, the corresponding set of the basis functions is flexible to choose (see, e.g., [17, 53]). Using this flexibility, one can customize the method according to the certain problem structure. However, one needs to prove that the chosen method works better for that specific problem than other ways of using the flexibility. In Chapters 3–5, we will consider the time-dependent Scrödinger equation in high-dimensional settings. The big challenge here is coming from the combination of discretization of both time and high-dimensional spatial domain.

For the time-stepping scheme we employ the exponential operator splitting which is known to work very well with pseudospectral methods; see, e.g., [8, 54, 89, 91]. The success of this combination is owning to the fact that the

differential operators A and B can be applied separately by carefully choosing the basis functions which give exact eigenvalues of the differential operators.

In Chapter 3, we first compare our method with the method based on sparse grids proposed by Gradinaru [26]. The sparse grid method is using Strang splitting for time stepping which is a second order splitting method. Hence we consider Strang splitting combined with lattice points. Our contributions are the following: (1) we prove that our method gives the second order convergence in time stepping by specifying explicit conditions for the convergence results; (2) we give an expression for the total error for the full discretization, not only time stepping, but also including the spatial discretization of the dynamics; (3) our numerical results show that our method gives second order convergence consistently even for higher-dimensional cases. The advantage of our method compared to the sparse grid method is partially coming from the unitarity of the Fourier transform on lattice points, whereas the Fourier transform on the sparse grid is not unitary. This unitarity of the Fourier transform is a critical issue for the time-stepping error, because non-unitarity amplifies the accumulated error at every time step.

In Chapter 4, we generalize the results of the second order convergence to higher-order convergence. Error analysis for the time-dependent Schrödinger equation using exponential splitting has been done in [89] by Thalhammer, where the spatial discretization is not considered. Afterwards, a more generalized analysis is done in [90] by the same author, including the spatial discretization where the pseudospectral Fourier method on regular grids is considered. We use higher-order exponential operator splitting combined with a pseudospectral Fourier method on rank-1 lattice points. We prove local error bounds and show numerical results which confirm the theory.

In Chapter 5, we consider the nonlinear Schrödinger equation. We apply the same scheme, higher-order exponential operator splitting combined with pseudospectral Fourier method on rank-1 lattice points. We first show several numerical results. Then we also discuss the numerical error by following the analysis of [90]. In the end, the same error bound holds if a conjecture on rank-1 lattice structure holds which we also observe numerically.

1.2 Lattice rules for integration over $\mathbb{R}^d$

In the QMC literature, integration over $\mathbb{R}^d$ is a relatively new problem compared with integration over the unit cube. Integration over $\mathbb{R}^d$ often appears in finance and statistics where probability distributions over $\mathbb{R}^d$ are involved. There are several results using QMC method for this problem [22, 47, 57, 58], but no

result is known achieving higher-order convergence with lattice rules where the integrand has only finite smoothness.

In Chapter 6, we consider this problem. Our aim is to obtain error bounds for the new cubature algorithm using scaled lattice points. Our method is two-fold: first we truncate the integration domain from $\mathbb{R}^d$ to a finite box, then we apply a lattice rule in the box. Inside the box we analyse the difference between two reproducing kernel Hilbert spaces, Korobov spaces and unanchored Sobolev spaces, which are commonly used in the QMC literature. By making use of the orthogonal projection from an unanchored Sobolev space to a Korobov space, we obtain cubature error bounds.

1.3 Notations

Throughout the thesis $\mathbb{R}$ denotes the set of real numbers, $\mathbb{Z}$ denotes the set of integers, $\mathbb{Z}_n := \{0, 1, \ldots, n-1\}$ is the set of integers modulo n, $\mathbb{N} := \{1, 2, \ldots\}$ the natural numbers and $\mathbb{Q}$ the rational numbers. By $\mathbb{T}$ we mean the torus $\mathbb{R}/\mathbb{Z}$. We use I_n or just I to denote the $n \times n$ identity matrix. We distinguish between the normal equivalence in congruence modulo n as $a \equiv b \pmod{n}$ and the binary operation modulo n denoted by $a \mod n$ which returns the corresponding value in $\mathbb{Z}_n$ for mod n and in $\mathbb{T}$ for mod 1.

By $\mathcal{H}(K)$ we mean a reproducing kernel Hilbert space with the reproducing kernel K. In that space, inner product is denoted by $\langle f, g\rangle_K$ and the norm by $\|f\|_K = \sqrt{\langle f, f\rangle_K}$.

When we write sums over multi-indices given a subset of dimensions $\mathfrak{u} \subseteq \{1, \ldots, d\}$ like $\sum_{\boldsymbol{\tau}_\mathfrak{u} \in \mathcal{G}^{|\mathfrak{u}|}} A(\boldsymbol{\tau})$ with $\mathcal{G}$ some one-dimensional set, then this should be read as $\sum_{\boldsymbol{\tau} \in \mathbb{N}^d,\, \tau_j \in \mathcal{G} \text{ for } j \in \mathfrak{u}, \tau_j = 0 \text{ for } j \notin \mathfrak{u}} A(\boldsymbol{\tau})$ such that for $\mathfrak{u} = \emptyset$ this sum has a single term with all $\tau_j = 0$.

Chapter 2

Preliminaries

In this chapter, we give the background theory of the basic concepts used in the following chapters where more particular problems are solved. We particularly consider the function spaces on the d-dimensional unit cube $[0,1]^d$.

2.1 Reproducing kernel Hilbert spaces

Throughout this thesis, we always work in a reproducing kernel Hilbert space (RKHS).

Definition 2.1. *A reproducing kernel Hilbert space $\mathcal{H}(K)$ is a Hilbert space with a kernel function $K : [0,1]^d \times [0,1]^d \to \mathbb{R}$ which satisfies*

- *for any fixed $\boldsymbol{y} \in [0,1]^d$, $K(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{H}(K)$,*
- *for any fixed $\boldsymbol{y} \in [0,1]^d$ and any $f \in \mathcal{H}(K)$, $f(\boldsymbol{y}) = \langle f, K(\cdot, \boldsymbol{y}) \rangle_K$,*

where $\langle \cdot, \cdot \rangle_K$ is the inner product in $\mathcal{H}(K)$.

RKHSs possess many useful properties. To explain these, we need one fundamental theorem.

Theorem 2.1 (Riesz representation theorem)**.** *For any continuous linear functional $\xi \in \mathcal{H}^*$, where we denote by $\mathcal{H}^*$ the dual space of $\mathcal{H}$ consisting*

of all continuous linear functionals from $\mathcal{H}$ to $\mathbb{R}$, there exists a unique $u \in \mathcal{H}$ which satisfies

$$\xi(f) = \langle f, u \rangle_{\mathcal{H}},$$

for all $f \in \mathcal{H}$, and for which

$$\|\xi\|_{\mathcal{H}^*} := \sup_{0 \neq f \in \mathcal{H}} \frac{|\xi(f)|}{\|f\|_{\mathcal{H}}} = \|u\|_{\mathcal{H}}.$$

The element $u \in \mathcal{H}$ is called the representer of $\xi \in \mathcal{H}^*$.

Proof. We refer to [76, Theorem 4.12]. □

By using Theorem 2.1, the unique existence of a reproducing kernel can be proved if and only if the function evaluation $f(\boldsymbol{x})$ for any $\boldsymbol{x} \in [0,1]^d$ and $f \in \mathcal{H}(K)$ is a continuous linear functional (see [2]).

Another useful property which we will use in Chapter 6 is the following: for a RKHS $\mathcal{H}(K) \subset \mathcal{H}'$ where $\mathcal{H}'$ is a Hilbert space, the function

$$g(y) = \langle f, K(\cdot, y) \rangle_{\mathcal{H}'}$$

is the orthogonal projection of $f \in \mathcal{H}'$ onto $\mathcal{H}(K)$ meaning that

$$\langle f - g, K(\cdot, y) \rangle_{\mathcal{H}'} = 0,$$

see [2, Section 2, (6)]. Conversely, if we have two complementary RKHSs $\mathcal{H}(K')$ and $\mathcal{H}(K'')$, then the direct sum of those two spaces is also a RKHS with the reproducing kernel $K = K' + K''$. For more interesting properties of RKHSs, we refer to [2].

We give several examples of RKHSs which are often used in the QMC literature. We first note that the term "Sobolev space" refers to spaces in which derivatives appear in the inner product.

Unanchored Sobolev space

The unanchored Sobolev space is a Sobolev space in which certain variables are integrated out. Here we define $\mathcal{H}(K_{\alpha,d}^{\mathrm{Sob}})$ as the function space where the inner

product is given by

$$\langle f, g\rangle_{K^{\mathrm{Sob}}_{\alpha,d}} := \sum_{\substack{\boldsymbol{\tau}\in\{0,\dots,\alpha\}^d \\ \mathfrak{w}=\{j:\tau_j=\alpha\}}} \int_{[\mathbf{0}_{\mathfrak{w}},\mathbf{1}_{\mathfrak{w}}]} \left(\int_{[\mathbf{0}_{-\mathfrak{w}},\mathbf{1}_{-\mathfrak{w}}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{w}}\right) \left(\int_{[\mathbf{0}_{-\mathfrak{w}},\mathbf{1}_{-\mathfrak{w}}]} g^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{w}}\right) \mathrm{d}\boldsymbol{x}_{\mathfrak{w}},$$

where

$$f^{(\boldsymbol{\tau})} := \frac{\partial^{\tau_1+\cdots+\tau_d} f}{\partial x_1^{\tau_1}\cdots\partial x_d^{\tau_d}},$$

and $[\boldsymbol{a}_{\mathfrak{w}}, \boldsymbol{b}_{\mathfrak{w}}] := \prod_{j\in\mathfrak{w}}[a_j, b_j]$, $[\boldsymbol{a}_{-\mathfrak{w}}, \boldsymbol{b}_{-\mathfrak{w}}] := \prod_{j\notin\mathfrak{w}}[a_j, b_j]$, and likewise for $\boldsymbol{x}_{\mathfrak{w}}$ and $\boldsymbol{x}_{-\mathfrak{w}}$. This space is used, e.g., in [4].

This unanchored Sobolev space is one kind of *ANOVA* (analysis of variance) space in which functions are decomposable in the following manner:

$$f(\boldsymbol{x}) = f_{\emptyset} + \sum_{\emptyset\neq\mathfrak{u}\subseteq\{1,2,\dots,d\}} f_{\mathfrak{u}}(\boldsymbol{x}_{\mathfrak{u}}), \tag{2.1}$$

where $f_{\emptyset} := \int_{[0,1]^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ and

$$f_{\mathfrak{u}}(\boldsymbol{x}_{\mathfrak{u}}) := \int_{[0,1]^{d-|\mathfrak{u}|}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{u}} - \sum_{\mathfrak{v}\subset\mathfrak{u}} f_{\mathfrak{v}}(\boldsymbol{x}_{\mathfrak{v}}),$$

and the norm is also decomposable in the same manner:

$$\begin{aligned}\|f(\boldsymbol{x})\|^2_{K^{\mathrm{Sob}}_{\alpha,d}} &= |f_{\emptyset}|^2 + \sum_{\emptyset\neq\mathfrak{u}\subseteq\{1,2,\dots,d\}} \|f_{\mathfrak{u}}(\boldsymbol{x}_{\mathfrak{u}})\|^2_{K^{\mathrm{Sob}}_{\alpha,|\mathfrak{u}|}} \\ &= \sum_{\mathfrak{u}\subseteq\{1,2,\dots,d\}} \|f_{\mathfrak{u}}(\boldsymbol{x}_{\mathfrak{u}})\|^2_{K^{\mathrm{Sob}}_{\alpha,|\mathfrak{u}|}}.\end{aligned} \tag{2.2}$$

We refer to [42, Example 4.3] for the proof of this norm decomposition. We remark that the decomposition of the function (2.1) is always possible for functions $f \in L_2([0,1]^d)$. However, the decomposition of the norm (2.2) does not always hold in general for a normed space $\mathcal{H} \subset L_2$.

As we will see in Chapter 6, functions $f \in \mathcal{H}(K^{\mathrm{Sob}}_{\alpha,d})$ can be alternatively represented by

$$f(\boldsymbol{x}) = \sum_{\mathfrak{u}\subseteq\{1,\dots,d\}} \sum_{\boldsymbol{\tau}_\mathfrak{u}\in\{0,\dots,\alpha\}^{|\mathfrak{u}|}} (-1)^{d-|\mathfrak{u}|} \left(\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right) \int_{[0,1]^d} f^{(\boldsymbol{\tau}_\mathfrak{u},\boldsymbol{\alpha}_{-\mathfrak{u}})}(\boldsymbol{y}) \left(\prod_{j\notin\mathfrak{u}} \frac{\widetilde{B}_\alpha(x_j-y_j)}{\alpha!}\right) \mathrm{d}\boldsymbol{y},$$

where $B_\tau(x)$ is the τth Bernoulli polynomial and $\widetilde{B}_\tau(x)$ is the 1-periodic extension of the τth Bernoulli polynomial.

As stated before, this space is a RKHS, and the reproducing kernel is given by

$$\begin{aligned} K^{\mathrm{Sob}}_{\alpha,d}(\boldsymbol{x},\boldsymbol{y}) &= \prod_{j=1}^{d} K^{\mathrm{Sob}}_{\alpha}(x_j,y_j) \\ &= \prod_{j=1}^{d} \left(1 + \sum_{\tau_j=1}^{\alpha} \frac{B_{\tau_j}(x_j)}{\tau_j!}\frac{B_{\tau_j}(y_j)}{\tau_j!} + (-1)^{\alpha+1}\frac{\widetilde{B}_{2\alpha}(x_j-y_j)}{(2\alpha)!}\right). \end{aligned}$$

Anchored Sobolev space

The anchored Sobolev space $\mathcal{H}(K^{\mathrm{ASob}}_{\alpha,d,\boldsymbol{c}})$ is another kind of Sobolev space which is not ANOVA space, meaning this space does not satisfy (2.2). The inner product is given by

$$\langle f,g\rangle_{K^{\mathrm{ASob}}_{\alpha,d,\boldsymbol{c}}} := \sum_{\substack{\boldsymbol{\tau}\in\{0,\dots,\alpha\}^d \\ \mathfrak{w}=\{j:\tau_j=\alpha\}}} \int_{[\mathbf{0}_\mathfrak{w},\mathbf{1}_\mathfrak{w}]} \left(f^{(\boldsymbol{\tau})}(\boldsymbol{x}_\mathfrak{w};\boldsymbol{c}_{-\mathfrak{w}})\right)\left(g^{(\boldsymbol{\tau})}(\boldsymbol{x}_\mathfrak{w};\boldsymbol{c}_{-\mathfrak{w}})\right) \mathrm{d}\boldsymbol{x}_\mathfrak{w},$$

where the fixed point $\boldsymbol{c} \in [0,1]^d$ is called the anchor point. The reproducing kernel is then given by $K^{\mathrm{ASob}}_{\alpha,d,\boldsymbol{c}}(\boldsymbol{x},\boldsymbol{y}) = \prod_{j=1}^{d} K^{\mathrm{ASob}}_{\alpha,c_j}(x_j,y_j)$ where

$$K^{\mathrm{ASob}}_{\alpha,c_j}(x_j,y_j) = \begin{cases} \sum_{\tau=1}^{\alpha-1} \frac{(x_j-c_j)^\tau}{\tau!}\frac{(y_j-c_j)^\tau}{\tau!} + \int_{c_j}^{\min(x_j,y_j)} \frac{(x_j-t)^{\alpha-1}}{(\alpha-1)!}\frac{(y_j-t)^{\alpha-1}}{(\alpha-1)!}\,\mathrm{d}t, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{for } x_j,y_j > c_j \\ \sum_{\tau=1}^{\alpha-1} \frac{(c_j-x_j)^\tau}{\tau!}\frac{(c_j-y_j)^\tau}{\tau!} + \int_{\min(x_j,y_j)}^{c_j} \frac{(t-x_j)^{\alpha-1}}{(\alpha-1)!}\frac{(t-y_j)^{\alpha-1}}{(\alpha-1)!}\,\mathrm{d}t, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{for } x_j,y_j < c_j \\ \sum_{\tau=1}^{\alpha-1} \frac{(x_j-c_j)^\tau}{\tau!}\frac{(y_j-c_j)^\tau}{\tau!}, \qquad \text{otherwise.} \end{cases}$$

For details and applications of this space, we refer to [92, Section 1] and [55, 56].

Korobov space

The Korobov space is a RKHS of Fourier series, which is alternatively called *a periodic Sobolev space of dominating mixed smoothness*. For a function in the Wiener algebra

$$A(\mathbb{T}^d) := \left\{ f : \sum_{\boldsymbol{h}\in\mathbb{Z}^d} |\widehat{f}(\boldsymbol{h})| < \infty \right\}$$

where

$$\widehat{f}(\boldsymbol{h}) = \int_{[0,1]^d} f(\boldsymbol{x}) \exp(2\pi \mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},$$

we define the inner product of the Korobov space by

$$\langle f,g\rangle_{K^{\mathrm{Kor}}_{\alpha,d}} := \sum_{\boldsymbol{h}\in\mathbb{Z}^d} \widehat{f}(\boldsymbol{h})\,\overline{\widehat{g}(\boldsymbol{h})}\;\left[r_{\alpha,d}(\boldsymbol{h})\right]^2,$$

where $[r_{\alpha,d}(\boldsymbol{h})]^2 = \prod_{\substack{j=1\\ h_j\neq 0}}^{d} |2\pi\, h_j|^{2\alpha}$ for Chapter 6. For this case, the reproducing kernel is given by

$$K^{\mathrm{Kor}}_{\alpha,d}(\boldsymbol{x},\boldsymbol{y}) = \prod_{j=1}^{d}\left(1 + (-1)^{\alpha+1}\frac{\widetilde{B}_{2\alpha}(x_j-y_j)}{(2\alpha)!}\right).$$

The details of this space and the relation to the unanchored Sobolev space will be explained in Chapter 6. We remark that in Chapters 3–5, we are going to use slightly modified Korobov spaces which we will denote by $E_\alpha(\mathbb{T}^d)$ with $[r_{\alpha,d}(\boldsymbol{h})]^2 = \prod_{j=1}^{d} \max(|h_j|^{2\alpha},1)$, for the sake of notational simplicity.

Half-period cosine space

The half-period cosine space is a non-periodic space where all functions in this space can be represented by half-period cosine series

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{Z}_+^d} \widetilde{f}(\boldsymbol{k})\ \phi_{\boldsymbol{k}}(\boldsymbol{x}),$$

where

$$\phi_{\boldsymbol{k}}(\boldsymbol{x}) := \sqrt{2}^{\|\boldsymbol{k}\|_0} \prod_{j=1}^{d} \cos(\pi k_j x_j), \quad \widetilde{f}(\boldsymbol{k}) := \int_{[0,1]^d} f(\boldsymbol{x})\ \phi_{\boldsymbol{k}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x},$$

and $\|\boldsymbol{k}\|_0$ is the number of non-zero components of $\boldsymbol{k}$. Similar to the Korobov space, the inner product in this space is defined by

$$\langle f, g \rangle_{K_{\alpha,d}^{\mathrm{Cos}}} := \sum_{\boldsymbol{k} \in \mathbb{Z}_+^d} \widetilde{f}(\boldsymbol{k})\ \overline{\widetilde{g}(\boldsymbol{k})}\ \left[r_{\alpha,d}(\boldsymbol{k})\right]^2,$$

Then the reproducing kernel is given by the following Mercer form:

$$K_{\alpha,d}^{\mathrm{Cos}} = \sum_{\boldsymbol{k} \in \mathbb{Z}_+^d} \phi_{\boldsymbol{k}}(\boldsymbol{x})\ \phi_{\boldsymbol{k}}(\boldsymbol{y})\ \left[r_{\alpha,d}(\boldsymbol{k})\right]^{-2},$$

and if $\left[r_{\alpha,d}(\boldsymbol{k})\right]^2 = \prod_{\substack{j=1 \\ h_j \neq 0}}^{d} |2\pi\, k_j|^{2\alpha}$, the reproducing kernel has the closed form

$$\begin{aligned}
K_{\alpha,d}^{\mathrm{Cos}}(\boldsymbol{x}, \boldsymbol{y}) &= \sum_{\boldsymbol{k} \in \mathbb{Z}_+^d} \phi_{\boldsymbol{k}}(\boldsymbol{x})\ \phi_{\boldsymbol{k}}(\boldsymbol{y}) \prod_{\substack{j=1 \\ h_j \neq 0}}^{d} |2\pi\, k_j|^{-2\alpha} \\
&= \prod_{j=1}^{d} \left(1 + \sum_{k_j \in \mathbb{Z}_+} \frac{\sqrt{2}\cos(\pi k_j x_j)}{(2\pi k_j)^{\alpha}}\ \frac{\sqrt{2}\cos(\pi k_j y_j)}{(2\pi k_j)^{\alpha}} \right) \\
&= \prod_{j=1}^{d} \left(1 + \sum_{k_j \in \mathbb{Z}_+} \frac{\cos(\pi k_j (x_j + y_j)) + \cos(\pi k_j (x_j - y_j))}{(2\pi k_j)^{2\alpha}} \right) \\
&= \prod_{j=1}^{d} \left(1 + (-1)^{\alpha+1} \widetilde{B}_{2\alpha}\left(\frac{x_j + y_j}{2} \right) + (-1)^{\alpha+1} \widetilde{B}_{2\alpha}\left(\frac{x_j - y_j}{2} \right) \right).
\end{aligned} \tag{2.3}$$

This space was first introduced by [20]. Therein, it was shown that $\mathcal{H}(K_{1,d}^{\mathrm{Sob}})$ and $\mathcal{H}(K_{1,d}^{\mathrm{Cos}})$ are the same, with the choice of $r_{\alpha,d}$ here.

2.2 Lattice points for integration

In this section we introduce rank-1 lattice point sets which we will denote by $\Lambda(\boldsymbol{z}, n)$. For a given $\boldsymbol{z} \in \mathbb{Z}^d$ and $n \in \mathbb{N}$, the point set is defined as

$$\Lambda(\boldsymbol{z}, n) := \left\{ \boldsymbol{p}_i := \frac{\boldsymbol{z}i}{n} \bmod 1 : i \in \mathbb{Z}_n \right\}.$$

The vector $\boldsymbol{z}$ is called *the generating vector*. This vector determines the quality of the method using these lattice points.

We want to approximate the following integral

$$I(f) := \int_{[0,1]^d} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$$

by the finite sum

$$Q(f; \Lambda(\boldsymbol{z}, n)) := \frac{1}{n} \sum_{i=1}^{n} f(\boldsymbol{p}_i),$$

If the integrand can be represented by absolutely convergent Fourier series

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{h} \in \mathbb{Z}^d} \widehat{f}(\boldsymbol{h}) \exp(2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{x}), \quad \widehat{f}(\boldsymbol{h}) := \int_{[0,1]^d} f(\boldsymbol{x}) \exp(-2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{x}) \, \mathrm{d}\boldsymbol{x},$$

then the integration error can be written as

$$\begin{aligned} \int_{[0,1]^d} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - \frac{1}{n} \sum_{i=1}^{n} f(\boldsymbol{p}_i) &= \widehat{f}(\boldsymbol{0}) - \frac{1}{n} \sum_{i=1}^{n} \sum_{\boldsymbol{h} \in \mathbb{Z}^d} \widehat{f}(\boldsymbol{h}) \exp(2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{z} i / n) \\ &= \widehat{f}(\boldsymbol{0}) - \sum_{\boldsymbol{h} \cdot \boldsymbol{z} \equiv 0 \pmod{n}} \widehat{f}(\boldsymbol{h}) \\ &= - \sum_{\substack{\boldsymbol{h} \neq \boldsymbol{0} \\ \boldsymbol{h} \cdot \boldsymbol{z} \equiv 0 \pmod{n}}} \widehat{f}(\boldsymbol{h}). \end{aligned}$$

Due to this error expression, it is possible to obtain a small integration error for functions of which the Fourier coefficients decay fast in some sense. This decay of coefficients and the corresponding criterion for choosing a good generating vector $\boldsymbol{z}$ have been considered in different manners. Korobov in [39] introduced the following decay condition

$$|\widehat{f}(\boldsymbol{h})| \leq c \prod_{j=1}^{d} \max(1, |h_j|)^{-\alpha},$$

where the parameter α determines the speed of decay. Therein, the corresponding error criterion is chosen to be

$$P_\alpha(\boldsymbol{z}, n) := \sum_{\substack{\boldsymbol{h} \neq \boldsymbol{0} \\ \boldsymbol{h}\cdot\boldsymbol{z} \equiv 0 \pmod n}} \prod_{j=1}^{d} \max(1, |h_j|)^{-\alpha},$$

because the integration error can be bounded by

$$\left| \int_{[0,1]^d} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - \frac{1}{n} \sum_{i=1}^{n} f(\boldsymbol{p}_i) \right| \leq c P_\alpha(\boldsymbol{z}, n).$$

Hence finding good generating vectors to obtain small $P_\alpha(\boldsymbol{z}, n)$ is an important task. Bounds on $P_\alpha(\boldsymbol{z}, n)$ and existence of such lattice rules were well studied in literature, e.g., [23, 59–62].

One important property is that for a reproducing kernel Hilbert space $\mathcal{H}(K)$ with reproducing kernel $K(\boldsymbol{x}, \boldsymbol{y})$, the worst case integration error (the operator norm of the cubature error in the space) is represented by

$$\begin{aligned} wce^2(\mathcal{H}(K), \{\boldsymbol{p}_i\}_{i=1}^n) &:= \left(\sup_{0 \neq f \in \mathcal{H}(K)} \frac{\left| \int_{[0,1]^d} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - \frac{1}{n} \sum_{i=1}^n f(\boldsymbol{p}_i) \right|}{\|f\|_K} \right)^2 \\ &= \int_{[0,1]^d} \int_{[0,1]^d} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{y} - \frac{2}{n} \sum_{i=1}^{n} K(\boldsymbol{x}, \boldsymbol{p}_i) \\ &\quad + \frac{1}{n^2} \sum_{i=1}^{n} \sum_{i'=1}^{n} K(\boldsymbol{p}_i, \boldsymbol{p}_{i'}). \end{aligned}$$

This is very useful for theoretical analysis and numerical experiments especially when the reproducing kernel has a closed form.

Random shift

A random shift is one way of randomizing lattice points. Let Δ be a randomly chosen point from the uniform distribution on $[0, 1]^d$. By adding this random shift to all lattice points and taking modulo 1, we obtain the randomly shifted lattice points:

$$\Lambda(\boldsymbol{z}, \Delta, n) := \left\{ \frac{\boldsymbol{z} i}{n} + \Delta \bmod 1 : i \in \mathbb{Z} \right\}.$$

A randomly shifted lattice rule is one kind of RQMC methods, which gives us an unbiased variance estimate of the error using n_{sh} independently chosen shifts $\Delta_1, ..., \Delta_{n_{sh}}$. First we see that $Q(f; \Lambda(\boldsymbol{z}, \Delta, n))$ is an unbiased estimate of $I(f)$;

$$\begin{aligned}\mathbb{E}^{\Delta}\left[Q(f;\Lambda(\boldsymbol{z},\Delta,n))\right] &= \int_{[0,1]^d} \frac{1}{n}\sum_{i=1}^{n} f(\{\boldsymbol{p}_i+\Delta\})\,\mathrm{d}\Delta \\ &= \int_{[0,1]^d} \frac{1}{n}\sum_{i=1}^{n} f(\{\Delta\})\,\mathrm{d}\Delta \\ &= \int_{[0,1]^d} f(\{\Delta\})\,\mathrm{d}\Delta = I(f), \end{aligned} \tag{2.4}$$

where $\{x\}$ is the fractional part of x. Then the unbiased variance estimate of $Q(f; \Lambda(\boldsymbol{z}, \Delta, n))$ is given by

$$\bar{\sigma}^2 := \frac{1}{n_{sh}-1}\sum_{i=1}^{n_{sh}}\left(Q(f;\Lambda(\boldsymbol{z},\Delta_i,n)) - \frac{1}{n_{sh}}\sum_{i'=1}^{n_{sh}} Q(f;\Lambda(\boldsymbol{z},\Delta_{i'},n))\right)^2,$$

where the unbiasedness

$$\mathbb{E}^{\Delta_1,...,\Delta_{n_{sh}}}[\bar{\sigma}^2] = \mathbb{E}\left[\int_{[0,1]^d}(Q(f;\Lambda(\boldsymbol{z},\Delta,n)) - I(f))^2\,\mathrm{d}\Delta\right] =: \sigma^2$$

can be shown by using (2.4). Hence the unbiased estimate of the variance of the error using n_{sh} shifts is

$$\mathbb{E}^{\Delta_1,...,\Delta_{n_{sh}}}\left[\left(\frac{1}{n_{sh}}\sum_{i'=1}^{n_{sh}} Q(f;\Lambda(\boldsymbol{z},\Delta_{i'},n)) - I(f)\right)^2\right] = \frac{\bar{\sigma}^2}{n_{sh}}.$$

Here we remark, often it is confused that $\bar{\sigma}$ is the unbiased estimator of the standard deviation σ, but this is not true in general;

$$\mathbb{E}^{\Delta_1,...,\Delta_{n_{sh}}}[\bar{\sigma}] \neq \sigma.$$

For more details of randomly shifted lattice rules, we refer to [34, 79].

Tent transformation

For non-periodic integrands, often lattice rules are used with the tent transformation, alternatively called baker's transformation (see also Figure 2.1)

$$t(x) = 1 - |2x-1|, \quad t(\boldsymbol{x}) = (t(x_1), ..., t(x_d)).$$

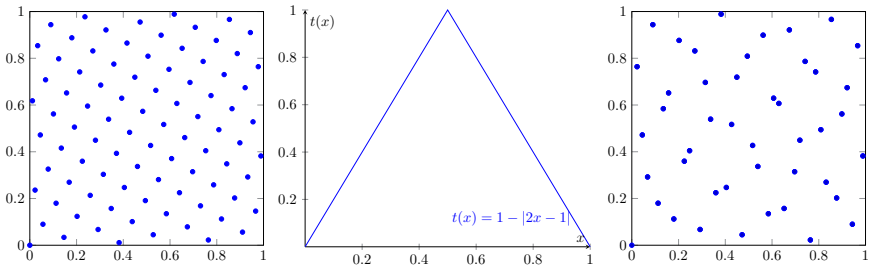

Figure 2.1: An example of a rank-1 lattice (left), the tent transform (middle) and the tent-transformed lattice (right), where $n = 89$ and $\boldsymbol{z}^\top = (1, 55)$.

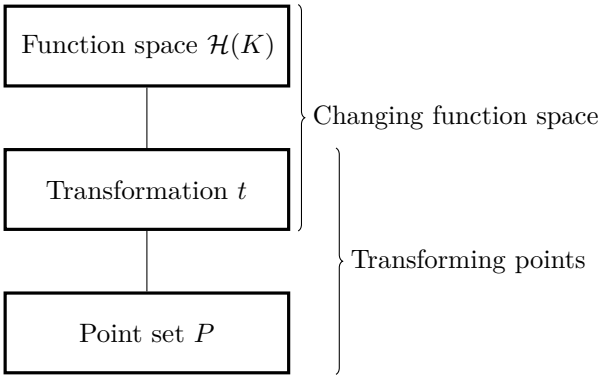

Figure 2.2: Two ways of seeing the transformation

This use of tent transformation for lattice rules was first introduced in [31], where first order convergence for $\mathcal{H}(K^{\mathrm{Sob}}_{1,d})$ and second order convergence for $\mathcal{H}(K^{\mathrm{Sob}}_{2,d})$ are achieved together with random shifts. In [20], αth order convergence for $\mathcal{H}(K^{\mathrm{Cos}}_{\alpha,d})$ is achieved with tent-transformed lattice rules but without random shifts. Due to this result combined with the fact that $\mathcal{H}(K^{\mathrm{Sob}}_{1,d})$ and $\mathcal{H}(K^{\mathrm{Cos}}_{1,d})$ are the same, first order convergence for $\mathcal{H}(K^{\mathrm{Sob}}_{1,d})$ with tent-transformed but without random shifting is proved. After [20, 31], it was shown in [24] that second order convergence for $\mathcal{H}(K^{\mathrm{Sob}}_{2,d})$ can be achieved without random shifting.

It is also beneficial to see it as change of an underlying function space (Figure 2.2). For instance, when both the tent transformation and random shift are applied to lattice points, one needs to apply the random shift first and then the tent transform. This order cannot be reversed to obtain second order convergence for $\mathcal{H}(K^{\mathrm{Sob}}_{2,d})$. To understand this, seeing the tent transformation as changing of

function spaces is useful. The tent transform makes the non-periodic function space periodic in a certain sense. Therefore one needs to use randomly shifted lattice rules, which requires periodicity to obtain second-order convergence, after the function has been made periodic. In general, applying some transformation and getting periodic functions from non-periodic ones is called a periodization strategy (see [46] and the references therein). Even though it is pointed out that the periodization strategy can be hugely cursed by dimensionality [7, 46], it is still a strong tool for lower dimensional problems.

2.3 Lattice points for function approximation

Function approximation is a different problem from integration but closely related. From the perspective of information based complexity, this is a problem of approximating a linear bounded operator, whereas integration is a problem of approximating a linear bounded functional [65, 66]. This difference makes it difficult to obtain worst case error expression using the reproducing kernels for function approximation, because the Riesz representation theorem is only applicable for linear bounded functionals.

To formulate the problem of function approximation on lattice points, let us consider a function f on a d-dimensional torus $\mathbb{T}^d$ which we identify with $[0,1]^d$. We wish to recover function f with limited number of function evaluations by the following equal weight linear combination:

$$f_a(\boldsymbol{x}) := \sum_{\boldsymbol{h}\in\mathcal{A}} \widehat{f}_a(\boldsymbol{h}) \exp(2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{x}), \quad \widehat{f}_a(\boldsymbol{h}) := \frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{p}_i)\exp(-2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{p}),$$

where $\mathcal{A}$ is an index set comprising frequency vectors in $\mathbb{Z}^d$. The goal is to minimize the approximation error

$$\|f - f_a\|$$

with some norm $\|\cdot\|$ depending on the context, e.g., L_2, L_p, or some Sobolev-type norm (see [9]).

In [9], it is shown that the lattice points can only achieve *the half rate* $\mathcal{O}(n^{-\alpha/2})$ convergence instead of the best possible rate $\mathcal{O}(n^{-\alpha})$ for functions in $\mathcal{H}(K^{\mathrm{Kor}}_{\alpha,d})$. Nevertheless, the flexible structure of lattice points is still attractive for actual applications.

Tent transformation

Tent-transformed lattices are also used for function approximation; see [43, 84]. Tent-transformed lattices are especially studied for approximating specific non-periodic functions by half-period cosine series. Interestingly, those half-period cosine functions can be mapped to Chebyshev functions, and therefore tent-transformed lattices also can be used for function approximation with Chebyshev series [43, 73].

Interpolation on lattices

Often, function approximation with rank-1 lattices uses oversampling, i.e., approximating a function f by

$$f_a(\boldsymbol{x}) := \sum_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)} \widehat{f}_a(\boldsymbol{h})\, \exp(2\pi \mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x})$$

where

$$\widehat{f}_a(\boldsymbol{h}) := \frac{1}{n} \sum_{\boldsymbol{p}\in\Lambda(\boldsymbol{z},n)} f(\boldsymbol{p})\, \exp(-2\pi \mathrm{i}\boldsymbol{h}\cdot\boldsymbol{p}),$$

and $\mathcal{A}(\boldsymbol{z},n)$ is some predetermined frequency set with $n > |\mathcal{A}(\boldsymbol{z},n)|$. This means the number of sampling points is taken larger than the number of basis functions to approximate the function, and the interpolation condition,

$$f_a(\boldsymbol{p}) = f(\boldsymbol{p}), \quad \text{for all } \boldsymbol{p}\in\Lambda(\boldsymbol{z},n),$$

is not satisfied in general. Depending on the aim of approximating a function, this interpolation condition can be a desirable property. For instance, pseudospectral methods need this condition. In Chapters 3–5, we will use the pseudospectral method on lattices where we call the index set $\mathcal{A}(\boldsymbol{z},n)$ *an anti-aliasing set* if the condition

$$(\boldsymbol{h}-\boldsymbol{h}')\cdot\boldsymbol{z} \not\equiv 0 \pmod{n},$$

is satisfied for all distinct vectors $\boldsymbol{h},\boldsymbol{h}'\in\Lambda(\boldsymbol{z},n)$. With the condition $|\mathcal{A}(\boldsymbol{z},n)| = n$, we essentially study interpolation on lattices.

From a group-theoretical view, the pair of a rank-1 lattice and the corresponding anti-aliasing set is called the Pontryagin dual. This means that the Fourier transform of each set is an isomorphism of the other, and therefore taking the Fourier transform twice of each set gives an isomorphism of itself (see Figure 2.3 for examples of corresponding dual sets).

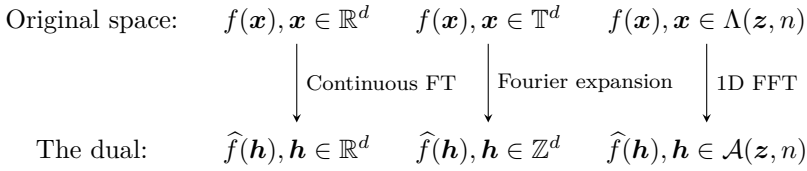

Figure 2.3: Fourier transforms and dual spaces on different abelian groups

There exist a few references on interpolation on lattices. In [50], a spectral method on lattices is introduced in a general setting. The choice of a frequency index set $\mathcal{A}$ is numerically studied in [53] to solve Poisson equations. One of the earliest application is in computer graphics [17]. However, except for the listed papers above, there are not that many references. Upon finalizing this thesis, we learned that the use of lattice points for the time-dependent Schrödinger equation was suggested in the section of *conclusions and related work* in [53]. We will show the effect of using interpolating lattices in Chapter 3, where the results are much better than those obtained by a sparse grid based method in [26].

Chapter 3

Strang splitting in combination with rank-1 and rank-r lattices for the time-dependent Schrödinger equation

In this chapter, we consider the time-dependent Schrödinger equation in a multidimensional setting. We approximate the solution for the time dependent Schrödinger equation (TDSE) in two steps. We first use a pseudospectral collocation method that uses samples of the functions on rank-1 or rank-r lattice points with unitary Fourier transforms. We then get a system of ordinary differential equations in time, which we solve approximately by stepping in time using the exponential operator splitting methods (Strang splitting or higher-order splitting). For the Strang splitting case, we prove that the numerical scheme proposed converges quadratically with respect to the time step size, given that the potential functions are in a Korobov space with smoothness parameter greater than 9/2. This result is then generalized to higher order splittings for obtaining p-th order convergence given $2p + 1/2$ Korobov smoothness. Particularly, we prove that the required degree of smoothness is independent of the dimension of the problem. We demonstrate our new method by comparing with results using sparse grids from [26], with several numerical examples showing large advantage for our new method and pushing the examples to

higher dimensionality. The proposed method has two distinctive features from a numerical perspective: (i) numerical results show the error convergence of time discretization is consistent even for higher-dimensional problems; (ii) by using the rank-1 lattice points, the solution can be efficiently computed (and further time stepped) using only 1-dimensional Fast Fourier Transforms. The contents of this chapter is part of the published paper [88].

3.1 Introduction

Approximating the solution of the many-particle Schrödinger equation is a challenging problem, where the dimension of the problem increases linearly with the number of particles in the system. Many attempts have been made to break the curse of dimensionality with respect to this problem [25, 26, 33]. This is also the focus of the present chapter and we propose a numerical method which provides a partial solution to this. Often in the context of physics, the time-dependent Schrödinger equation (TDSE) is referred to as the following equation:

$$\mathrm{i}\,\hbar\,\frac{\partial\psi}{\partial t}(\boldsymbol{x},t) = -\frac{\hbar^2}{2m}\,\nabla^2\psi(\boldsymbol{x},t) + v(\boldsymbol{x})\,\psi(\boldsymbol{x},t),$$

where $\hbar$ is the reduced Planck constant and m is the mass. By scaling the time by $1/\sqrt{m}$ and setting $\gamma = \hbar/\sqrt{m}$ this is equivalent to the following form for which $\psi(\boldsymbol{x},t) = u(\boldsymbol{x},t/\sqrt{m})$. We therefore consider the following equivalent equation in this chapter (as was done in [25, 26, 33]):

$$\mathrm{i}\,\gamma\,\frac{\partial u}{\partial t}(\boldsymbol{x},t) = -\frac{\gamma^2}{2}\,\nabla^2 u(\boldsymbol{x},t) + v(\boldsymbol{x})\,u(\boldsymbol{x},t), \tag{3.1}$$

with positions $\boldsymbol{x} \in \mathbb{T}^d = \mathbb{T}([0,1)^d)$, time $t \in [0,T]$, $\gamma = \hbar/\sqrt{m} > 0$ a small positive parameter, i the imaginary unit and ∇^2 is the Laplace operator w.r.t. the positions $\boldsymbol{x}$, i.e., $\nabla^2 = \sum_{i=1}^{M}\sum_{j=1}^{D}\partial^2/\partial x_{i,j}^2$ where M is the number of particles and D is the physical dimensionality. For notational simplicity we set $d = M \times D$. The function $u(\boldsymbol{x},t)$ is the wave function which we seek to approximate, $v(\boldsymbol{x})$ the potential and $g(\boldsymbol{x})$ the initial condition at time $t = 0$; specific details about these functions will be covered in the later sections. In addition, the boundary conditions are assumed to be periodic. This periodic boundary makes the problem equivalent to identify the domain of $\boldsymbol{x}$ as the d-dimensional torus $\mathbb{T}^d = \mathbb{T}([0,1)^d) \simeq [0,1)^d$ with period 1. The TDSE in the above form appears in quantum mechanics and molecular chemistry, and is general enough to include the case of the quantum-mechanical harmonic oscillator, see, e.g., [5, 93]. We note that this form of equations can be interpreted in several ways: one-particle

in d-dimensional space; multiple d particles in one-dimensional space (e.g., [10]); and the combination of those two (multiple particles in multi-dimensional space, e.g., [38, 93]).

In [33], Jahnke and Lubich applied the Strang splitting method which is an operator splitting method, to approximate the solution of the TDSE where a collocation method using regular grids was first used to discretize the spatial dimensions of the initial wave function and the Strang splitting method was then applied to propagate the wave function in time. In [25, 26], sparse grids were used instead of regular grids to overcome the curse of dimensionality but with limited success. The numerical experiments on the TDSE were limited to dimension 5.

We are interested in using rank-1 lattices for function approximation. Lattice rules have traditionally been used for numerical integration of periodic functions, see, e.g., [12, 68, 80]. Rank-1 lattice rules have been studied for the integration of functions belonging to smooth permutation invariant function spaces in [70]. This research is also relevant to our work, since a system with identical particles admits to the setting where (groups of) coordinates, i.e., per particle, are permutation invariant. Additionally, lattice rules have been used for function approximation in recent years, e.g., [44, 45, 72]. A spectral collocation method using a rank-1 lattice was developed by [50] to approximate the solution of partial differential equations in a periodic space. In addition to the periodic setting, rank-1 lattices, after an appropriate transformation, were found suitable for integration and approximation of non periodic functions from smooth half-period cosine spaces (which includes the usual Sobolev space with bounded mixed first derivatives), see respectively [20] and [16, 84].

The above research motivates the use of rank-1 lattices for solving the TDSE where some symmetry is exhibited due to the physical nature, see [93]. We derive a spectral collocation method based on rank-1 and more general rank-r lattice rules. The general rank-r lattice points also include the (possibly anisotropic) regular grids. The computation of the involved spectral coefficients can be efficiently calculated using unitary Fast Fourier Transformations (FFTs) owing to the special structure of lattice points. Further, we conduct the error analysis of the numerical scheme. Our focus is on the error coming from the time discretization. The main theoretical result is that the error of the time-discretization converges with rate of $\mathcal{O}((\Delta t)^2)$ where Δt is the discretization step of the time $t \in [0, T]$. Our analysis shows that the convergence rate requires some smoothness of the potential function $v(\boldsymbol{x})$, but this smoothness does not depend on the dimension d, where the results in [26], where the collocation was done using sparse grids, need the smoothness to be higher when d increases. We provide numerical results in various settings, showing that the convergence rate against the time propagation is very stable and not affected by the dimension d.

3.2 The method

In this section, we will describe the numerical method used for solving the TDSE. First, we introduce the key concepts that are required throughout this chapter: lattice point sets, the Fourier pseudospectral method on lattices, and the Strang splitting.

3.2.1 Lattices

The main building blocks of the proposed method are *integration lattices*. They are the intersection of a lattice $A\,\mathbb{Z}^d$ with the unit cube $[0,1)^d$ where $A \in \mathbb{Q}^{d\times r}$, $1 \leq r \leq d$, is a rational matrix and were originally proposed to approximate periodic integrals on $[0,1)^d$. For more detailed information we refer to [12, 52, 80].

For the main part of this chapter we make use of a *rank*-1 lattice

$$\Lambda(\boldsymbol{z}, n) := \left\{ \frac{\boldsymbol{z}k}{n} \bmod 1 : k \in \mathbb{Z} \right\},$$

which is completely defined by its integer *generating vector* $\boldsymbol{z} \in \mathbb{Z}^d$ and the modulus n. We take the components of $\boldsymbol{z}$ relatively prime to n, such that the total number of points is n.

We derive theory for both rank-1 and rank-r lattices, enabling us to state all results for regular (anisotropic) grids as well since they can be represented by rank-r lattices. We therefore introduce the definition of a rank-r lattice, see [80].

Definition 3.1 (Canonical form of rank-r lattice)**.** *A d-dimensional integration lattice can be written in terms of a generator*

$$A = \begin{pmatrix} \boldsymbol{z}_1/n_1 & \boldsymbol{z}_2/n_2 & \cdots & \boldsymbol{z}_r/n_r \end{pmatrix} \in \mathbb{Q}^{d\times r},$$

which is specified by the generating vectors $\boldsymbol{Z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r) \in \mathbb{Z}^{d\times r}$ and moduli $\boldsymbol{n} = (n_1, \ldots, n_r) \in \mathbb{N}^r$, such that $A = \boldsymbol{Z}\,\mathrm{diag}(\boldsymbol{n})^{-1}$, with the corresponding lattice point set $\Lambda(\boldsymbol{Z}, \boldsymbol{n})$ given by

$$\Lambda(\boldsymbol{Z}, \boldsymbol{n}) := \left\{ A\boldsymbol{k} \bmod 1 : \boldsymbol{k} \in \mathbb{Z}^r \right\} \subset [0,1)^d.$$

This form is the canonical form of a rank-r lattice provided the moduli satisfy n_{i+1} divides n_i for $i = 1, \ldots, r-1$, the generating vectors $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r \in \mathbb{Z}^d$ are

linearly independent over the rational numbers and the components of each $\boldsymbol{z}_i$ *are relatively prime to* n_i*. Then* r *is the minimum number of generating vectors needed to describe this lattice point set and its total number of unique points in the unit cube* $[0,1)^d$ *is* $n = \prod_{i=1}^{r} n_i$*.*

For further details we refer to [80, Theorem 3.2] and the related part there. We interpret the collection of generating vectors $\boldsymbol{Z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r) \in \mathbb{Z}^{d\times r}$ as a matrix where the generating vectors constitute the columns of the matrix. The associated rank-r "lattice rule" is the equal-weight cubature rule to approximate an integral of a function f over the unit cube. For a rank-r lattice in its canonical form we can iterate over all points by a multiindex $\boldsymbol{k} \in \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_r}$ and therefore the cubature rule based on this lattice point set can be written as

$$Q(f; \boldsymbol{Z}, \boldsymbol{n}) := \frac{1}{n_1} \sum_{k_1=0}^{n_1-1} \cdots \frac{1}{n_r} \sum_{k_r=0}^{n_r-1} f\left(\left(\frac{\boldsymbol{z}_1 k_1}{n_1} + \cdots + \frac{\boldsymbol{z}_r k_r}{n_r} \right) \bmod 1 \right).$$

In this chapter we will always assume that a rank-r lattice is given in canonical form, i.e., $\boldsymbol{Z}$ and $\boldsymbol{n}$ satisfy the properties of Definition 3.1, and there is thus a one-to-one correspondence between the lattice points and the multiindex $\boldsymbol{k} \in \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_r}$. We also introduce an associated anti-aliasing index set for the rank-r lattice $\Lambda(\boldsymbol{Z}, \boldsymbol{n})$ which we will denote by $\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$. The anti-aliasing set is not unique.

Definition 3.2 (Anti-aliasing set)**.** *An anti-aliasing set* $\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n}) \in \mathbb{Z}^d$ *associated with the rank-*r *lattice* $\Lambda(\boldsymbol{Z}, \boldsymbol{n})$ *in canonical form has the property that for all distinct vectors* $\boldsymbol{h}, \boldsymbol{h}' \in \mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$ *it never holds that*

$$\boldsymbol{Z}^\top (\boldsymbol{h} - \boldsymbol{h}') \equiv \boldsymbol{0} \pmod{\boldsymbol{n}} \qquad \equiv \qquad \begin{cases} 0 \pmod{n_1}, \\ \vdots \qquad\qquad \vdots \\ 0 \pmod{n_r}, \end{cases}$$

where the equivalence is to be interpreted component-wise and $\boldsymbol{0}$ *is the* r*-dimensional zero-vector.*

The anti-aliasing condition states that every $\boldsymbol{h} \in \mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$ can be associated with a unique multiindex $\boldsymbol{\xi} \in \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_r}$, similarly as how we iterate over the points of the rank-r lattice. Therefore the maximum size of $\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$ is $n = \prod_{i=1}^{r} n_i$. Furthermore, if $|\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})| = n$ we can divide $\mathbb{Z}^d$ into conjugacy

classes with respect to $\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$ in the following three ways

$$\begin{aligned}\mathbb{Z}^d &= \biguplus_{\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{Z},\boldsymbol{n})} (\boldsymbol{h} + \mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})) \\ &= \biguplus_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \{\boldsymbol{h}' \in \mathbb{Z}^d : \boldsymbol{Z}^\top \boldsymbol{h}' \equiv \boldsymbol{Z}^\top \boldsymbol{h} \pmod{\boldsymbol{n}}\} \\ &= \biguplus_{\boldsymbol{\xi}\in\mathbb{Z}_{n_1}\oplus\cdots\oplus\mathbb{Z}_{n_r}} \{\boldsymbol{h} \in \mathbb{Z}^d : \boldsymbol{Z}^\top \boldsymbol{h} \equiv \boldsymbol{\xi} \pmod{\boldsymbol{n}}\},\end{aligned} \tag{3.2}$$

where $\uplus$ means that all sets are disjunct.

This set in the rank-1 case, $\mathcal{A}(\boldsymbol{z}, n)$, has been studied before, e.g., [13, 16, 50, 84]. It is sometimes also called a reconstructing rank-1 lattice, e.g., in [9, 36]. By using the concept of the *dual* of the lattice, defined by

$$\Lambda^\perp(\boldsymbol{Z}, \boldsymbol{n}) := \{\boldsymbol{h} \in \mathbb{Z}^d : \boldsymbol{Z}^\top \boldsymbol{h} \equiv \boldsymbol{0} \pmod{\boldsymbol{n}}\}, \tag{3.3}$$

where again the equivalence is to be interpreted component-wise with respect to the $\boldsymbol{n} = (n_1, \ldots, n_r)$, an anti-aliasing set can be equivalently defined as a set for which for all distinct $\boldsymbol{h}, \boldsymbol{h}' \in \mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$ we have

$$\boldsymbol{h} - \boldsymbol{h}' \notin \Lambda^\perp(\boldsymbol{Z}, \boldsymbol{n}).$$

In addition, we make extensive use of the *character property* of a rank-r lattice which is given by

$$\frac{1}{n} \sum_{\boldsymbol{p}\in\Lambda(\boldsymbol{Z},\boldsymbol{n})} \exp(2\pi\mathrm{i}(\boldsymbol{h}\cdot\boldsymbol{p})) = \begin{cases} 1 & \text{if } \boldsymbol{h} \in \Lambda^\perp(\boldsymbol{Z}, \boldsymbol{n}), \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

As an example of rank-1 lattice points and corresponding anti-aliasing set, we exhibit the case for $n = 55$ and $\boldsymbol{z}^\top = (1, 34)$ in Figure 3.1. Typically, pseudospectral Fourier methods use regular grids in the spatial domain, and the unitary Fourier transform maps these points to integer points in a hyper-rectangle in the frequency domain. Our method uses rank-r lattices instead of regular grids in the spatial domain. Similar to the typical pseudospectral Fourier methods, a unitary discrete Fourier transformation maps the lattice points to anti-aliasing integer point sets in the frequency domain, see Theorem 3.2.

3.2.2 The Fourier pseudospectral method on lattice point sets

A *pseudospectral method* is a way to approximate solutions of partial differential equations in terms of a finite number of basis functions. This was applied to

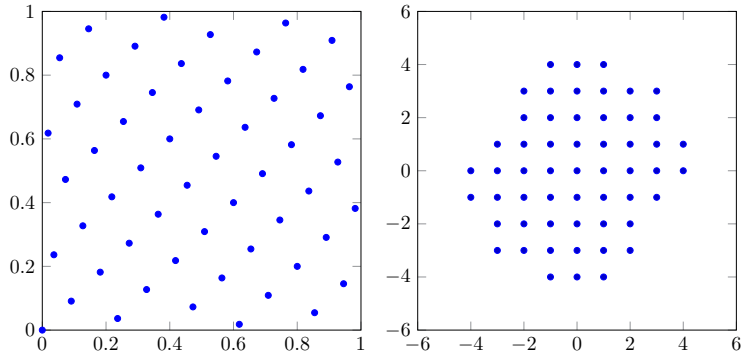

Figure 3.1: An example of rank-1 lattice and the corresponding anti-aliasing set with full cardinality, where $n = 55$ and $\boldsymbol{z}^\top = (1, 34)$.

approximate the solution of the TDSE in [26, 33] by expanding all functions into Fourier series. To apply the Fourier pseudospectral method, we require some properties. The minimum requirement we need is that any considered function is continuous and its Fourier series converges pointwise to the original function

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{h}\in\mathbb{Z}^d} \widehat{f}(\boldsymbol{h}) \exp(2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{x}), \qquad \text{for all } \boldsymbol{x}\in\mathbb{T}^d,$$

where the Fourier coefficients of f are given by

$$\widehat{f}(\boldsymbol{h}) = \int_{[0,1]^d} f(\boldsymbol{x}) \exp(-2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}.$$

This condition is guaranteed if we assume that the Fourier coefficients of the function f are absolutely summable

$$\|f\|_{A(\mathbb{T}^d)} := \sum_{\boldsymbol{h}\in\mathbb{Z}^d} |\widehat{f}(\boldsymbol{h})| < \infty.$$

The space of functions satisfying this condition is called the *Wiener algebra* $A(\mathbb{T}^d)$. For a detailed discussion, we refer to [27, Section 3.3].

To assure that the solution $u(\boldsymbol{x}, t) \in A(\mathbb{T}^d)$, we have the following lemma which makes use of the *Korobov space*, a reproducing kernel Hilbert space of Fourier series with a certain decay

$$E_\alpha(\mathbb{T}^d) := \left\{ f \in L_2(\mathbb{T}^d) : \|f\|^2_{E_\alpha(\mathbb{T}^d)} := \sum_{\boldsymbol{h}\in\mathbb{Z}^d} |\widehat{f}(\boldsymbol{h})|^2 \, r^2_\alpha(\boldsymbol{h}) < \infty \right\},$$

where

$$r_\alpha^2(\boldsymbol{h}) := \prod_{j=1}^{d} \max(|h_j|^{2\alpha}, 1). \tag{3.5}$$

The parameter $\alpha > 1/2$, the smoothness parameter, determines the rate of decay of the Fourier coefficients. For $\alpha > 1/2$ we have $E_\alpha(\mathbb{T}^d) \subset A(\mathbb{T}^d)$. This space is also referred to as a kind of unanchored periodic Sobolev space with dominating mixed-smoothness. In particular when $\alpha \in \mathbb{N}$ the norm can be expressed in terms of derivatives. Furthermore, if we define

$$(r_\alpha^*(\boldsymbol{h}))^2 := \prod_{j=1}^{d} (1 + |2\pi\, h_j|^{2\alpha}) \geq r_\alpha^2(\boldsymbol{h}),$$

then the associated norm defined as above is always larger than for r_α, and when $\alpha \in \mathbb{N}$ this norm then reads as

$$\|f\|^2_{E^*_\alpha(\mathbb{T}^d)} := \sum_{\boldsymbol{h}\in\mathbb{Z}^d} |\widehat{f}(\boldsymbol{h})|^2\, (r_\alpha^*(\boldsymbol{h}))^2 = \sum_{\boldsymbol{\tau}\in\{0,\alpha\}^d} \|D^{\boldsymbol{\tau}} f\|^2_{L_2(\mathbb{T}^d)}.$$

In fact this could be used as an alternative norm throughout this chapter. For a detailed discussion about Korobov spaces, see [64] and references therein. To assure that the term $\nabla^2 u$ in (3.1) makes sense, we require $\alpha \geq 2$. In Section 3.2.3, this space plays an important role to prove the convergence of our proposed method.

Lemma 3.1 (Regularity of solution and Fourier expansion)**.** *Given the TDSE* (3.1) *with $v, g \in E_\alpha(\mathbb{T}^d)$ and $\alpha \geq 2$, then the solution $u(\boldsymbol{x}, t) \in E_\alpha(\mathbb{T}^d)$ for all finite $t \geq 0$ and therefore*

$$u(\boldsymbol{x}, t) = \sum_{\boldsymbol{h}\in\mathbb{Z}^d} \widehat{u}(\boldsymbol{h}, t)\, \exp(2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{x}), \tag{3.6}$$

with

$$\mathrm{i}\,\gamma\, \widehat{u}'(\boldsymbol{h}, t) = 2\pi^2\gamma^2\, \|\boldsymbol{h}\|_2^2\, \widehat{u}(\boldsymbol{h}, t) + \widehat{f}(\boldsymbol{h}, t), \tag{3.7}$$

for all $\boldsymbol{h} \in \mathbb{Z}^d$, with $\widehat{u}'(\boldsymbol{h}, t) = (\partial/\partial t)\, \widehat{u}(\boldsymbol{h}, t)$ and $\widehat{f}(\boldsymbol{h}, t)$ the Fourier coefficients of $f(\boldsymbol{x}, t) := u(\boldsymbol{x}, t)\, v(\boldsymbol{x})$.

Proof. To prove that $u(\boldsymbol{x}, t) \in E_\alpha(\mathbb{T}^d)$, we first rewrite the TDSE (3.1),

$$\frac{\partial}{\partial t} u(\boldsymbol{x}, t) = \mathrm{i}\,\frac{\gamma}{2}\,\nabla^2 u(\boldsymbol{x}, t) - \frac{\mathrm{i}}{\gamma} v(\boldsymbol{x})\, u(\boldsymbol{x}, t) = Au(\boldsymbol{x}, t) + Bu(\boldsymbol{x}, t),$$

where $Au(\boldsymbol{x},t) = \mathrm{i}\,\frac{\gamma}{2}\,\nabla^2 u(\boldsymbol{x},t)$ and $Bu(\boldsymbol{x},t) = -\frac{\mathrm{i}}{\gamma} v(\boldsymbol{x}) u(\boldsymbol{x},t)$. We let $(\mathrm{e}^{At})_{t\geq 0}$ and $(\mathrm{e}^{Bt})_{t\geq 0}$ denote strongly continuous semigroups generated by A and B respectively. We note that the solution of (3.1), then, can be written as $u(\boldsymbol{x},t) = \mathrm{e}^{(A+B)\,t}\, g(\boldsymbol{x})$. Observe that e^{At} is unitary on the Korobov space $E_\alpha(\mathbb{T}^d)$, i.e., for any $t \geq 0$,

$$\begin{aligned}\|\mathrm{e}^{At} g(\boldsymbol{x})\|^2_{E_\alpha(\mathbb{T}^d)} &= \sum_{\boldsymbol{h}\in\mathbb{Z}^d} |\mathrm{e}^{\frac{\gamma}{2}\mathrm{i}\,\|\boldsymbol{h}\|_2^2\,t}\,\widehat{g}(\boldsymbol{h})|^2\; r_\alpha^2(\boldsymbol{h}) \\ &= \sum_{\boldsymbol{h}\in\mathbb{Z}^d} |\widehat{g}(\boldsymbol{h})|^2 r_\alpha^2(\boldsymbol{h}) = \|g(\boldsymbol{x})\|^2_{E_\alpha(\mathbb{T}^d)}.\end{aligned}$$

Also we know that the Korobov space is an algebra (see [67, Appendix 2]) such that for any $f, g \in E_\alpha(\mathbb{T}^d)$ also their product is in $E_\alpha(\mathbb{T}^d)$,

$$\|f\,g\|_{E_\alpha(\mathbb{T}^d)} \leq C_{d,\alpha}\,\|f\|_{E_\alpha(\mathbb{T}^d)}\,\|g\|_{E_\alpha(\mathbb{T}^d)},$$

where the constant $C_{d,\alpha} = 2^{d\alpha}(1+2\zeta(2\alpha))^{d/2}$. Hence this result holds for any $t \geq 0$ and $f(\boldsymbol{x}) = v(\boldsymbol{x})$ and $g = u(\boldsymbol{x},t)$. By using the Lie–Trotter product formula for unbounded self-adjoint operators (see e.g., [77, Theorem 6.4]), we obtain the following bound:

$$\begin{aligned}\left\|\mathrm{e}^{(A+B)t} u(\boldsymbol{x},0)\right\|_{E_\alpha(\mathbb{T}^d)} &= \left\|\lim_{n\to\infty}\left(\mathrm{e}^{A\frac{t}{n}}\mathrm{e}^{B\frac{t}{n}}\right)^n u(\boldsymbol{x},0)\right\|_{E_\alpha(\mathbb{T}^d)} \\ &\leq \lim_{n\to\infty}\left(\|\mathrm{e}^{A\frac{t}{n}}\|_{E_\alpha(\mathbb{T}^d)\to E_\alpha(\mathbb{T}^d)}\;\;\|\mathrm{e}^{B\frac{t}{n}}\|_{E_\alpha(\mathbb{T}^d)\to E_\alpha(\mathbb{T}^d)}\right)^n \|u(\boldsymbol{x},0)\|_{E_\alpha(\mathbb{T}^d)} \\ &\leq \lim_{n\to\infty}\left(\|\mathrm{e}^{B\frac{t}{n}}\|_{E_\alpha(\mathbb{T}^d)\to E_\alpha(\mathbb{T}^d)}\right)^n \|u(\boldsymbol{x},0)\|_{E_\alpha(\mathbb{T}^d)} \\ &\leq \lim_{n\to\infty}\left(\mathrm{e}^{\|B\|_{E_\alpha(\mathbb{T}^d)\to E_\alpha(\mathbb{T}^d)}\frac{t}{n}}\right)^n \|u(\boldsymbol{x},0)\|_{E_\alpha(\mathbb{T}^d)} \\ &\leq \mathrm{e}^{\frac{1}{\gamma}C_{d,\alpha}\|v\|_{E_\alpha(\mathbb{T}^d)}t}\|u(\boldsymbol{x},0)\|_{E_\alpha(\mathbb{T}^d)},\end{aligned}$$

where $\|V\|_{X\to X} := \sup_{0\neq u\in X}\|V(u)\|_X/\|u\|_X$ is the induced operator norm, and the above bound is finite for finite time t. Thus we have $u(\boldsymbol{x},t) \in E_\alpha(\mathbb{T}^d)$ for any finite $t > 0$.
By expanding the left hand side of (3.1), we have

$$\mathrm{i}\,\gamma\,\frac{\partial u}{\partial t} = \mathrm{i}\,\gamma\sum_{\boldsymbol{h}\in\mathbb{Z}^d}\widehat{u}'(\boldsymbol{h},t)\,\exp(2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}).$$

By also expanding the right hand side of (3.1), we obtain

$$\begin{aligned}&\mathrm{i}\,\gamma \sum_{\boldsymbol{h}\in\mathbb{Z}^d} \widehat{u}'(\boldsymbol{h},t)\,\exp(2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x})\\ &= \sum_{\boldsymbol{h}\in\mathbb{Z}^d} \Big(2\pi^2\gamma^2\,\|\boldsymbol{h}\|_2^2\,\widehat{u}(\boldsymbol{h},t) + \widehat{f}(\boldsymbol{h},t)\Big)\,\exp(2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}).\end{aligned}$$

This holds for all $\boldsymbol{x}\in\mathbb{T}^d$, therefore by comparing each of the coefficients, we obtain

$$\mathrm{i}\,\gamma\,\widehat{u}'(\boldsymbol{h},t) = 2\pi^2\gamma^2\,\|\boldsymbol{h}\|_2^2\,\widehat{u}(\boldsymbol{h},t) + \widehat{f}(\boldsymbol{h},t), \tag{3.8}$$

for all $\boldsymbol{h}\in\mathbb{Z}^d$. □

We approximate the Fourier series (3.6) using a rank-r lattice $\Lambda(\boldsymbol{Z},\boldsymbol{n})$ and a corresponding well chosen anti-aliasing set $\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})$. Let the approximation of the solution be given by

$$u_a(\boldsymbol{x},t) := \sum_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \widehat{u}_a(\boldsymbol{h},t)\,\exp(2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}), \tag{3.9}$$

with the approximated coefficients calculated by a rank-r lattice rule

$$\widehat{u}_a(\boldsymbol{h},t) := \frac{1}{n}\sum_{\boldsymbol{p}\in\Lambda(\boldsymbol{Z},\boldsymbol{n})} u(\boldsymbol{p},t)\,\exp(-2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{p}). \tag{3.10}$$

We thus write $u_a(\boldsymbol{x},t)$ and $\widehat{u}_a(\boldsymbol{h},t)$ to denote the approximations to $u(\boldsymbol{x},t)$ and $\widehat{u}(\boldsymbol{h},t)$ respectively. For notational simplification we fix the time t and omit this argument in the remainder of this section.

We define the unitary discrete Fourier transform (DFT) to map an r-dimensional tensor $\boldsymbol{x}\in\mathbb{C}^{n_1\times\cdots\times n_r}$ to a similarly shaped tensor $\boldsymbol{X}\in\mathbb{C}^{n_1\times\cdots\times n_r}$ by the transform

$$\begin{aligned}X_{\xi_1,\ldots,\xi_r} = \frac{1}{\sqrt{n_1}}\sum_{k_1=0}^{n_1-1}\exp(-2\pi\mathrm{i}k_1\xi_1/n_1)\,\times\cdots\\ \times\,\frac{1}{\sqrt{n_r}}\sum_{k_r=0}^{n_r-1}\exp(-2\pi\mathrm{i}k_r\xi_r/n_r)\,x_{k_1,\ldots,k_r},\end{aligned} \tag{3.11}$$

for $\boldsymbol{\xi}\in\mathbb{Z}^{n_1}\oplus\cdots\oplus\mathbb{Z}^{n_r}$ and with the obvious modification for $r=1$. We define the unitary one-dimensional Fourier matrix and its inverse by

$$F_n := \frac{1}{\sqrt{n}}\Big(\exp(-2\pi\mathrm{i}\,k\xi/n\Big)_{k=0}^{n-1},\quad F_n^{-1} := \frac{1}{\sqrt{n}}\Big(\exp(2\pi\mathrm{i}\,k\xi/n\Big)_{\xi=0}^{n-1}, \tag{3.12}$$

and the r-dimensional Fourier matrix of size $n_1 \times \cdots \times n_r$ as the tensor product $F_{\boldsymbol{n}} = \otimes_{i=1}^{r} F_{n_i}$. We can then write $\boldsymbol{X} = F_{\boldsymbol{n}}\, \boldsymbol{x}$ for (3.11), and $\boldsymbol{x} = F_{\boldsymbol{n}}^{-1}\, \boldsymbol{X}$ when "vectorizing" the tensors in lexicographical ordering. The fast implementation of transforming $\boldsymbol{x}$ into $\boldsymbol{X}$, as well as its inverse, in $\mathcal{O}(n \log n)$, where $n = \prod_{i=1}^{r} n_i$, is called the fast Fourier transform (FFT) and is well known (although the direction and the normalization vary from implementation to implementation).

In the next theorem we show how to use r-dimensional FFTs to map from a rank-r lattice (in space) to a corresponding anti-aliasing set of full cardinality (in the frequency domain), and back. Note that a regular grid would be represented as a lattice with $r = d$, and in this setting the usage of the d-dimensional FFT is well known. The use of one-dimensional FFTs with a rank-1 lattice and a corresponding anti-aliasing set is also known, see, e.g., [17, 50]. We extend this for rank-r lattices by using the r-dimensional FFT. The following theorem shows three essential properties which make use of the fact that $|\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})| = n$.

Theorem 3.2. *Given a rank-r lattice point set $\Lambda(\boldsymbol{Z}, \boldsymbol{n})$ in canonical form and a corresponding anti-aliasing set $\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$ with $|\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})| = n$, the following properties hold.*

(i) (Dual character property) Define the corresponding d-dimensional Dirichlet kernel by

$$D_{\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})}(\boldsymbol{x}) := \sum_{\boldsymbol{h} \in \mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \exp(2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{x}).$$

Then for any two lattice points $\boldsymbol{p}, \boldsymbol{p}' \in \Lambda(\boldsymbol{Z}, \boldsymbol{n})$

$$\frac{1}{n} D_{\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})}(\boldsymbol{p} - \boldsymbol{p}') = \frac{1}{n} \sum_{\boldsymbol{h} \in \mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \exp(2\pi \mathrm{i} \boldsymbol{h} \cdot (\boldsymbol{p} - \boldsymbol{p}')) = \delta_{\boldsymbol{p},\boldsymbol{p}'}, \tag{3.13}$$

where $\delta_{\boldsymbol{p},\boldsymbol{p}'}$ is the Kronecker delta function that is 1 if $\boldsymbol{p} = \boldsymbol{p}'$ and 0 otherwise.

(ii) (Interpolation condition) If u_a is the approximation of a function $u \in A(\mathbb{T}^d)$ by truncating its Fourier series expansion to the anti-aliasing set $\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$ and by calculating the coefficients by the rank-r lattice rule, cfr. (3.9) *and* (3.10)*, then for any $\boldsymbol{p} \in \Lambda(\boldsymbol{Z}, \boldsymbol{n})$*

$$u_a(\boldsymbol{p}) = u(\boldsymbol{p}). \tag{3.14}$$

(iii) (Mapping through FFT) Define the r-dimensional tensors

$$\boldsymbol{u} := \big(u(\boldsymbol{p}_{(k_1,\ldots,k_r)})\big)_{k_1=0,\ldots,n_1-1,\ldots,k_r=0,\ldots,n_r-1},$$

$$\boldsymbol{u}_a := \big(u_a(\boldsymbol{p}_{(k_1,\ldots,k_r)})\big)_{k_1=0,\ldots,n_1-1,\ldots,k_r=0,\ldots,n_r-1},$$

$$\widehat{\boldsymbol{u}}_a := \big(\widehat{u}_a(\boldsymbol{h}_{(\xi_1,\ldots,\xi_r)})\big)_{\xi_1=0,\ldots,n_1-1,\ldots,\xi_r=0,\ldots,n_r-1},$$

with $\boldsymbol{p}_{\boldsymbol{k}} = (\boldsymbol{z}_1 k_1/n_1 + \cdots + \boldsymbol{z}_r k_r/n_r) \bmod 1 \in \Lambda(\boldsymbol{Z}, \boldsymbol{n})$*, and where* $\boldsymbol{h}_{\boldsymbol{\xi}} \in \mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$ *is such that* $\boldsymbol{\xi} = (\boldsymbol{h} \cdot \boldsymbol{z}_1 \bmod n_1, \ldots, \boldsymbol{h} \cdot \boldsymbol{z}_r \bmod n_r)$*. Then* $\boldsymbol{u} = \boldsymbol{u}_a$ *(by (ii)) is the collection of function values* $u(\boldsymbol{p})$ *on the lattice points* $\boldsymbol{p} \in \Lambda(\boldsymbol{Z}, \boldsymbol{n})$ *and* $\widehat{\boldsymbol{u}}_a$ *is the collection of Fourier coefficients* $\widehat{u}_a(\boldsymbol{h})$ *(by using the lattice rule, cfr.* (3.9) *and* (3.10)*) on the anti-aliasing indices* $\boldsymbol{h} \in \mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$*. The* r*-dimensional discrete Fourier transform and its inverse now map tensors* $\boldsymbol{u}_a \in \mathbb{C}^{n_1 \times \cdots \times n_r}$ *to tensors* $\widehat{\boldsymbol{u}}_a \in \mathbb{C}^{n_1 \times \cdots \times n_r}$ *and back.*

Proof.

(i) The proof is based on [11, Theorem 7.3]. Remember that $n = \prod_{i=1}^{r} n_i$. Now associate an arbitrary but fixed ordering such that we can enumerate the lattice points by $\boldsymbol{p}^{(\kappa)}$ for $\kappa = 0, \ldots, n-1$. Likewise, associate an arbitrary but fixed ordering such that we can enumerate the Fourier indices in the anti-aliasing set by $\boldsymbol{h}^{(\chi)}$ for $\chi = 0, \ldots, n-1$. Then

$$\frac{1}{n} \sum_{\kappa=0}^{n-1} \exp(2\pi \mathrm{i} \boldsymbol{h}^{(\chi)} \cdot \boldsymbol{p}^{(\kappa)}) \exp(-2\pi \mathrm{i} \boldsymbol{h}^{(\chi')} \cdot \boldsymbol{p}^{(\kappa)}) = \delta_{\chi,\chi'}, \tag{3.15}$$

for all $\chi, \chi' = 0, \ldots, n-1$, because of the character property (3.4) and since $\boldsymbol{h}^{(\chi)} - \boldsymbol{h}^{(\chi')} \notin \Lambda^{\perp}(\boldsymbol{Z}, \boldsymbol{n})$ for $\chi \neq \chi'$ because of the anti-aliasing condition. We rewrite (3.15) as

$$PMP^* = I_n,$$

where $M = \mathrm{diag}(1/n, \ldots, 1/n)$ and

$$\begin{aligned} P &= \Big(\exp(2\pi \mathrm{i} \boldsymbol{h}^{(\chi)} \cdot \boldsymbol{p}^{(\kappa)})\Big)_{\substack{\chi=0,\ldots,n-1 \\ \kappa=0,\ldots,n-1}} \\ &= \begin{pmatrix} \exp(2\pi \mathrm{i} \boldsymbol{h}^{(0)} \cdot \boldsymbol{p}^{(0)}) & \cdots & \exp(2\pi \mathrm{i} \boldsymbol{h}^{(0)} \cdot \boldsymbol{p}^{(n-1)}) \\ \vdots & \ddots & \vdots \\ \exp(2\pi \mathrm{i} \boldsymbol{h}^{(n-1)} \cdot \boldsymbol{p}^{(0)}) & \cdots & \exp(2\pi \mathrm{i} \boldsymbol{h}^{(n-1)} \cdot \boldsymbol{p}^{(n-1)}) \end{pmatrix}, \end{aligned}$$

with P^* the Hermitian conjugate of P. We note that once (3.15) holds, then the matrix P is non-singular. Therefore we obtain

$$P^* P = M^{-1},$$

which can be written as

$$\frac{1}{n} \sum_{\chi=0}^{n-1} \exp(-2\pi \mathrm{i} \boldsymbol{h}^{(\chi)} \cdot \boldsymbol{p}^{(\kappa)}) \exp(2\pi \mathrm{i} \boldsymbol{h}^{(\chi)} \cdot \boldsymbol{p}^{(\kappa')}) = \delta_{\kappa,\kappa'},$$

for all $\kappa, \kappa' = 0, \ldots, n-1$, which is equivalent to (3.13).

(*ii*) From (3.9) and (3.10) it follows

$$\begin{aligned} u_a(\boldsymbol{p}) &= \sum_{\boldsymbol{h} \in \mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \widehat{u}_a(\boldsymbol{h}) \, \exp(2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{p}) \\ &= \sum_{\boldsymbol{h} \in \mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \left(\frac{1}{n} \sum_{\boldsymbol{p}' \in \Lambda(\boldsymbol{Z},\boldsymbol{n})} u(\boldsymbol{p}') \, \exp(-2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{p}') \right) \exp(2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{p}) \\ &= \sum_{\boldsymbol{p}' \in \Lambda(\boldsymbol{Z},\boldsymbol{n})} u(\boldsymbol{p}') \, \frac{1}{n} \sum_{\boldsymbol{h} \in \mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \exp(-2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{p}') \, \exp(2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{p}) \\ &= \sum_{\boldsymbol{p}' \in \Lambda(\boldsymbol{Z},\boldsymbol{n})} u(\boldsymbol{p}') \, \delta_{\boldsymbol{p},\boldsymbol{p}'} \\ &= u(\boldsymbol{p}), \end{aligned}$$

where the dual character property (3.13) is used for the second to last equality.

(*iii*) Consider approximating the Fourier coefficient $\widehat{u}(\boldsymbol{h})$ by the rank-r lattice rule,

$$\widehat{u}_a(\boldsymbol{h}) = \frac{1}{n_1} \sum_{k_1=0}^{n_1-1} \cdots \frac{1}{n_r} \sum_{k_r=0}^{n_r-1} u(A\boldsymbol{k} \bmod 1) \, \exp(-2\pi \mathrm{i} \boldsymbol{h}^\top A \boldsymbol{k}).$$

Now define the r-dimensional function $v(k_1/n_1, \ldots, k_r/n_r) := u(A\boldsymbol{k} \bmod 1)$, then we can identify the above equation with

$$\begin{aligned} \widehat{u}_a(\boldsymbol{h}) = \frac{1}{n_1} \sum_{k_1=0}^{n_1-1} \cdots \frac{1}{n_r} \sum_{k_r=0}^{n_r-1} v(k_1/n_1, \ldots, k_r/n_r) \, \prod_{j=1}^{r} \exp(-2\pi \mathrm{i} (\boldsymbol{h} \cdot \boldsymbol{z}_j) \, k_j/n_j) \\ = \widehat{v}(\boldsymbol{h} \cdot \boldsymbol{z}_1 \bmod n_1, \ldots, \boldsymbol{h} \cdot \boldsymbol{z}_r \bmod n_r), \end{aligned}$$

where $\widehat{v}(\xi_1, \ldots, \xi_r)$ are the discrete Fourier coefficients of v. Now because of the anti-aliasing condition we can identify each $\boldsymbol{h} \in \mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$ uniquely with an index $\boldsymbol{\xi} \in \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_r}$ through $(\boldsymbol{h} \cdot \boldsymbol{z}_1 \bmod n_1, \ldots, \boldsymbol{h} \cdot \boldsymbol{z}_r \bmod n_r) = (\xi_1, \ldots, \xi_r)$. Therefore the transformation is an r-dimensional $n_1 \times \cdots \times n_r$ discrete Fourier transform. □

Finally we show the relation between the approximated coefficients $\widehat{u}_a(\boldsymbol{h})$ and the coefficients $\widehat{u}(\boldsymbol{h})$. The approximated coefficients would be exact in case the function u is solely supported on the anti-aliasing set $\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$, but in general this is not the case and we will have aliasing errors.

Lemma 3.3 (Aliasing)**.** *The approximated Fourier coefficients* (3.10) *through the lattice rule* $\Lambda(\boldsymbol{Z}, \boldsymbol{n})$ *alias the true Fourier coefficients in the following way*

$$\widehat{u}_a(\boldsymbol{h}) = \sum_{\boldsymbol{h}' \in \Lambda^\perp(\boldsymbol{Z},\boldsymbol{n})} \widehat{u}(\boldsymbol{h}+\boldsymbol{h}') = \widehat{u}(\boldsymbol{h}) + \sum_{\boldsymbol{0}\neq\boldsymbol{h}' \in \Lambda^\perp(\boldsymbol{Z},\boldsymbol{n})} \widehat{u}(\boldsymbol{h}+\boldsymbol{h}').$$

Proof. This follows from a straightforward calculation:

$$\begin{aligned}
\widehat{u}_a(\boldsymbol{h}) &= \frac{1}{n} \sum_{\boldsymbol{p}\in\Lambda(\boldsymbol{Z},\boldsymbol{n})} u(\boldsymbol{p}) \exp(-2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{p}) \\
&= \frac{1}{n} \sum_{\boldsymbol{p}\in\Lambda(\boldsymbol{Z},\boldsymbol{n})} \sum_{\boldsymbol{h}'\in\mathbb{Z}^d} \widehat{u}(\boldsymbol{h}') \exp(2\pi \mathrm{i} \boldsymbol{h}'\cdot\boldsymbol{p}) \exp(-2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{p}) \\
&= \sum_{\boldsymbol{h}'\in\mathbb{Z}^d} \widehat{u}(\boldsymbol{h}') \frac{1}{n} \sum_{\boldsymbol{p}\in\Lambda(\boldsymbol{Z},\boldsymbol{n})} \exp(2\pi \mathrm{i} (\boldsymbol{h}'-\boldsymbol{h})\cdot\boldsymbol{p}) \\
&= \sum_{\boldsymbol{h}'\in\Lambda^\perp(\boldsymbol{Z},\boldsymbol{n})} \widehat{u}(\boldsymbol{h}+\boldsymbol{h}'),
\end{aligned}$$

where the character property (3.4) is used in the last equality. □

This last lemma shows that for the approximation u_a to be meaningful the Fourier coefficients necessarily have to decay at a certain rate such that the error in the approximation $\widehat{u}_a(\boldsymbol{h})$ can be bounded. This decay is not enforced by asking $u \in A(\mathbb{T}^d)$, but it is enforced by asking u to be in the Korobov space $E_\alpha(\mathbb{T}^d)$.

3.2.3 Strang splitting

We will use Strang splitting to do time stepping on our discretized function. The idea of Strang splitting [83] is to break up the solution operator for an ODE which consists of a sum of two differential operators into applying them each separately in a way to be specified below and thereby achieving second order convergence with respect to the time step. Strang splitting can be applied to initial value problems of the form

$$y'(t) = (A+B)\, y(t), \qquad y(0) = y_0,$$

where A and B are differential operators.

We first explain a splitting method which attains first order convergence in the time step and then show the Strang operator splitting which gives second order convergence in the time step. If A and B are constant coefficient matrices, as is the case in our application, then the solution is given by

$$y(t) = \mathrm{e}^{(A+B)\,t}\, y_0,$$

where $\mathrm{e}^{(A+B)}$ is the matrix exponential. If A and B commute, i.e., $AB = BA$, then $y(t) = \mathrm{e}^{A\,t}\,\mathrm{e}^{B\,t}\, y_0$. This follows from the Baker–Campbell–Hausdorff formula from Lie group analysis

$$\log(\mathrm{e}^{A\,t}\,\mathrm{e}^{B\,t}) = (A+B)\,t + [A,B]\,\frac{t^2}{2} + ([A,[A,B]] + [B,[B,A]])\,\frac{t^3}{12} + \cdots,$$

where the commutator of two operators A and B is defined by $[A,B] := AB - BA$, which reduces to $\mathrm{e}^{A\,t}\,\mathrm{e}^{B\,t} = \mathrm{e}^{(A+B)\,t}$ if $[A,B] = 0$, where 0 should be interpreted as the zero matrix. If $[A,B]$ is nonzero then, given an initial solution $y(t)$, we can write

$$\begin{aligned} y(t+\Delta t) &= \mathrm{e}^{(A+B)\,\Delta t}\, y(t) \\ &= \exp\left(\log(\mathrm{e}^{A\,\Delta t}\,\mathrm{e}^{B\,\Delta t}) - [A,B]\,\frac{(\Delta t)^2}{2} - \mathcal{O}((\Delta t)^3)\right) y(t), \end{aligned}$$

which for a discrete time stepping scheme $y_t \approx y(t)$ can be used to show a global error of first order in Δt for bounded (fixed) commutator (see e.g., [33]).

For Strang operator splitting we first write $A + B = \frac{1}{2}B + A + \frac{1}{2}B$. We now want to approximate $\mathrm{e}^{(A+B)\,t}$ by $\mathrm{e}^{\frac{1}{2}B\,t}\mathrm{e}^{A\,t}\mathrm{e}^{\frac{1}{2}B\,t}$, in effect taking twice half a time step for B and sandwiching a full time step for A in the middle. The Strang splitting method for a time discretization Δt then operates as follows

$$y_{k+1} = \mathrm{e}^{\frac{1}{2}B\,\Delta t}\,\mathrm{e}^{A\,\Delta t}\,\mathrm{e}^{\frac{1}{2}B\,\Delta t}\, y_k, \tag{3.16}$$

where $y_k \approx y(k\Delta t)$ and $y_0 = y(0)$ is the initial value. Note that if A and B would commute, the scheme is exact. We have the following local error bound (per time step) from [33, Theorem 2.1].

Theorem 3.4 (Strang splitting local error bounds)**.** *Let X be a Banach space equipped with the norm $\|\cdot\|$, A the generator of the strongly continuous semigroup $\mathrm{e}^{A\,t}$ on the Banach space X, and B be a bounded linear operator on X with induced operator norm $\|B\| := \sup_{0\neq y\in X} \|B\,y\|/\|y\| < \infty$. Let ω be an arbitrary constant. Then the following hold:*

(i) If there exist constants $\alpha \geq 0$ and c_1 satisfying

$$\|[A,B]\,y\| \leq c_1\,\|(A+\omega\,I)^{\alpha}\,y\| \qquad \text{for all } y \in X,$$

then the local error of the Strang splitting method is bounded as follows

$$\left\| \mathrm{e}^{\frac{1}{2}B\,\tau}\,\mathrm{e}^{A\,\tau}\,\mathrm{e}^{\frac{1}{2}B\,\tau} y - \mathrm{e}^{(A+B)\,\tau}\,y \right\| \le C_1\,\tau^2\,\|(A+\omega\,I)^{\alpha}\,y\| \qquad \textit{for all } y \in X,$$

where C_1 depends only on c_1 and $\|B\|$.

(ii) Under the condition above and additionally if there exist constants $\beta \ge 1 \ge \alpha$, and c_2 satisfying

$$\|[A,[A,B]]\,y\| \le c_2\,\|(A+\omega\,I)^{\beta} y\| \qquad \textit{for all } y \in X,$$

then the local error of the Strang splitting method is bounded as follows

$$\left\| \mathrm{e}^{\frac{1}{2}B\,\tau}\,\mathrm{e}^{A\,\tau}\,\mathrm{e}^{\frac{1}{2}B\,\tau} y - \mathrm{e}^{(A+B)\,\tau}\,y \right\| \le C_2\,\tau^3\,\|(A+\omega\,I)^{\beta}\,y\| \qquad \textit{for all } y \in X$$

where C_2 depends only on c_1, c_2 and $\|B\|$.

Proof. See [33, Theorem 2.1]. □

Now assume $\widehat{\boldsymbol{u}}_t$ are the approximate Fourier coefficients of $u(\boldsymbol{x},t)$ at time t. The previous theorem shows that we need a bound on $\|(A+\omega\,I)^{\alpha}\,\widehat{\boldsymbol{u}}_t\|_2$ and that $\|B\|_2$ should be bounded to get first order convergence for the global error of the time stepping scheme using the Strang splitting method. It also shows that if we have a bound on $\|(A+\omega I)^{\beta}\,\widehat{\boldsymbol{u}}_t\|_2$ we obtain second order convergence for the global error of the time stepping scheme.

In Lemma 3.6 we will first derive the key ingredient for our main result when the discretization in space is done by a rank-1 lattice $\Lambda(\boldsymbol{z},n)$ with corresponding finite Fourier series on an associated anti-aliasing set $\mathcal{A}(\boldsymbol{z},n)$. In Lemma 3.7 we will extend the result to general rank-r lattices $\Lambda(\boldsymbol{Z},\boldsymbol{n})$ which include any regular (possibly anisotropic) grid.

3.2.4 Strang splitting and rank-1 lattices

Denote by $\widehat{\boldsymbol{u}}_t := \big(\widehat{u}_a(\boldsymbol{h}^{(0)},t),\,\ldots,\widehat{u}_a(\boldsymbol{h}^{(n-1)},t)\big)$ the approximated solution at time t using a fixed anti-aliasing set $\mathcal{A}(\boldsymbol{z},n) = \{\boldsymbol{h}_\xi : \xi = 0,\ldots,n-1\}$ of full size n, where $\boldsymbol{h}_\xi \in \mathcal{A}(\boldsymbol{z},n)$ is such that $\boldsymbol{h}_\xi \cdot \boldsymbol{z} \equiv \xi \pmod n$. Demanding that (3.7) holds for all $\boldsymbol{h} \in \mathcal{A}(\boldsymbol{z},n)$, we have the following relation

$$\mathrm{i}\,\gamma\,\widehat{\boldsymbol{u}}_t' = \frac{1}{2}\gamma^2 D_n \widehat{\boldsymbol{u}}_t + W_n \widehat{\boldsymbol{u}}_t, \tag{3.17}$$

with the initial condition $\widehat{\boldsymbol{u}}_0 = \widehat{\boldsymbol{g}}_a := (\widehat{g}_a(\boldsymbol{h}^{(0)}), \ldots, \widehat{g}_a(\boldsymbol{h}^{(n-1)}))$,

$$D_n := \operatorname{diag}\left((4\pi^2 \|\boldsymbol{h}_\xi\|_2^2)_{\xi=0,\ldots,n-1}\right), \tag{3.18}$$

and the linear operator $W_n := F_n V_n F_n^{-1}$ with

$$V_n := \operatorname{diag}\left((v(\boldsymbol{p}_k))_{k=0,\ldots,n-1}\right), \tag{3.19}$$

where F_n is the unitary Fourier matrix. We mention that it is cheap to compute matrix exponentials of D_n and W_n exactly. For the derivation of W_n, we have the following Lemma.

Lemma 3.5 (Multiplication operator on rank-1 lattices)**.** *Given a rank-1 lattice point set $\Lambda(\boldsymbol{z}, n)$ and corresponding anti-aliasing set $\mathcal{A}(\boldsymbol{z}, n)$ of full size, a potential function $v \in E_\alpha(\mathbb{T}^d)$ with $\alpha \geq 2$ and a function $u_a \in E_\beta(\mathbb{T}^d)$ with $\beta \geq 2$ with Fourier coefficients only supported on $\mathcal{A}(\boldsymbol{z}, n)$. Then the action in the Fourier domain restricted to $\mathcal{A}(\boldsymbol{z}, n)$ of multiplying with v, that is $f_a(\boldsymbol{x}) = v(\boldsymbol{x})\, u_a(\boldsymbol{x})$, on the nodes of the rank-1 lattice, and with f_a having Fourier coefficients restricted to the set $\mathcal{A}(\boldsymbol{z}, n)$, can be described by a circulant matrix $W_n \in \mathbb{C}^{n\times n}$ with $W_n = F_n\, V_n\, F_n^{-1}$, with V_n given by* (3.19) *and F_n the unitary Fourier matrix* (3.12)*, where the element at position (ξ, ξ') of W_n is given by*

$$w_{\xi,\xi'} = w_{(\xi-\xi') \bmod n} = \sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^d \\ \boldsymbol{h}\cdot\boldsymbol{z}\equiv\xi-\xi' \pmod n}} \widehat{v}(\boldsymbol{h}). \tag{3.20}$$

Proof. In the following, F_n is the unitary discrete Fourier transformation matrix. We denote the coefficients of the product $v(\boldsymbol{x})u_a(\boldsymbol{x})$ by $\widehat{f}(\boldsymbol{h})$. For each $\boldsymbol{h} \in \mathcal{A}(\boldsymbol{z}, n)$ we have

$$\begin{aligned}
\widehat{f}(\boldsymbol{h}) &= \int_{[0,1]^d} u_a(\boldsymbol{x}) v(\boldsymbol{x}) \exp(-2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \\
&= \int_{[0,1]^d} \left(\sum_{\boldsymbol{h}'\in\mathcal{A}(\boldsymbol{z},n)} \widehat{u}_a(\boldsymbol{h}') \exp(2\pi \mathrm{i} \boldsymbol{h}'\cdot\boldsymbol{x}) \right) \\
&\qquad\qquad \times \left(\sum_{\boldsymbol{h}''\in\mathbb{Z}^d} \widehat{v}(\boldsymbol{h}'') \exp(2\pi \mathrm{i} \boldsymbol{h}''\cdot\boldsymbol{x}) \right) \exp(-2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \\
&= \sum_{\boldsymbol{h}'\in\mathcal{A}(\boldsymbol{z},n)} \widehat{v}(\boldsymbol{h}-\boldsymbol{h}')\, \widehat{u}_a(\boldsymbol{h}').
\end{aligned}$$

By Lemma 3.3 The coefficients calculated on the rank-1 lattice points have the form

$$\widehat{v}_a(\boldsymbol{h}-\boldsymbol{h}') = \sum_{\boldsymbol{\ell}\in\Lambda^\perp(\boldsymbol{z},n)} \widehat{v}(\boldsymbol{h}-\boldsymbol{h}'+\boldsymbol{\ell}).$$

Therefore we have the following approximation for $\widehat{f}(\boldsymbol{h})$:

$$\widehat{f}_a(\boldsymbol{h}) = \sum_{\boldsymbol{h}'\in\mathcal{A}(\boldsymbol{z},n)} \sum_{\boldsymbol{\ell}\in\Lambda^\perp(\boldsymbol{z},n)} \widehat{v}(\boldsymbol{h}-\boldsymbol{h}'+\boldsymbol{\ell})\widehat{u}_a(\boldsymbol{h}',t).$$

We have hence proved the claims of the lemma. □

The exact solution of the ordinary differential equation (3.17) is

$$\widehat{\boldsymbol{u}}_t = \mathrm{e}^{-\frac{\mathrm{i}}{\gamma}W_n\, t-\frac{\mathrm{i}\gamma}{2}D_n\, t}\,\widehat{\boldsymbol{u}}_0.$$

Applying the Strange splitting method (3.16) then gives us

$$\widehat{\boldsymbol{u}}_a^{k+1} = \mathrm{e}^{-\frac{\mathrm{i}}{2\gamma}W_n\Delta t}\,\mathrm{e}^{-\frac{\mathrm{i}\gamma}{2}D_n\Delta t}\,\mathrm{e}^{-\frac{\mathrm{i}}{2\gamma}W_n\Delta t}\,\widehat{\boldsymbol{u}}_a^k \qquad \text{for } k=0,1,\ldots,m-1, \quad (3.21)$$

where

$$\mathrm{e}^{-\frac{\mathrm{i}}{2}W_n\Delta t} = F_n\,\mathrm{diag}\left((\mathrm{e}^{-\frac{\mathrm{i}}{2}v(\boldsymbol{p}_k)\Delta t})_{k=0,\ldots,n-1}\right)F_n^{-1}.$$

We then approximate the solution of the differential system at time $t=k\Delta t$ by stepping with a time step of Δt iteratively.

To bound the error of the Strang splitting we need to bound the effect of the commutators as specified in Theorem 3.4, for this we will make use of the Korobov space. Now we are ready to state our key theoretical result, namely that the Strang splitting has bounded error of the time evolution when the discretization in space is done by a rank-1 lattice rule and the truncation of the Fourier series is done on an associated anti-aliasing set. First we show that the commutators of the operators $D=\frac{\gamma}{2}D_n$ and $W=\frac{1}{\gamma}W_n$ are bounded in the sense of Theorem 3.4 with $A=D$, $B=W$ and $\omega=1$.

Lemma 3.6 (Rank-1 lattice commutator bounds)**.** *Given a rank-1 lattice with generating vector $\boldsymbol{z}\in\mathbb{Z}^d$ and modulus n and a TDSE with a potential function $v\in E_\alpha(\mathbb{T}^d)$ with $\alpha\geq 2$ and an initial condition $g\in E_\beta(\mathbb{T}^d)$ with $\beta\geq 2$. Let $D=\frac{\gamma}{2}D_n$ and $W=\frac{1}{\gamma}W_n$ with D_n and $W_n=F_nV_nF_n^{-1}$ as defined in* (3.18) *and* (3.20)*, and with V_n as defined in* (3.19) *using the potential function v.*

If the anti-aliasing set $\mathcal{A}(\boldsymbol{z},n)=\{\boldsymbol{h}_\xi\in\mathbb{Z}^d : \boldsymbol{h}_\xi\cdot\boldsymbol{z}\equiv\xi \pmod n$ for $\xi=0,\ldots,n-1\}$, with full cardinality, is chosen such that it has minimal ℓ_2 norm, i.e.,

$$\|\boldsymbol{h}_\xi\|_2 = \min_{\boldsymbol{h}'\in A(\boldsymbol{z},n,\xi)}\|\boldsymbol{h}'\|_2, \quad (3.22)$$

with

$$A(\boldsymbol{z}, n, \xi) := \big\{\boldsymbol{h} \in \mathbb{Z}^d : \boldsymbol{h} \cdot \boldsymbol{z} \equiv \xi \pmod{n}\big\},$$

then the following hold:

(i) If $v \in E_\alpha(\mathbb{T}^d)$ with parameter $\alpha > 5/2$, then for all $\boldsymbol{y} \in \mathbb{R}^n$ we have

$$\|[D, W]\,\boldsymbol{y}\|_2 \le c_1\,\|(D + I)\,\boldsymbol{y}\|_2,$$

where c_1 is a constant independent of n and $\boldsymbol{y}$.

(ii) If $v \in E_\alpha(\mathbb{T}^d)$ with $\alpha > 9/2$, then for all $\boldsymbol{y} \in \mathbb{R}^n$ we have

$$\|[D, [D, W]]\,\boldsymbol{y}\|_2 \le c_2\,\|(D + I)^2\,\boldsymbol{y}\|_2,$$

where c_2 is a constant, independent of n and $\boldsymbol{y}$.

Proof. We first prove the first order result (i) and then prove the second order result (ii). We note that the regularity of the solution $u \in E_2(\mathbb{T}^d)$ is guaranteed by Lemma 3.1 due to the regularities of the potential v and initial condition g.

(i) Since $(D + I)$ is non-singular, we show

$$\|[D, W]\,(D + I)^{-1}\,\boldsymbol{y}\|_2 \le c_1 \|\boldsymbol{y}\|_2 \qquad \text{for all } \boldsymbol{y} \in \mathbb{R}^n.$$

Hence we need to bound the induced matrix p-norm $\|A\|_p := \sup_{\boldsymbol{0} \neq \boldsymbol{y} \in \mathbb{R}^n} \|A\,\boldsymbol{y}\|_p / \|\boldsymbol{y}\|_p$ for $p = 2$ for the matrix $A = [D, W]\,(D + I)^{-1} \in \mathbb{R}^{n \times n}$ by an absolute constant. We have

$$[D, W] = D\,W - W\,D = \Big(2\pi^2 \left(\|\boldsymbol{h}_\xi\|_2^2 - \|\boldsymbol{h}_{\xi'}\|_2^2\right)\, w_{\xi - \xi'}\Big)_{\xi, \xi' = 0, \ldots, n-1},$$

where the subscript of w should be interpreted modulo n, see (3.20). For ease of notation we now multiply by γ and consider the matrix M defined by

$$M := \gamma\,(D\,W - W\,D)\,(D + I)^{-1} = \left(\frac{\|\boldsymbol{h}_\xi\|_2^2 - \|\boldsymbol{h}_{\xi'}\|_2^2}{\|\boldsymbol{h}_{\xi'}\|_2^2 + c}\, w_{\xi - \xi'}\right)_{\xi, \xi' = 0, \ldots, n-1},$$

where $c = 1/(2\pi^2\gamma) > 0$. Note that $\|[D, W]\,(D+I)^{-1}\|_2 = \frac{1}{\gamma}\|M\|_2$. By Hölder's inequality, we have $\|M\|_2 \le \sqrt{\|M\|_1 \|M\|_\infty}$, therefore we will bound $\|M\|_1 = \max_{\xi' = 0, \ldots, n-1} \sum_{\xi=0}^{n-1} |M_{\xi, \xi'}|$ and $\|M\|_\infty = \max_{\xi = 0, \ldots, n-1} \sum_{\xi'=0}^{n-1} |M_{\xi, \xi'}|$. Clearly the diagonal elements of M are zero and we can exclude those cases in the

following. For $\|M\|_1$ we obtain

$$\|M\|_1 = \max_{\xi' \in \mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi \neq \xi'}}^{n-1} \left| \frac{\|\boldsymbol{h}_\xi\|_2^2 - \|\boldsymbol{h}_{\xi'}\|_2^2}{\|\boldsymbol{h}_{\xi'}\|_2^2 + c} \, w_{\xi-\xi'} \right|$$

$$\leq \max_{\xi' \in \mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi \neq \xi'}}^{n-1} \left[(1 + \frac{\|\boldsymbol{h}_\xi\|_2^2}{\max(\|\boldsymbol{h}_{\xi'}\|_2^2, c)}) \, |w_{\xi-\xi'}| \right].$$

Note that

$$\max_{\xi' \in \mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi \neq \xi'}}^{n-1} |w_{\xi-\xi'}| = \sum_{i=1}^{n-1} |w_i| \leq \sum_{\boldsymbol{h} \in \mathbb{Z}^d} |\widehat{v}(\boldsymbol{h})| = \|v\|_{A(\mathbb{T}^d)} < \infty,$$

since we assume $v \in A(\mathbb{T}^d)$, see Lemma 3.1. We still need to bound

$$\max_{\xi' \in \mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi \neq \xi'}}^{n-1} \frac{\|\boldsymbol{h}_\xi\|_2^2}{\max(\|\boldsymbol{h}_{\xi'}\|_2^2, \, c)} \, |w_{\xi-\xi'}| = \max_{\xi' \in \mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi \neq \xi'}}^{n-1} \frac{\|\boldsymbol{h}_\xi\|_2^2}{\max(\|\boldsymbol{h}_{\xi'}\|_2^2, \, c)} \frac{\|\boldsymbol{h}_{\xi-\xi'}\|_2^2}{\|\boldsymbol{h}_{\xi-\xi'}\|_2^2} \, |w_{\xi-\xi'}|, \tag{3.23}$$

where also $\boldsymbol{h}_{\xi-\xi'}$ has to be read as $\boldsymbol{h}_{(\xi-\xi') \bmod n}$. Note that $\boldsymbol{h}_{\xi-\xi} = \boldsymbol{h}_0 = \boldsymbol{0}$ is excluded from the sum. Since the anti-aliasing set is such that $\boldsymbol{h}_\xi$ has minimal ℓ_2 norm by (4.5) we can bound $\|\boldsymbol{h}_\xi\|_2 \leq \|\boldsymbol{h}'_\xi\|_2$ for any $\boldsymbol{h}'_\xi \in A(\boldsymbol{z}, n, \xi)$ with the property $\boldsymbol{h}'_\xi \cdot \boldsymbol{z} \equiv \xi \pmod n$. In particular for $\boldsymbol{h}'_\xi = \boldsymbol{h}_{\xi-\xi'} + \boldsymbol{h}_{\xi'}$ since $(\boldsymbol{h}_{\xi-\xi'} + \boldsymbol{h}_{\xi'}) \cdot \boldsymbol{z} \equiv \xi \pmod n$ for any choice of $\xi' = 0, \ldots, n-1$. Therefore

$$\|\boldsymbol{h}_\xi\|_2^2 \leq \|\boldsymbol{h}'_\xi\|_2^2 = \|\boldsymbol{h}_{\xi-\xi'} + \boldsymbol{h}_{\xi'}\|_2^2 \leq \|\boldsymbol{h}_{\xi-\xi'}\|_2^2 + 2\,\|\boldsymbol{h}_{\xi-\xi'}\|_2\,\|\boldsymbol{h}_{\xi'}\|_2 + \|\boldsymbol{h}_{\xi'}\|_2^2,$$

and thus (remembering we have $\boldsymbol{h}_{\xi-\xi'} \neq \boldsymbol{0}$)

$$\frac{\|\boldsymbol{h}_\xi\|_2^2}{\max(\|\boldsymbol{h}_{\xi'}\|_2^2, \, c)\,\|\boldsymbol{h}_{\xi-\xi'}\|_2^2} \leq \frac{\|\boldsymbol{h}_{\xi-\xi'}\|_2^2 + 2\,\|\boldsymbol{h}_{\xi-\xi'}\|_2\,\|\boldsymbol{h}_{\xi'}\|_2 + \|\boldsymbol{h}_{\xi'}\|_2^2}{\max(\|\boldsymbol{h}_{\xi'}\|_2^2, \, c)\,\|\boldsymbol{h}_{\xi-\xi'}\|_2^2}$$

$$\leq \max(1/c, \, 4).$$

Let $c' := \max(1/c, 4)$. We continue from (3.23) to obtain

$$\begin{aligned}
&\max_{\xi' \in \mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi \neq \xi'}}^{n-1} \frac{\|\boldsymbol{h}_\xi\|_2^2}{\max(\|\boldsymbol{h}_{\xi'}\|_2^2,\, c)} \frac{\|\boldsymbol{h}_{\xi-\xi'}\|_2^2}{\|\boldsymbol{h}_{\xi-\xi'}\|_2^2} \, |w_{\xi-\xi'}| \\
&\leq c' \max_{\xi' \in \mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi \neq \xi'}}^{n-1} \|\boldsymbol{h}_{\xi-\xi'}\|_2^2 \left| \sum_{\boldsymbol{h} \in A(\boldsymbol{z}, n, \xi-\xi')} \widehat{v}(\boldsymbol{h}) \right| \\
&= c' \max_{\xi' \in \mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi \neq \xi'}}^{n-1} \|\boldsymbol{h}_{\xi-\xi'}\|_2^2 \left| \sum_{\boldsymbol{h} \in A(\boldsymbol{z}, n, \xi-\xi')} \widehat{v}(\boldsymbol{h}) \right| \\
&\leq c' \max_{\xi' \in \mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi \neq \xi'}}^{n-1} \sum_{\boldsymbol{h} \in A(\boldsymbol{z}, n, \xi-\xi')} \|\boldsymbol{h}\|_2^2 \, |\widehat{v}(\boldsymbol{h})| \\
&\leq c' \max_{\xi' \in \mathbb{Z}_n} \sum_{\boldsymbol{h} \in \mathbb{Z}^d} \|\boldsymbol{h}\|_2^2 \, |\widehat{v}(\boldsymbol{h})|.
\end{aligned}$$

The last inequality follows from (3.2) and is independent from ξ' such that we can drop the maximum. For the function $v \in E_\alpha(\mathbb{T}^d)$ with $\alpha > 5/2$ the following holds by applying the Cauchy–Schwarz inequality and multiplying and dividing by r_α, as defined in (3.5),

$$\begin{aligned}
\sum_{\boldsymbol{h} \in \mathbb{Z}^d} \|\boldsymbol{h}\|_2^2 \, |\widehat{v}(\boldsymbol{h})| &\leq \left(\sum_{\boldsymbol{h} \in \mathbb{Z}^d} r_\alpha^2(\boldsymbol{h}) \, |\widehat{v}(\boldsymbol{h})|^2 \right)^{1/2} \left(\sum_{\boldsymbol{h} \in \mathbb{Z}^d} \frac{\|\boldsymbol{h}\|_2^4}{r_\alpha^2(\boldsymbol{h})} \right)^{1/2} \\
&\leq \|v\|_{E_\alpha(\mathbb{T}^d)} \left(\sum_{\boldsymbol{h} \in \mathbb{Z}^d} \frac{(\sqrt{d}\, \|\boldsymbol{h}\|_\infty)^4}{r_\alpha^2(\boldsymbol{h})} \right)^{1/2} \\
&\leq \|v\|_{E_\alpha(\mathbb{T}^d)} \left(\sum_{\boldsymbol{h} \in \mathbb{Z}^d} \frac{d^2}{r_{\alpha-2}^2(\boldsymbol{h})} \right)^{1/2} \\
&\leq \|v\|_{E_\alpha(\mathbb{T}^d)} \left(d^2 \left(1 + 2\, \zeta(2\alpha - 4)\right)^d \right)^{1/2} < \infty.
\end{aligned}$$

Therefore we have bounded $\|M\|_1$ independent of n. For $\|M\|_\infty$ we can proceed in a similar way to obtain

$$\|M\|_\infty = \max_{\xi\in\mathbb{Z}_n} \sum_{\substack{\xi'=0\\ \xi'\neq\xi}}^{n-1} \left| \frac{\|\boldsymbol{h}_\xi\|_2^2 - \|\boldsymbol{h}_{\xi'}\|_2^2}{\|\boldsymbol{h}_{\xi'}\|_2^2 + c}\, w_{\xi-\xi'} \right|$$

$$\leq \|v\|_{A(\mathbb{T}^d)} + c' \|v\|_{E_\alpha(\mathbb{T}^d)} \left(d^2\,(1+2\,\zeta(2\alpha-4))^d\right)^{1/2}.$$

Therefore, for any $\boldsymbol{y}\in\mathbb{R}^n$ it holds that

$$\|(D\,W - W\,D)\,\boldsymbol{y}\|_2 \leq c_1\,\|(D+I)\,\boldsymbol{y}\|_2,$$

where c_1 is a constant independent of $\boldsymbol{y}$ and n.

(*ii*) Similar argument holds for second order convergence. Then

$$[D,\,[D,W]]\,(D+I)^{-2} = \left(\frac{(\|\boldsymbol{h}_\xi\|_2^2 - \|\boldsymbol{h}_{\xi'}\|_2^2)^2}{2\pi^2\gamma\,(\|\boldsymbol{h}_{\xi'}\|_2^2 + c)^2}\, w_{\xi-\xi'} \right)_{\xi,\xi'=0,\ldots,n-1},$$

with the same constant $c = 1/(2\pi^2\gamma)$. For $\xi\neq\xi'$ we can multiply and divide by $\|\boldsymbol{h}_{\xi-\xi'}\|_2^4$ and then

$$\frac{(\|\boldsymbol{h}_\xi\|_2^2 - \|\boldsymbol{h}_{\xi'}\|_2^2)^2}{(\|\boldsymbol{h}_{\xi'}\|_2^2 + c)^2 \|\boldsymbol{h}_{\xi-\xi'}\|_2^4} \leq \frac{(\|\boldsymbol{h}_{\xi-\xi'} + \boldsymbol{h}_{\xi'}\|_2^2 + \|\boldsymbol{h}_{\xi'}\|_2^2)^2}{(\|\boldsymbol{h}_{\xi'}\|_2^2 + c)^2 \|\boldsymbol{h}_{\xi-\xi'}\|_2^4}$$

has an upper bound of $\max(25, 1/c^2)$. Therefore, the ℓ_1 and ℓ_∞ induced norms of this matrix can be bounded if the potential function $v(\boldsymbol{x})$ is in Korobov space $E_\alpha(\mathbb{T}^d)$ with the smoothness parameter $\alpha > 9/2$:

$$\max_{\xi'\in\mathbb{Z}_n} \sum_{\substack{\xi=0\\ \xi\neq\xi'}}^{n-1} \|\boldsymbol{h}_{\xi-\xi'}\|_2^4\,|w_{\xi-\xi'}| \leq \sum_{\boldsymbol{h}\in\mathbb{Z}^d} \|\boldsymbol{h}\|_2^4\,|\widehat{v}(\boldsymbol{h})|$$

$$\leq \|v\|_{E_\alpha(\mathbb{T}^d)} \left(d^4\,(1+2\zeta(2\alpha-8))^d\right)^{1/2}.$$

We have hence proved the claims of the lemma. □

In Algorithm 1, our procedure of the time-stepping on rank-1 lattice points is shown. Each time-step is done with complexity $\mathcal{O}(n\log n)$. Matrices $\mathrm{e}^{-\frac{\mathrm{i}}{2\gamma}V_n\Delta t}$ and $\mathrm{e}^{-\frac{\mathrm{i}\gamma}{2}D_n\Delta t}$ are diagonal, hence there is no need to store n-by-n matrices.

Algorithm 1 Strang splitting on rank-1 lattice points

Input:
$\Delta t, m, g$ ▷ $m\Delta t = T$ is the final time, g is the initial condition
$\mathcal{A}(\boldsymbol{z}, n) = \{\boldsymbol{h}_0, ..., \boldsymbol{h}_{n-1}\} \subset \mathbb{Z}^{d\times n}$ ▷ Anti-aliasing set of full cardinality
$\Lambda(\boldsymbol{z}, n) = \{\boldsymbol{p}_0, ..., \boldsymbol{p}_{n-1}\} \subset \mathbb{T}^{d\times n}$ ▷ Rank-1 lattice points

$V_n = \operatorname{diag}\left((v(\boldsymbol{p}_k))_{k=0,\dots,n-1}\right)$ ▷ The potential matrix on lattice points
$D_n = \operatorname{diag}\left((4\pi^2\|\boldsymbol{h}_\xi\|_2^2)_{\xi=0,\dots,n-1}\right)$ ▷ The Laplacian matrix
$\widehat{\boldsymbol{u}}_a^0 = \widehat{\boldsymbol{g}}_a = (\widehat{g}_a(\boldsymbol{h}_0), \dots, \widehat{g}_a(\boldsymbol{h}_{n-1})) = F_n(g(\boldsymbol{p}_0), \dots, g(\boldsymbol{p}_{n-1}))/\sqrt{n}$

for $k = 1, 2, ..., m$ **do**
$\quad \widehat{\boldsymbol{u}}_a^k = F_n \mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} V_n \Delta t} F_n^{-1} \mathrm{e}^{-\frac{\mathrm{i}\gamma}{2} D_n \Delta t} F_n \mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} V_n \Delta t} F_n^{-1} \widehat{\boldsymbol{u}}_a^{k-1}$
end for

Output: $\widehat{\boldsymbol{u}}_a^m$

3.2.5 Strang splitting and rank-r lattices

In this section, we generalize the results of the previous section for rank-r lattices. Consider a rank-r lattice $\Lambda(\boldsymbol{Z}, \boldsymbol{n})$ in canonical form, and the corresponding anti-aliasing set $\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$ with full cardinality $n = \prod_{i=1}^r n_i$. We enumerate the anti-aliasing set in "lexicographical ordering" by identifying $\boldsymbol{h}^{(\chi)} = \boldsymbol{h}_{\boldsymbol{\xi}}$ for $\chi = 0, \dots, n-1$ and $\boldsymbol{\xi} = \boldsymbol{h}_{\boldsymbol{\xi}} \cdot \boldsymbol{z} \bmod \boldsymbol{n}$ for all $\boldsymbol{h}_{\boldsymbol{\xi}} \in \mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})$ such that

$$\chi = \xi_1\, n_2 \cdots n_r + \xi_2\, n_3 \cdots n_r + \cdots + \xi_r = \sum_{i=1}^{r} \left(\xi_i \prod_{j=i+1}^{r} n_j \right),$$

for all $\boldsymbol{\xi} \in \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_r}$. Likewise, we enumerate the lattice points by identifying $\boldsymbol{p}^{(\kappa)} = \boldsymbol{p}_{\boldsymbol{k}}$ for $\kappa = 0, \dots, n-1$ such that

$$\kappa = \kappa_1\, n_2 \cdots n_r + \kappa_2\, n_3 \cdots n_r + \cdots + \kappa_r = \sum_{i=1}^{r} \left(\kappa_i \prod_{j=i+1}^{r} n_j \right),$$

for all $\boldsymbol{k} \in \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_r}$. Then the ordinary differential equation (3.17) holds with, $\widehat{\boldsymbol{u}}_0 = \widehat{\boldsymbol{g}}_a := (\widehat{g}_a(\boldsymbol{h}^{(0)}), \dots, \widehat{g}_a(\boldsymbol{h}^{(n-1)}))$,

$$D_{\boldsymbol{n}} := \operatorname{diag}\left((4\pi^2 \|\boldsymbol{h}^{(\chi)}\|_2^2)_{\chi=0,\dots,n-1} \right), \tag{3.24}$$

and

$$W_{\boldsymbol{n}} := F_{\boldsymbol{n}} V_{\boldsymbol{n}} F_{\boldsymbol{n}}^{-1}, \tag{3.25}$$

with

$$V_{\boldsymbol{n}} := \operatorname{diag}\left(\left(v(\boldsymbol{p}^{(\kappa)})\right)_{\kappa=0,\ldots,n-1}\right), \tag{3.26}$$

where $F_{\boldsymbol{n}}$ is the r-dimensional discrete Fourier transform. With these notations we have the following generalization of Lemma 3.6.

Lemma 3.7 (Rank-r commutator bounds)**.** *Given a rank-r lattice $\Lambda(\boldsymbol{Z}, \boldsymbol{n})$ in canonical form with the number of points $n = \prod_{i=1}^{r} n_i$, and a TDSE with a potential function $v \in E_\alpha(\mathbb{T}^d)$ with $\alpha \geq 2$ and an initial condition $g \in E_\beta(\mathbb{T}^d)$ with $\beta \geq 2$. Let $D = \frac{\gamma}{2} D_{\boldsymbol{n}}$ and $W = \frac{1}{\gamma} W_{\boldsymbol{n}}$ with $D_{\boldsymbol{n}}$ and $W_{\boldsymbol{n}} = F_{\boldsymbol{n}} V_{\boldsymbol{n}} F_{\boldsymbol{n}}^{-1}$ as defined in* (3.24) *and* (3.25)*, and with $V_{\boldsymbol{n}}$ as defined in* (3.26) *using the potential function v.*

If the anti-aliasing set $\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n}) = \{\boldsymbol{h}_{\boldsymbol{\xi}} \in \mathbb{Z}^d : \boldsymbol{Z}^\top \boldsymbol{h}_{\boldsymbol{\xi}} \equiv \boldsymbol{\xi} \pmod{\boldsymbol{n}}$ for $\boldsymbol{\xi} \in \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_r}\}$, with full cardinality, is chosen such that each $\boldsymbol{h}_{\boldsymbol{\xi}}$ with $\boldsymbol{\xi} \in \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_r}$ has minimal ℓ_2 norm, i.e.,

$$\|\boldsymbol{h}_{\boldsymbol{\xi}}\|_2 = \min_{\boldsymbol{h}' \in A(\boldsymbol{Z}, \boldsymbol{n}, \boldsymbol{\xi})} \|\boldsymbol{h}'\|_2,$$

with

$$A(\boldsymbol{Z}, \boldsymbol{n}, \boldsymbol{\xi}) := \left\{\boldsymbol{h} \in \mathbb{Z}^d : \boldsymbol{Z}^\top \boldsymbol{h} \equiv \boldsymbol{\xi} \pmod{\boldsymbol{n}}\right\},$$

then the following hold.

(i) If $v \in E_\alpha(\mathbb{T}^d)$ with parameter $\alpha > 5/2$ then, for all $\boldsymbol{y} \in \mathbb{R}^n$ we have

$$\|[D, W]\, \boldsymbol{y}\|_2 \leq c_1 \|(D + I)\, \boldsymbol{y}\|_2,$$

where c_1 is a constant independent of $\boldsymbol{n}$ and $\boldsymbol{y}$.

(ii) If $v \in E_\alpha(\mathbb{T}^d)$ with parameter $\alpha > 9/2$ then, for all $\boldsymbol{y} \in \mathbb{R}^n$ we have

$$\|[D, [D, W]]\, \boldsymbol{y}\|_2 \leq c_2 \|(D + I)^2\, \boldsymbol{y}\|_2,$$

where c_2 is a constant independent of $\boldsymbol{n}$ and $\boldsymbol{y}$.

Proof. Due to the lexicographical ordering on matrices $D_{\boldsymbol{n}}, W_{\boldsymbol{n}}$ and $V_{\boldsymbol{n}}$, we operate in the same way as in Lemma 3.6. □

We note that Algorithm 1 works in the same manner by replacing the inputs to the rank-r setting and using r-dimensional FFTs. In Algorithm 2, our procedure of the time-stepping on rank-r lattice points is shown.

Algorithm 2 Strang splitting on rank-r lattice points

Input:
$\Delta t, m, g$ ▷ $m\Delta t = T$ is the final time, g is the initial condition
$\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n}) = \{\boldsymbol{h}^{(0)}, ..., \boldsymbol{h}^{(n-1)}\} \subset \mathbb{Z}^{d\times n}$ ▷ Anti-aliasing set of full cardinality
$\Lambda(\boldsymbol{Z}, \boldsymbol{n}) = \{\boldsymbol{p}^{(0)}, ..., \boldsymbol{p}^{(n-1)}\} \subset \mathbb{T}^{d\times n}$ ▷ Rank-r lattice points

$V_{\boldsymbol{n}} = \operatorname{diag}\left(\left(v(\boldsymbol{p}^{(\kappa)})\right)_{\kappa=0,\dots,n-1}\right)$ ▷ The potential matrix on lattice points
$D_{\boldsymbol{n}} = \operatorname{diag}\left((4\pi^2\|\boldsymbol{h}^{(\chi)}\|_2^2)_{\chi=0,\dots,n-1}\right)$ ▷ The Laplacian matrix
$\widehat{\boldsymbol{u}}_a^0 = \widehat{\boldsymbol{g}}_a = (\widehat{g}_a(\boldsymbol{h}^{(0)}), \dots, \widehat{g}_a(\boldsymbol{h}^{(n-1)})) = F_{\boldsymbol{n}}(g(\boldsymbol{p}^{(0)}), \dots, g(\boldsymbol{p}^{(n-1)}))/\sqrt{n}$

for $k = 1, 2, ..., m$ **do**
$\quad \widehat{\boldsymbol{u}}_a^k = F_{\boldsymbol{n}} \mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} V_{\boldsymbol{n}} \Delta t} F_{\boldsymbol{n}}^{-1} \mathrm{e}^{-\frac{\mathrm{i}\gamma}{2} D_{\boldsymbol{n}} \Delta t} F_{\boldsymbol{n}} \mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} V_{\boldsymbol{n}} \Delta t} F_{\boldsymbol{n}}^{-1} \widehat{\boldsymbol{u}}_a^{k-1}$
end for

Output: $\widehat{\boldsymbol{u}}_a^m$

3.2.6 Total time discretization error bound

Combining Theorem 3.4 with Lemmas 3.6 and 3.7 we obtain the following global error bound.

Theorem 3.8 (Total error bounds)**.** *Given a rank-r lattice $\Lambda(\boldsymbol{Z}, \boldsymbol{n})$ in canonical form with number of points $n = \prod_{i=1}^{r} n_i$, and a TDSE with a potential function $v \in E_\alpha(\mathbb{T}^d)$ with $\alpha \geq 2$ and an initial condition $g \in E_\beta(\mathbb{T}^d)$ with $\beta \geq 2$. Let $D = \frac{\gamma}{2} D_{\boldsymbol{n}}$ and $W = \frac{1}{\gamma} W_{\boldsymbol{n}}$ with $D_{\boldsymbol{n}}$ and $W_{\boldsymbol{n}} = F_{\boldsymbol{n}} V_{\boldsymbol{n}} F_{\boldsymbol{n}}^{-1}$ as defined in* (3.24) *and* (3.25)*, and with $V_{\boldsymbol{n}}$ as defined in* (3.26) *using the potential function v.*

If the anti-aliasing set $\mathcal{A}(\boldsymbol{Z}, \boldsymbol{n}) = \{\boldsymbol{h}_{\boldsymbol{\xi}} \in \mathbb{Z}^d : \boldsymbol{Z}^\top \boldsymbol{h}_{\boldsymbol{\xi}} \equiv \boldsymbol{\xi} \pmod{\boldsymbol{n}}$ for $\boldsymbol{\xi} \in \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_r}\}$, with full cardinality, is chosen such that each $\boldsymbol{h}_{\boldsymbol{\xi}}$ with $\boldsymbol{\xi} \in \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_r}$ has minimal ℓ_2 norm, i.e.,

$$\|\boldsymbol{h}_{\boldsymbol{\xi}}\|_2 = \min_{\boldsymbol{h}' \in A(\boldsymbol{Z}, \boldsymbol{n}, \boldsymbol{\xi})} \|\boldsymbol{h}'\|_2,$$

with

$$A(\boldsymbol{Z}, \boldsymbol{n}, \boldsymbol{\xi}) := \left\{\boldsymbol{h} \in \mathbb{Z}^d : \boldsymbol{Z}^\top \boldsymbol{h} \equiv \boldsymbol{\xi} \pmod{\boldsymbol{n}}\right\},$$

then, by applying the Strang Splitting

$$\widehat{\boldsymbol{u}}_a^{k+1} = \mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} W_{\boldsymbol{n}} \Delta t} \mathrm{e}^{-\frac{\mathrm{i}\gamma}{2} D_{\boldsymbol{n}} \Delta t} \mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} W_{\boldsymbol{n}} \Delta t} \widehat{\boldsymbol{u}}_a^k \qquad \text{for } k = 0, 1, \dots, m-1,$$

the following hold:

(i) If $v \in E_\alpha(\mathbb{T}^d)$ with parameter $\alpha > 5/2$, then the error is bounded for $t = k\Delta t$ by

$$\|u_a^k(\cdot) - u_a(\cdot, t)\|_{L_2} \leq \Delta t \, C_1 t \max_{0 \leq t' \leq t} \|(D + I)\, \widehat{\boldsymbol{u}}_{t'}\|_2,$$

where C_1 is a constant independent of n, k, and Δt.

(ii) If $v \in E_\alpha(\mathbb{T}^d)$ with parameter $\alpha > 9/2$, then the error is bounded for $t = k\Delta t$ by

$$\|u_a^k(\cdot) - u_a(\cdot, t)\|_{L_2} \leq (\Delta t)^2 \, C_2 t \max_{0 \leq t' \leq t} \|(D + I)^2\, \widehat{\boldsymbol{u}}_{t'}\|_2,$$

where C_2 is a constant independent of n, k and Δt.

Proof.

(i) Let us denote the Strang splitting operator by $S = \mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} W_{\boldsymbol{n}} \Delta t} \mathrm{e}^{-\frac{\mathrm{i}\gamma}{2} D_{\boldsymbol{n}} \Delta t} \mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} W_{\boldsymbol{n}} \Delta t}$ and the true solution operator by $T = \mathrm{e}^{-(\frac{\mathrm{i}}{\gamma} W_{\boldsymbol{n}} - \frac{\mathrm{i}\gamma}{2} D_{\boldsymbol{n}}) \Delta t}$. We have the following for first order convergence:

$$\begin{aligned}
\|u_a^k(\cdot) - u_a(\cdot, k\Delta t)\|_{L_2} &= \|S^k\, \widehat{\boldsymbol{g}} - T^k\, \widehat{\boldsymbol{g}}\|_2 \\
&= \left\| \sum_{j=0}^{k-1} S^{k-j-1} (S - T) T^j\, \widehat{\boldsymbol{g}} \right\|_2 \\
&\leq \left(\max_{0 \leq t' \leq t} \|(S - T)\, \widehat{\boldsymbol{u}}_{t'}\|_2 \right) \sum_{j=0}^{k-1} \left\| S^{k-j-1} \right\|_2,
\end{aligned}$$

by using a telescoping sum, for the maximum argument t' there exists $\ell \in \{0, 1, ..., k\}$ such that $t' = \ell \Delta t$. Applying Theorem 3.4 and Lemmas 3.6 and 3.7, we have $\|S\boldsymbol{y} - T\boldsymbol{y}\| \leq C_1 (\Delta t)^2 \|(D + I)\boldsymbol{y}\|$ for the first order convergence condition and, $\|S\boldsymbol{y} - T\boldsymbol{y}\| \leq C_2 (\Delta t)^3 \|(D + I)^2 \boldsymbol{y}\|$ for the second order convergence condition, for all $\boldsymbol{y} \in \mathbb{R}^d$. Note that $\|S\|_2 \leq \|\mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} W_{\boldsymbol{n}} \Delta t}\|_2 \|\mathrm{e}^{-\frac{\mathrm{i}\gamma}{2} D_{\boldsymbol{n}} \Delta t}\|_2 \|\mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} W_{\boldsymbol{n}} \Delta t}\|_2 = 1$, because the ℓ_2 norm of a matrix is the largest singular value of the matrix, e.g.,

$$\begin{aligned}
\left\| \mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} W_{\boldsymbol{n}} \Delta t} \right\|_2 &= \sqrt{\lambda_{\max} \left(\mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} W_{\boldsymbol{n}} \Delta t} \left(\mathrm{e}^{-\frac{\mathrm{i}}{2\gamma} W_{\boldsymbol{n}} \Delta t} \right)^* \right)} \\
&= \sqrt{\lambda_{\max} \left(F_{\boldsymbol{n}} J_{v, \boldsymbol{n}} F_{\boldsymbol{n}}^{-1} \left(F_{\boldsymbol{n}} J_{v, \boldsymbol{n}} F_{\boldsymbol{n}}^{-1} \right)^* \right)} \\
&= \sqrt{\lambda_{\max}(I)} = 1,
\end{aligned}$$

where $\lambda_{\max}$ denotes the largest eigenvalue, $J_{v,\boldsymbol{n}} = \operatorname{diag}[(\mathrm{e}^{-\frac{\mathrm{i}}{2}v(\boldsymbol{x})\Delta t})_{\boldsymbol{x}\in\Lambda(\boldsymbol{Z},\boldsymbol{n})}]$ and A^* is the Hermitian conjugate of A. Hence we obtain the following for the first case:

$$\left(\sum_{j=0}^{k-1}\|S^{k-j-1}\|_2\right)\max_{0\le t'\le t}\|(S-T)\,\widehat{\boldsymbol{u}}_{t'}\|_2 \le kC_1(\Delta t)^2\max_{0\le t'\le t}\|(D+I)\,\widehat{\boldsymbol{u}}_{t'}\|_2$$

$$= C_1 t\Delta t\max_{0\le t'\le t}\|(D+I)\,\widehat{\boldsymbol{u}}_{t'}\|_2.$$

(*ii*) For second order convergence a similar argument holds and we obtain

$$\left\|u_a^k(\cdot)-u_a(\cdot,k\Delta t)\right\|_{L_2} \le kC_2(\Delta t)^3\max_{0\le t'\le t}\left\|(D+I)^2\,\widehat{\boldsymbol{u}}_{t'}\right\|_2$$

$$= C_2(\Delta t)^2 t\max_{0\le t'\le t}\left\|(D+I)^2\,\widehat{\boldsymbol{u}}_{t'}\right\|_2.$$

This concludes the proof. □

Note that this shows that the smoothness of the potential v required for second order convergence is independent of the number of dimensions. This is a big improvement compared to the results shown in [26] with respect to sparse grids, where the smoothness α needs to increase for increasing dimension to obtain second order convergence.

3.3 Numerical results for the Strang splitting

In this section, we demonstrate the method with numerical results. We particularly consider three quantities of interest: approximation error against the time step; evolution of the norm and the energy of the wave function over the time period; and the error which is caused by the physical discretization. To compare with the results from [26] using sparse grids, we choose the same experiments, but since our method allows the results to also be calculated for higher d than in [26] we extended the experiments.

3.3.1 Component-by-component construction

For constructing the rank-1 lattice and the anti-aliasing set, we employ the fast component-by-component construction for lattice sequences, see, e.g., [14].

We use the script `fastrank1expt.m`, available online [69] for fast component-by-component construction of a rank-1 lattice sequence with a prime power of points. We use powers of 2. The lattice point set is optimized for integration in the (unweighted) Korobov space with smoothness $\alpha = 1$ (in a common alternative notation this is $\alpha = 2$, as is the case for the construction script). After having obtained the generating vectors we construct the corresponding anti-aliasing sets in accordance with Lemma 3.6 in the following manner:

1. Generate all $\boldsymbol{h} \in \mathbb{Z}^d$ for which $\|\boldsymbol{h}\|_2 \leq r$ for some well chosen r.
2. Sort the points according to the ℓ_2-norm in ascending order.
3. Calculate $m_{\boldsymbol{h}} \equiv \boldsymbol{h} \cdot \boldsymbol{z} \pmod{n}$ in sorted order and add $\boldsymbol{h}$ to $\mathcal{A}(\boldsymbol{z}, n)$ if the value $m_{\boldsymbol{h}}$ has not been seen before. Repeat this step until the set has the cardinality n.

We refer to [15, Section 2.6] for iteratively constructing $\boldsymbol{h}$ in a bounded region.

To compare our results with the results in [26], we regenerated the data from that paper as accurately as possible from the graphs therein. In Figures 3.2 and 3.3, we denote with SG the results from [26] using sparse grids, and by LR our method using lattice rules. To make a fair comparison, we choose as close as possible the same number of basis functions n as in [26] whenever this is known. We calculate the number of basis functions n_{SG} for the d-dimensional sparse grid with level ℓ by

$$n_{\mathrm{SG}} = \sum_{i=0}^{\ell-1} 2^i \binom{d-1+i}{i}.$$

The corresponding numbers of basis functions for both methods and the generating vectors for the rank-1 lattice used in the experiments are exhibited in Table 3.1.

3.3.2 Convergence with respect to time step size

As is in [25, 26, 33] we consider the error of the calculated solution in terms of decreasing time steps against a reference solution. We choose two types of the initial condition g from [26], the "Gaussian" initial condition given by:

$$g_1(\boldsymbol{x}) := \left(\frac{2}{\pi\gamma}\right)^{d/4} \exp\left(-\frac{\left(2\pi x_1 - \frac{3\pi}{2}\right)^2 + \sum_{j=2}^{d} (2\pi x_j - \pi)^2}{\gamma}\right) \frac{1}{c_1},$$

d	n	$\boldsymbol{z}^\top$	n_{SG} from [26]
2	2^{18}	$(1, 100135)$	$2^{17.7}$ or $2^{19.9}$ *
	2^{20}	$(1, 443165)$	$2^{19.9}$
3	2^{22}	$(1, 1737355, 261247)$	$2^{22.9}$
	2^{25}	$(1, 12386359, 15699201)$	$2^{25.4}$
4 to 12	2^{25}	$(1, 12386359, 15699201, 6807287,$ $13966305, 6107923, 4432603, 2304135$ $7323801, 5705679, 5643703, 3867405)$	Not available

Table 3.1: Parameters of the numerical results. For (*) the level of the sparse grid is not specified as one number in [26]. For $d \geq 4$, we always choose $n = 2^{25}$, and $\boldsymbol{z}$ is chosen to be the first d components, e.g., for $d = 4$, $\boldsymbol{z}^\top = (1, 12386359, 15699201, 6807287)$.

and the "Hat" initial condition given by:

$$g_2(\boldsymbol{x}) := \left(\frac{3}{\pi\sqrt{\gamma}}\right)^{d/2} \left(1 - \frac{2}{\pi\sqrt{\gamma}} \left|2\pi x_1 - \frac{3\pi}{2}\right|\right) \prod_{j=2}^{d} \left(1 - \frac{2}{\pi\sqrt{\gamma}} |2\pi x_j - \pi|\right) \frac{1}{c_2},$$

for $\boldsymbol{x} \in [0,1)^d$ where c_1 and c_2 are normalizing constants to make the L_2 norms of both functions equal to 1. We remark that in [26], the domain was erroneously stated as $[-\pi, \pi]^d$ which would be equivalent to $[-1/2, 1/2)^d$ in our setting. However, we conclude that the actual calculation was done in $[0, 2\pi]^d$, as can be confirmed by the fact that the calculated norm of the Gaussian function was 1 in [26, Figure 6.8] therein, and the fact that the same author has exactly the same result in another paper [25] where the domain is stated as $[0, 2\pi]^d$ with the same Gaussian initial condition, which corresponds to $[0, 1)^d$ in our case. Therefore we conclude that our experiment is the same experiment as in [26]. For the potential function v, we consider a "smooth" potential function

$$v_1(\boldsymbol{x}) = \prod_{i=1}^{d} (1 - \cos(2\pi x_j)),$$

and a "harmonic" potential function

$$v_2 = \frac{1}{2} \sum_{j=1}^{d} (2\pi x_j - \pi)^2.$$

To show the time discretization error $\|u_a(\boldsymbol{x}, t) - u_a^m(\boldsymbol{x})\|_{L_2}$ at time $t = m\,\Delta t = 1$ being fixed, we calculate a reference solution $u_a^M(\boldsymbol{x})$ with the finest time step

size $\Delta t = 1/M = 1/10000$, as an approximation of $u_a(\boldsymbol{x}, t)$. Then we calculate $u_a^m(\boldsymbol{x})$ with various time step sizes $\Delta t = 1/m = 1/5, ..., 1/1000$ to be able to plot the convergence rate of $\|u_a^M(\boldsymbol{x}) - u_a^m(\boldsymbol{x})\|_{L_2}$.

The result is exhibited in Figures 3.2 and 3.3. We observe that the convergence rate for our new method consistently shows second order convergence $\mathcal{O}((\Delta t)^2)$. On the other hand the sparse grid results from [26] do not; for instance, see the case $d = 3$ with $\gamma = 0.01$. We remark that the initial condition g_1 combined with the potentials v_1 and v_2 satisfy the conditions of Lemma 3.1 and Theorem 3.8. Therefore we expect to see second order convergence in those cases. However, the hat initial condition g_2 does not satisfy the required regularity, nevertheless we have second order convergence in all cases. Moreover, our method achieves the second order convergence consistently even for high-dimensional cases, going from $d = 4$ in Figure 3.2 up to $d = 12$ in Figure 3.3. We note that for $d = 10$ and $d = 12$ the convergence graph for the potential v_1 does show some irregular behaviour. This comes from the numerical exuberance of the function v_1 itself when the dimension is high; the function rapidly increases to 2^d when the position $\boldsymbol{x}$ is close to $(1/2, \ldots, 1/2)$. This phenomenon does not happen with the harmonic potential v_2, which is more relevant for physics applications.

3.3.3 Norm and energy conservation

The TDSE, as a physical system, needs to conserve the norm and energy of the system. To test our algorithm we look at how well these quantities are preserved numerically. Denote the Hamiltonian by $H := -\frac{1}{2}\gamma\nabla^2 + \frac{1}{\gamma}v$, then $\frac{\partial u}{\partial t} = -\mathrm{i}Hu$. We study the time evolution of the L_2 norm of the wave function $\|u_a\|_{L_2}$ and the energy $\langle Hu_a, u_a\rangle_{L_2}$, where $\langle\ ,\ \rangle_{L_2}$ denotes the Hermitian inner product in the L_2 space. These two quantities are supposed to be conserved over the time period since

$$\frac{\partial}{\partial t}\langle u, u\rangle_{L_2} = \langle -\mathrm{i}Hu, u\rangle_{L_2} + \langle u, -\mathrm{i}Hu\rangle_{L_2} = 0,$$

and

$$\frac{\partial}{\partial t}\langle Hu, u\rangle_{L_2} = \langle -\mathrm{i}Hu, Hu\rangle_{L_2} + \langle Hu, -\mathrm{i}Hu\rangle_{L_2} = 0,$$

for the self-adjoint Hamiltonian H. For the self-adjointness of the Hamiltonian, we refer to [74]. Our numerical results are presented in Figure 3.4. To compare with the result from [26], we traced the graph therein, but we also need to remark that the absolute value in there was not accurate; the axis of the graph in [26] is not informative enough for this purpose. However, since the value of (max − min)/mean was exhibited in the article, we can compare the variation.

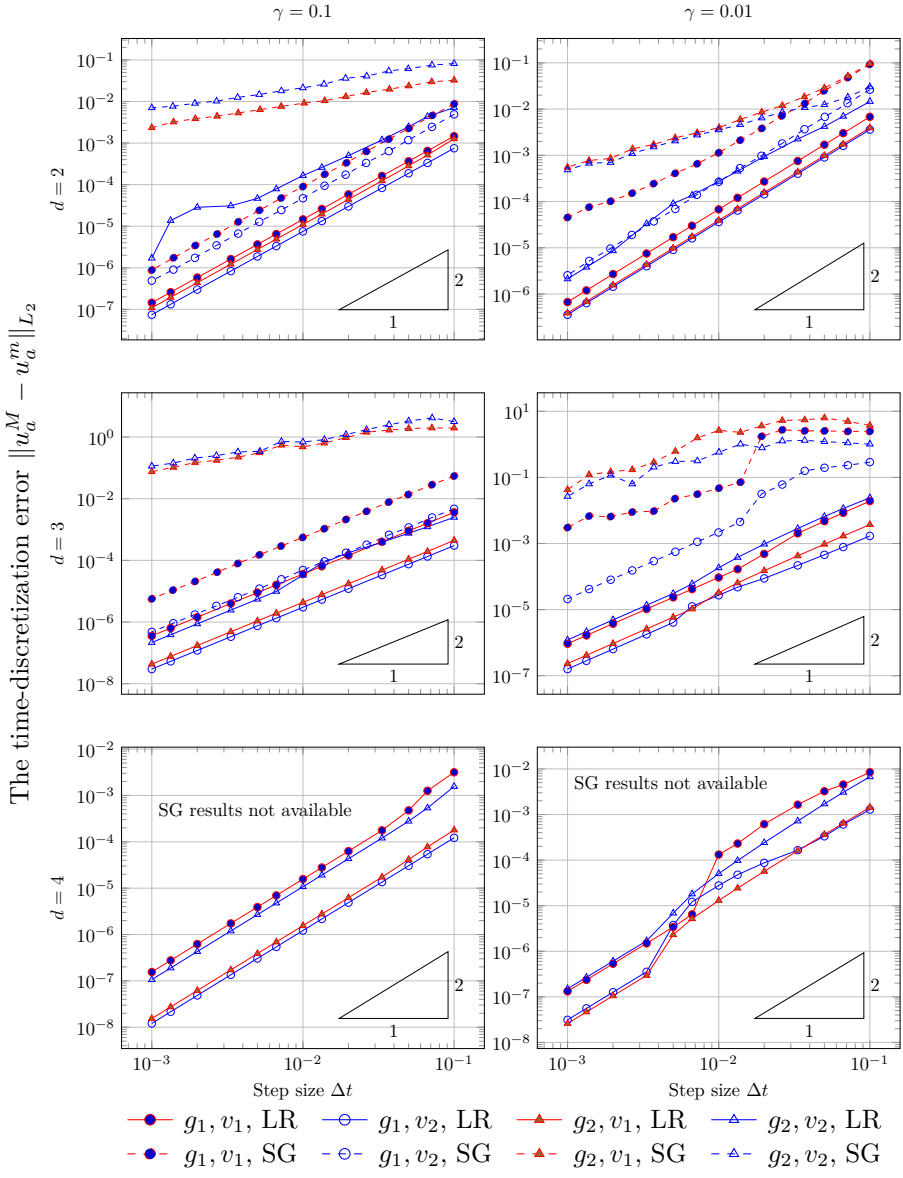

Figure 3.2: The time-discretization error. Our method (LR) is presented by the solid line, and the results by sparse grid (SG) from [26] by the dotted line. Note that the initial condition g_2 does not satisfy the regularity condition.

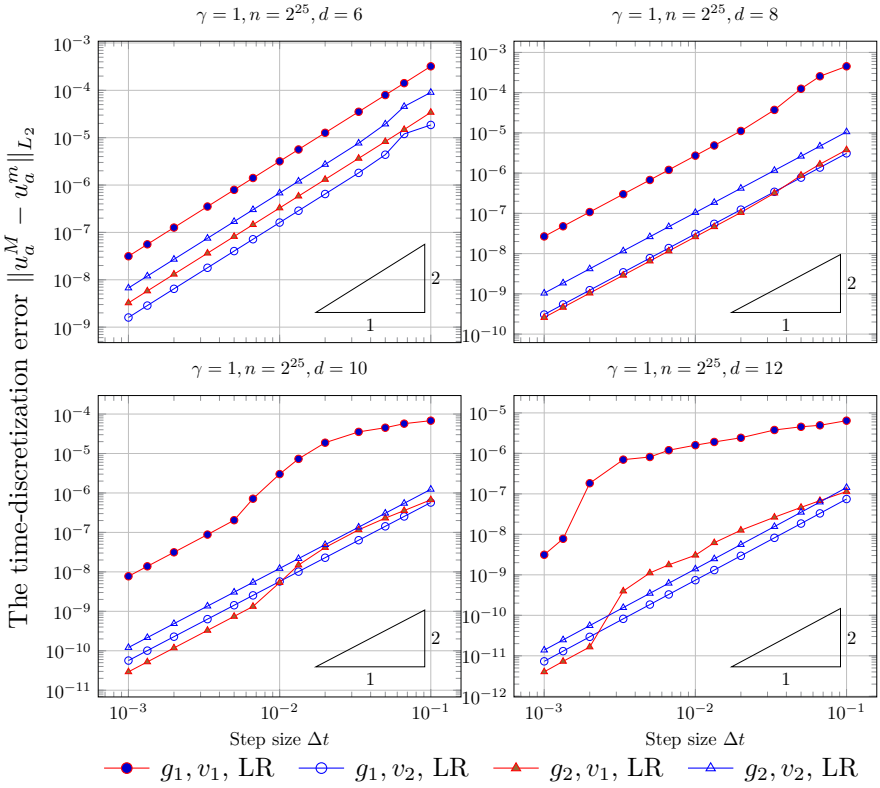

Figure 3.3: The time-discretization error in high-dimensional cases. Results by sparse grid [26] is not available for these higher-dimensional cases.

Therefore, we plot the time-evolution of the norm and the energy where the initial values are adjusted to zero.

In Figure 3.4 we see the two quantities are conserved much more accurately using our algorithm than when using the sparse grid approach in [26]. We calculate the quantity $\delta := (\max - \min)/\text{mean}$ to give an indication of the variation. Our method conserves more accurately than the sparse grid approach, for the norm conservation we have a factor of 10^{-5} smaller variation and for the energy conservation we have a factor of 10^{-4}.

The reason of the stability of our method is coming from the unitarity of the Fourier transform on our lattice points. Due to unitarity, the potential operator in the frequency domain, $F_{\boldsymbol{n}} V_{\boldsymbol{n}} F_{\boldsymbol{n}}^{-1}$, becomes Hermitian. Therefore the operator matrix $F_{\boldsymbol{n}} V_{\boldsymbol{n}} F_{\boldsymbol{n}}^{-1} + D_{\boldsymbol{n}}$ is also Hermitian and hence the spectral theorem tells

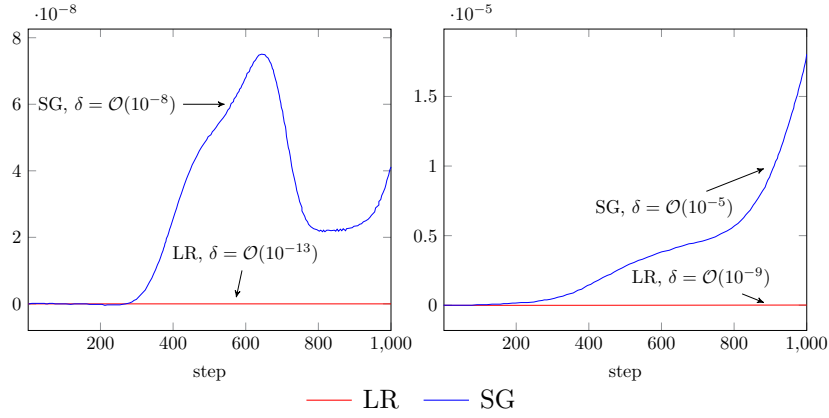

Figure 3.4: Variation of the norm (left) and the energy (right) for $\gamma = 0.5, d = 5$ with g_1 and v_2.

us that the eigenvalues of the operator matrix are all real. Finally, the time evolution operator is norm and energy conserving, i.e., $\|\mathrm{e}^{-\frac{\mathrm{i}}{\gamma}W_{\boldsymbol{n}}\,t-\frac{\mathrm{i}\gamma}{2}D_{\boldsymbol{n}}\,t}\|_2 = 1$. In contrast, the Fourier transform on the sparse grid in [26] is not unitary. The lack of unitarity can lead to numerical issues and can even lead to have the exponential error growth, instead of linear, in time [51, Section III.1.4].

3.3.4 Discussion on the initial discretization

Here we study the initial error which is caused by the initial discretization in space. The total mean square error of the initial (spatial) discretization is given by

$$\begin{aligned} e_{\text{total}}^2 &= \|g - g_a\|_{L_2}^2 \\ &= \int_{[0,1]^s} \left| \sum_{\boldsymbol{h}\in\mathbb{Z}^d} \widehat{g}(\boldsymbol{h})\, \exp(2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}) - \sum_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)} \widehat{g}_a(\boldsymbol{h})\, \exp(2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}) \right|^2 \mathrm{d}\boldsymbol{x} \\ &= \sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{z},n)} |\widehat{g}(\boldsymbol{h})|^2 + \sum_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)} |\widehat{g}(\boldsymbol{h}) - \widehat{g}_a(\boldsymbol{h})|^2. \end{aligned}$$

We plot the error e_{total} in Figure 3.5 with different dimensionality for the Gaussian initial condition. Approximating functions still requires many basis functions when the dimension becomes higher. However, intuitively we might

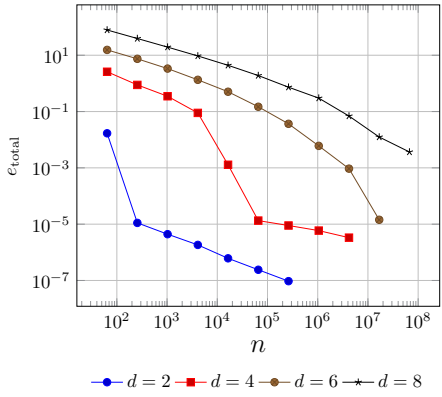

Figure 3.5: The initial discretization error e_{total} for $\gamma = 1$ with Gaussian initial condition g_1.

argue that our way of choosing the basis functions according to the ℓ_2 distance works well particularly for the Gaussian initial condition since the magnitude of the Fourier coefficients of a Gaussian is also a Gaussian (i.e., only depends on the ℓ_2 norm of the frequency, and decays exponentially fast).

3.4 The total error of full discretization

The total error of the method is coming from the discretization both in space and time. Here we recall our notation for approximating the solution:

1. $u(\boldsymbol{x}, t)$ is the true solution of (3.1);
2. $u_a(\boldsymbol{x}, t)$ is the spatially discretized solution including the dynamics as (3.17);
3. $u_a^k(\boldsymbol{x})$ is the fully discretized solution with Strang splitting (3.21).

First we denote by $\mathcal{I}_{\boldsymbol{n}}$ the interpolation operator on the lattice points, for a function f,

$$\mathcal{I}_{\boldsymbol{n}}(f)(\boldsymbol{x}, t) := \sum_{\boldsymbol{h} \in \mathcal{A}(\boldsymbol{Z}, \boldsymbol{n})} \widehat{f}_a(\boldsymbol{h}, t) \, \exp(2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{x}) \tag{3.27}$$

where,

$$\widehat{f}_a(\boldsymbol{h}, t) := \frac{1}{n} \sum_{\boldsymbol{p} \in \Lambda(\boldsymbol{Z}, \boldsymbol{n})} f(\boldsymbol{p}, t) \, \exp(-2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{p}).$$

By using the interpolation operator, we can bound

$$\|u(\cdot,t)-u_a(\cdot,t)\|_{L_2} \le \|u(\cdot,t)-\mathcal{I}_{\boldsymbol{n}}(u)(\cdot,t)\|_{L_2} + \|\mathcal{I}_{\boldsymbol{n}}(u)(\cdot,t)-u_a(\cdot,t)\|_{L_2}.$$

The error $\|u_a(\cdot,t)-u_a^k(\cdot)\|_{L_2}$ is already bounded by Theorem 3.8. Using the triangle inequality we can then bound the total error.

Theorem 3.9 (Total error)**.** *Given a rank-r lattice $\Lambda(\boldsymbol{Z},\boldsymbol{n})$ in canonical form with the number of points $n=\prod_{i=1}^r n_i$, and a TDSE with a potential function $v\in E_\alpha(\mathbb{T}^d)$ with $\alpha>9/2$ and an initial condition $g\in E_\beta(\mathbb{T}^d)$ with $\beta\ge 2$. Let $D=\frac{\gamma}{2}D_{\boldsymbol{n}}$ and $W=\frac{1}{\gamma}W_{\boldsymbol{n}}$ with $D_{\boldsymbol{n}}$ and $W_{\boldsymbol{n}}=F_{\boldsymbol{n}}V_{\boldsymbol{n}}F_{\boldsymbol{n}}^{-1}$ as defined in* (3.24) *and* (3.25)*, and with $V_{\boldsymbol{n}}$ as defined in* (3.26) *using the potential function v.*

If the anti-aliasing set $\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})=\{\boldsymbol{h}_{\boldsymbol{\xi}}\in\mathbb{Z}^d : \boldsymbol{Z}^\top\boldsymbol{h}_{\boldsymbol{\xi}}\equiv\boldsymbol{\xi} \pmod{\boldsymbol{n}}$ for $\boldsymbol{\xi}\in \mathbb{Z}_{n_1}\oplus\cdots\oplus\mathbb{Z}_{n_r}\}$, with full cardinality, is chosen such that each $\boldsymbol{h}_{\boldsymbol{\xi}}$ with $\boldsymbol{\xi}\in\mathbb{Z}_{n_1}\oplus\cdots\oplus\mathbb{Z}_{n_r}$ has minimal ℓ_2 norm, i.e.,

$$\|\boldsymbol{h}_{\boldsymbol{\xi}}\|_2 = \min_{\boldsymbol{h}'\in A(\boldsymbol{Z},\boldsymbol{n},\boldsymbol{\xi})}\|\boldsymbol{h}'\|_2, \tag{3.28}$$

with

$$A(\boldsymbol{Z},\boldsymbol{n},\boldsymbol{\xi}) := \big\{\boldsymbol{h}\in\mathbb{Z}^d : \boldsymbol{Z}^\top\boldsymbol{h}\equiv\boldsymbol{\xi} \pmod{\boldsymbol{n}}\big\},$$

then the following bound holds:

$$\begin{aligned}
&\|u(\cdot,t)-u_a^k(\cdot)\|_{L_2}\\
&\le \|u(\cdot,t)-\mathcal{I}_{\boldsymbol{n}}(u)(\cdot,t)\|_{L_2} + \|\mathcal{I}_{\boldsymbol{n}}(u)(\cdot,t)-u_a(\cdot,t)\|_{L_2} + \|u_a(\cdot,t)-u_a^k(\cdot)\|_{L_2}\\
&\le 2\sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})}|\widehat{u}(\boldsymbol{h},t)| + \frac{t\gamma}{2}\max_{0\le t'\le t}\sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})}\|\boldsymbol{h}\|_2^2\,|\widehat{u}(\boldsymbol{h},t')|\\
&\qquad + (\Delta t)^2\,C_2 t\max_{0\le t'\le t}\|(D+I)^2\,\widehat{\boldsymbol{u}}_{t'}\|_2,
\end{aligned}$$

where C_2 is a constant independent of n, k and Δt.

Proof. To bound the error, we follow a similar way of the proof for [51, Theorem 1.8] where the one-dimensional pseudospectral Fourier method for the TDSE is analysed. Applying the interpolation operator to (3.1) on both sides, we have

$$\begin{aligned}
\frac{\partial\mathcal{I}_{\boldsymbol{n}}(u)(\boldsymbol{x},t)}{\partial t} &= \frac{\gamma\mathrm{i}}{2}\,\mathcal{I}_{\boldsymbol{n}}(\nabla^2 u)(\boldsymbol{x},t) - \frac{\mathrm{i}}{\gamma}\mathcal{I}_{\boldsymbol{n}}(v\,u)(\boldsymbol{x},t)\\
&= \frac{\gamma\mathrm{i}}{2}\,(\nabla^2\mathcal{I}_{\boldsymbol{n}}(u))(\boldsymbol{x},t) - \frac{\mathrm{i}}{\gamma}\mathcal{I}_{\boldsymbol{n}}(v\,(\mathcal{I}_{\boldsymbol{n}}(u)))(\boldsymbol{x},t) + \delta_{\boldsymbol{n}}(\boldsymbol{x},t),
\end{aligned} \tag{3.29}$$

where $\delta_{\boldsymbol{n}}(\boldsymbol{x},t) = \frac{\gamma \mathrm{i}}{2}\,\mathcal{I}_{\boldsymbol{n}}(\nabla^2 u(\boldsymbol{x},t)) - \frac{\gamma \mathrm{i}}{2}\,(\nabla^2 \mathcal{I}_{\boldsymbol{n}}(u)(\boldsymbol{x},t))$ is called the *defect* which can be seen as a commutator of the interpolation operator and the Laplacian applied to the solution, and we used $\mathcal{I}_{\boldsymbol{n}}(v\,u)(\boldsymbol{x},t) = \mathcal{I}_{\boldsymbol{n}}(v\,(\mathcal{I}_{\boldsymbol{n}}(u)))(\boldsymbol{x},t)$. At the same time, we can express the dynamics of $\widehat{u}_a(\boldsymbol{h},t)$ given in (3.17) in the original space by

$$\frac{\partial u_a(\boldsymbol{x},t)}{\partial t} = \frac{\gamma \mathrm{i}}{2}\,\nabla^2 u_a(\boldsymbol{x},t) - \frac{\mathrm{i}}{\gamma}\mathcal{I}_{\boldsymbol{n}}(v\,u_a)(\boldsymbol{x},t). \tag{3.30}$$

Here we see two different dynamics in (3.29) and (3.30), therefore, by letting $\theta_{\boldsymbol{n}}(\boldsymbol{x},t) := \mathcal{I}_{\boldsymbol{n}}(u)(\boldsymbol{x},t) - u_a(\boldsymbol{x},t)$ and comparing (3.29) with (3.30), we have

$$\frac{\partial \theta(\boldsymbol{x},t)}{\partial t} = \frac{\gamma \mathrm{i}}{2}\,\nabla^2 \theta(\boldsymbol{x},t) - \frac{\mathrm{i}}{\gamma}\mathcal{I}_{\boldsymbol{n}}(v\,\theta)(\boldsymbol{x},t) + \delta_{\boldsymbol{n}}(\boldsymbol{x},t). \tag{3.31}$$

We note that $\theta(\boldsymbol{x},0) = 0$. Using the relation

$$\frac{1}{2}\frac{\partial |\theta(\boldsymbol{x},t)|^2}{\partial t} = \mathrm{Re}\left(\theta(\boldsymbol{x},t)\overline{\frac{\partial \theta(\boldsymbol{x},t)}{\partial t}}\right),$$

where $\overline{x}$ denotes the complex conjugate, and using the chain rule we obtain the following inequality

$$\begin{aligned}
&\|\theta(\cdot,t)\|_{L_2}\frac{\partial \|\theta(\cdot,t)\|_{L_2}}{\partial t} = \frac{1}{2}\frac{\partial \|\theta(\cdot,t)\|_{L_2}^2}{\partial t} \\
&= \mathrm{Re}\left(\left\langle \theta(\cdot,t), \frac{\partial \theta(\cdot,t)}{\partial t}\right\rangle_{L_2}\right) \\
&= \mathrm{Re}\left(\left\langle \theta(\cdot,t), \frac{\gamma \mathrm{i}}{2}\,\nabla^2\theta(\cdot,t) - \frac{\mathrm{i}}{\gamma}\mathcal{I}_{\boldsymbol{n}}(v\,\theta)(\cdot,t)\right\rangle_{L_2}\right) + \mathrm{Re}\left(\left\langle \theta(\cdot,t), \delta_{\boldsymbol{n}}(\cdot,t)\right\rangle_{L_2}\right) \\
&= \mathrm{Re}\left(\left\langle \theta(\cdot,t), \delta_{\boldsymbol{n}}(\cdot,t)\right\rangle_{L_2}\right) \\
&\le \|\theta(\cdot,t)\|_{L_2}\|\delta_{\boldsymbol{n}}(\cdot,t)\|_{L_2},
\end{aligned}$$

where we used the fact that our discrete Fourier matrix $F_{\boldsymbol{n}}$ is unitary which makes the operator $\frac{\gamma \mathrm{i}}{2}\nabla^2(\cdot) - \frac{\mathrm{i}}{\gamma}\mathcal{I}_{\boldsymbol{n}}(v\ \ (\cdot))(\boldsymbol{x},t)$ (e.g., (3.17) and (3.30)) self-adjoint, and consequently the energy $\left\langle \theta(\cdot,t), \frac{\gamma}{2}\,\nabla^2\theta(\cdot,t) - \frac{1}{\gamma}\mathcal{I}_{\boldsymbol{n}}(v\,\theta)(\cdot,t)\right\rangle_{L_2}$ is always real. Dividing both sides of the above inequality by $\|\theta(\cdot,t)\|_{L_2}$ and integrating over time, we obtain

$$\int_0^t \frac{\partial \|\theta(\cdot,t')\|_{L_2}}{\partial t}\,\mathrm{d}t' = \|\theta(\cdot,t)\|_{L_2} \le \int_0^t \|\delta_{\boldsymbol{n}}(\cdot,t')\|_{L_2}\,\mathrm{d}t'.$$

Using Lemma 3.3, we can explicitly calculate the defect

$$\delta_{\boldsymbol{n}}(\boldsymbol{x},t) = \frac{\gamma \mathrm{i}}{2}\,\mathcal{I}_{\boldsymbol{n}}(\nabla^2 u(\boldsymbol{x},t)) - \frac{\gamma \mathrm{i}}{2}\,(\nabla^2 \mathcal{I}_{\boldsymbol{n}}(u)(\boldsymbol{x},t))$$

$$= \frac{\gamma \mathrm{i}}{2} \sum_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \left(\sum_{\boldsymbol{\ell}\in\Lambda^\perp(\boldsymbol{Z},\boldsymbol{n})} (\|\boldsymbol{h}+\boldsymbol{\ell}\|_2^2 - \|\boldsymbol{h}\|_2^2)\, \widehat{u}(\boldsymbol{h}+\boldsymbol{\ell},t) \right) \exp(2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{x}).$$

For $\boldsymbol{\ell}=\boldsymbol{0}$, all terms become zero and we drop those. Now we use (3.28) such that $|\,\|\boldsymbol{h}+\boldsymbol{\ell}\|_2^2 - \|\boldsymbol{h}\|_2^2\,| \le \|\boldsymbol{h}+\boldsymbol{\ell}\|_2^2$ for any $\boldsymbol{h}\in\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})$ and $\boldsymbol{\ell}\in\Lambda^\perp(\boldsymbol{Z},\boldsymbol{n})$. This means

$$\|\delta_{\boldsymbol{n}}(\cdot,t)\|_{L_2} \le \frac{\gamma}{2} \left(\sum_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \left(\sum_{\boldsymbol{0}\neq\boldsymbol{\ell}\in\Lambda^\perp(\boldsymbol{Z},\boldsymbol{n})} \|\boldsymbol{h}+\boldsymbol{\ell}\|_2^2\, |\widehat{u}(\boldsymbol{h}+\boldsymbol{\ell},t)| \right)^2 \right)^{1/2}$$

$$\le \frac{\gamma}{2} \sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \|\boldsymbol{h}\|_2^2\, |\widehat{u}(\boldsymbol{h},t)|.$$

Therefore, we have

$$\|\theta(\cdot,t)\|_{L_2} \le \int_0^t \|\delta_{\boldsymbol{n}}(\cdot,t')\|_{L_2}\,\mathrm{d}t' \le t \max_{0\le t'\le t} \frac{\gamma}{2} \sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \|\boldsymbol{h}\|_2^2\, |\widehat{u}(\boldsymbol{h},t')|.$$

For the remaining term $\|u(\cdot,t) - \mathcal{I}_{\boldsymbol{n}}(u)(\cdot,t)\|_{L_2}$, we have

$$\|u(\cdot,t) - \mathcal{I}_{\boldsymbol{n}}(u)(\cdot,t)\|_{L_2}$$

$$= \left(\sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{z},n)} |\widehat{u}(\boldsymbol{h},t)|^2 + \sum_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)} \left| \sum_{\boldsymbol{0}\neq\boldsymbol{\ell}\in\Lambda^\perp(\boldsymbol{Z},\boldsymbol{n})} \widehat{u}(\boldsymbol{h}+\boldsymbol{\ell},t) \right|^2 \right)^{1/2}$$

$$\le \left(\sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{z},n)} |\widehat{u}(\boldsymbol{h},t)|^2 \right)^{1/2} + \left(\sum_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)} \left| \sum_{\boldsymbol{0}\neq\boldsymbol{\ell}\in\Lambda^\perp(\boldsymbol{Z},\boldsymbol{n})} \widehat{u}(\boldsymbol{h}+\boldsymbol{\ell},t) \right|^2 \right)^{1/2}$$

$$\le \sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{z},n)} |\widehat{u}(\boldsymbol{h},t)| + \left(\sum_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)} \left(\sum_{\boldsymbol{0}\neq\boldsymbol{\ell}\in\Lambda^\perp(\boldsymbol{Z},\boldsymbol{n})} |\widehat{u}(\boldsymbol{h}+\boldsymbol{\ell},t)| \right)^2 \right)^{1/2}$$

$$\le 2 \sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{z},n)} |\widehat{u}(\boldsymbol{h},t)|. \tag{3.32}$$

Using the triangle inequality, we obtain

$$\|u(\cdot,t) - u_a(\cdot,t)\|_{L_2} \leq \|u(\cdot,t) - \mathcal{I}_{\boldsymbol{n}}(u)(\cdot,t)\|_{L_2} + \|\mathcal{I}_{\boldsymbol{n}}(u)(\cdot,t) - u_a(\cdot,t)\|_{L_2}$$

$$\leq 2 \sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} |\widehat{u}(\boldsymbol{h},t)| + \frac{t\gamma}{2} \max_{0\leq t'\leq t} \sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \|\boldsymbol{h}\|_2^2 \, |\widehat{u}(\boldsymbol{h},t')|.$$

This completes the proof. □

The bound (3.32) is further bounded by

$$\|u(\cdot,t) - u_a(\cdot,t)\|_{L_2} \leq (2 + \frac{t\gamma}{2}) \max_{0\leq t'\leq t} \sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{Z},\boldsymbol{n})} \|\boldsymbol{h}\|_2^2 \, |\widehat{u}(\boldsymbol{h},t')|.$$

This is similar to the result of [51, Theorem 1.8] for the one-dimensional case which states

$$\|u(\cdot,t) - u_a(\cdot,t)\|_{L_2} \leq C(1+t) \max_{0\leq t'\leq t} \|\frac{\partial^2 u(\cdot,t')}{\partial x^2} - (\frac{\partial^2 \mathcal{I}_{\boldsymbol{n}}(u)}{\partial x^2})(\cdot,t')\|_{L_2}.$$

For certain function spaces, the approximation errors of lattice points are explicitly known, e.g., [9, 45]. It might be possible to construct approximation lattices according to the referenced papers and then to extend the frequency index set to fulfill the needed conditions. However, this is not the focus of the present chapter. The focus is the interplay between the spatial discretization and the time-stepping error, because the time-stepping error itself is heavily affected by the spatial discretization as we can see from the comparison with [26].

3.5 Chapter overview

We approximated the solution of the time-dependent Schrödinger equation by using rank-1 and rank-r lattices for the space discretization and Strang splitting for the time discretization. We combined the anti-aliasing set of the lattices together with FFTs to obtain both theoretical advantages and computational efficiency. We showed that the time discretization of our method has second-order convergence for a potential function $v \in E_\alpha(\mathbb{T}^d)$ with $\alpha > 9/2$ which is independent of the dimension d. The numerical experiments confirm the theory. We observed second order convergence with respect to the time step in cases up to 12 dimensions. Previous results based on sparse grids [26] have difficulty for cases higher than 5 dimensions.

Here we also remark limitations of our method. We exploited the structure of lattices to mitigate the curse of dimensionality, but we do not completely remove the curse. This means, we can solve rather higher-dimensional problems than regular grids and sparse grids in [26] can, but not too high. Also, our focus of the present chapter is on the time-dependent problems. The algorithm is especially made for obtaining a small time-stepping error. Therefore, we cannot expect that our method works better for the time-independent problems than existing methods such as [3, 28], for this the lattice points have to be constructed with this in mind.

Our method can be applied to different problems which would be more interesting for physics applications. In Chapter 5, we apply our method to the time-dependent non-linear Schrödinger equation for simulating Bose – Einstein condensates. In [90], Thalhammer showed that pseudospectral Fourier methods using regular grids with exponential splitting can obtain the higher order convergence in time stepping. We change the regular grid to lattice points to obtain the efficient simulation scheme with keeping the same convergence order. Another possibility is using our method for time-dependent potentials. For instance, the time-dependent harmonic oscillator is used for considering multiphoton excitation of molecules, see [41].

Chapter 4

Rank-1 lattices and higher-order exponential splitting for the time-dependent Schrödinger equation

In this chapter, we focus on two perspectives; high-dimensionality and higher-order convergence in time stepping. For the first point, Gradinaru [26] proposed to use sparse grids for the physical space. In Chapter 3, we used lattice points to prove second order convergence for the time discretization using Strang splitting and numerically compared results with the sparse grid approach from [26]. The numerical result using rank-1 lattices showed the expected second order convergence even up to 12 dimensions. Hence rank-1 lattice points perform thereby much better than the sparse grid approach. The second point, higher-order convergence in time stepping, is successfully achieved by Thalhammer [89] using higher-order exponential operator splitting. Therein, the spatial discretization was done by a full grid and therefore was limited to lower dimensional cases ($d \leq 3$). The contents of this chapter is part of the published paper [87].

Continuing from Chapter 3, we make extensive use of interpolation properties of lattice points (Theorem 3.2 and Lemma 3.3). We remark that Theorem 3.2 can

also be understood in terms of Fourier analysis on a finite Abelian group where the group, normally denoted as G, is the rank-1 lattice point set $\Lambda(\boldsymbol{z}, n)$ and the associated character group (Pontryagin dual) $\widehat{G} := \{\exp(2\pi \mathrm{i}\boldsymbol{h}\cdot\circ) : \boldsymbol{h} \in \mathcal{A}(\boldsymbol{z}, n)\}$ with $|\mathcal{A}(\boldsymbol{z}, n)| = n$. The (dual) character property is then to be understood as orthonormality of $\widehat{G}$ on $L_2(G)$. The interpolation condition can be seen as the representability of functions by using Fourier series. Due to this structure, the Plancherel theorem also holds:

$$\sum_{\boldsymbol{p}\in\Lambda(\boldsymbol{z},n)} f(\boldsymbol{p})\,\overline{g(\boldsymbol{p})} = \sum_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)} \widehat{f_a}(\boldsymbol{h})\,\overline{\widehat{g_a}(\boldsymbol{h})}$$

for $f, g \in L_2(G)$.

For readers who are not familiar with Fourier transforms on a rank-1 lattice, one intuitive way of seeing why one-dimensional FFTs are available is the following. The usual one-dimensional Fourier transform for equidistant points which is a scalar multiple of a unitary Fourier transform, for a function $f : \mathbb{T} \to \mathbb{C}$, can be written as

$$\widehat{f}(h) = \frac{1}{n}\sum_{k=0}^{n-1} f(k/n)\exp(-2\pi \mathrm{i}hk/n),$$

and the inverse

$$f(k/n) = \sum_{h=0}^{n-1} \widehat{f}(h)\exp(2\pi \mathrm{i}hk/n).$$

Now we see that the Fourier transform on a rank-1 lattice has the exact same structure for a function $f : \mathbb{T}^d \to \mathbb{C}$,

$$\widehat{f}(\boldsymbol{h}_\xi) = \frac{1}{n}\sum_{k=0}^{n-1} f(\boldsymbol{p}_k)\exp(-2\pi \mathrm{i}\boldsymbol{h}_\xi\cdot\boldsymbol{p}_k) = \frac{1}{n}\sum_{k=0}^{n-1} f(\boldsymbol{p}_k)\exp(-2\pi \mathrm{i}\xi k/n),$$

and

$$f(\boldsymbol{p}_k) = \sum_{\xi=0}^{n-1} \widehat{f}(\boldsymbol{h}_\xi)\exp(2\pi \mathrm{i}\boldsymbol{h}_\xi\cdot\boldsymbol{p}_k) = \sum_{k=0}^{n-1} \widehat{f}(\boldsymbol{h}_\xi)\exp(2\pi \mathrm{i}\xi k/n),$$

where we note $\boldsymbol{p}_k = \boldsymbol{z}k/n \bmod 1$ and $\boldsymbol{h}_\xi\cdot\boldsymbol{z} \equiv \xi \pmod n$. Hence we only need one-dimensional FFTs to transform functions on $\mathbb{T}^d$.

4.1 The method

Here we generalize the results from the Strang splitting to higher-order exponential splitting schemes (also called an exponential propagator), see,

e.g., [5, 85, 89]. To describe the higher-order exponential splitting, let us consider the following ordinary differential equation:

$$y'(t) = (A + B)\, y(t), \qquad y(0) = y_0, \tag{4.1}$$

where A and B are differential operators. The solution for the equation (4.1) is $y(t) = \mathrm{e}^{(A+B)t} y_0$. However, often it is not possible to compute this exactly, and one needs to approximate the quantity with cheap computational cost. When both e^{At} and e^{Bt} can be computed easily, the higher-order exponential splitting is a powerful tool to approximate the solution $\mathrm{e}^{(A+B)t} y_0$. The approximated solution for this case is given by:

$$y(t + \Delta t) \approx \mathrm{e}^{b_1 B\, \Delta t}\, \mathrm{e}^{a_1 A\, \Delta t}\, \cdots \mathrm{e}^{b_s B\, \Delta t} \mathrm{e}^{a_s A\, \Delta t}\, y(t), \tag{4.2}$$

where a_i and b_i, $i = 1, \ldots, s$, are coefficients determined by the desired order of convergence p. In other words, if the splitting (4.2) satisfies

$$\| \mathrm{e}^{b_1 B\, \Delta t}\, \mathrm{e}^{a_1 A\, \Delta t}\, \cdots \mathrm{e}^{b_s B\, \Delta t} \mathrm{e}^{a_s A\, \Delta t}\, y(t) - \mathrm{e}^{(A+B)\Delta t}\, y(t) \|_X \le C (\Delta t)^{p+1}, \tag{4.3}$$

for some normed space X, where the constant C is independent of Δt, then the splitting is said to have p-th order. The number of steps s and the coefficients a_i, b_i can be determined according to the order p, see [29] for details. We evolve the time using this discretization from time 0, i.e.,

$$y_{k+1} = \mathrm{e}^{b_1 B\, \Delta t}\, \mathrm{e}^{a_1 A\, \Delta t}\, \cdots \mathrm{e}^{b_s B\, \Delta t} \mathrm{e}^{a_s A\, \Delta t}\, y_k, \qquad y_{\{k=0\}} = y_0.$$

By summing up the local errors (4.3) of each step $k = 1, \ldots, m$, where $t = m\Delta t$, gives the total error:

$$\| y_m - y(t) \|_X \le C\, m\Delta t\, (\Delta t)^p = C t (\Delta t)^p.$$

We call this quantity the total error in the L_2 sense, and this is the reason why the splitting is called to be of p-th order. The error coming from the exponential splitting can be related to commutators of two operators A and B, namely $[A, B] := AB - BA$, $[A, [A, B]] := A^2 B - 2ABA + BA^2$, etc. We introduce the notation for the p-th commutator by following [89]:

$$\mathrm{ad}_A^p(B) = [A, \mathrm{ad}_A^{p-1}(B)], \;\; \mathrm{ad}_A^0(B) = B,$$

where $p \ge 1$. When the p-th commutator is bounded, it is known that the p-th order exponential splitting gives the desired order, see [89, Lemma 1 and Theorem 1]. We also refer to [33, Theorem 2.1] for the second-order splitting (namely, Strang splitting) in a more abstract setting.

We approximate the solution of the ordinary differential equation (3.17)

$$\widehat{\boldsymbol{u}}_t = \mathrm{e}^{-\frac{\mathrm{i}}{\gamma} W_n\, t - \frac{\mathrm{i}\gamma}{2} D_n\, t}\, \widehat{\boldsymbol{u}}_0,$$

by applying the higher-order exponential splitting method (4.2):

$$\widehat{\boldsymbol{u}}_a^{k+1} = \mathrm{e}^{-b_1 \frac{\mathrm{i}}{\gamma} W_n \, \Delta t} \, \mathrm{e}^{-a_1 \frac{\mathrm{i}\gamma}{2} D_n \, \Delta t} \, \cdots \mathrm{e}^{-b_s \frac{\mathrm{i}}{\gamma} W_n \, \Delta t} \mathrm{e}^{-a_s \frac{\mathrm{i}\gamma}{2} D_n \, \Delta t} \widehat{\boldsymbol{u}}_a^k \tag{4.4}$$

for $k = 0, 1, \ldots, m-1$, where again the coefficients a_i, b_i are determined according to the desired order of convergence, and

$$\mathrm{e}^{-\frac{\mathrm{i}}{2} W_n \Delta t} = F_n \, \mathrm{diag}\left((\mathrm{e}^{-\frac{\mathrm{i}}{2} v(\boldsymbol{p}_k) \Delta t})_{k=0,\ldots,n-1} \right) F_n^{-1}.$$

The approximated solution at the time $t = k\Delta t$ is then obtained by stepping time Δt iteratively by (4.4). In the following we show the *commutator bounds* which correspond to [89, Hypothesis 3] and lead us to the total bound as in [89, Theorem 1].

Theorem 4.1 (p-th commutator bound and total error bound)**.** *Given a rank-1 lattice with generating vector $\boldsymbol{z} \in \mathbb{Z}^d$ and modulus n and a TDSE with a potential function $v \in E_\alpha(\mathbb{T}^d)$ with $\alpha > 2p + 1/2$ and an initial condition $g \in E_\beta(\mathbb{T}^d)$ with $\beta \geq 2$. Let $D = \frac{\gamma}{2} D_n$ and $W = \frac{1}{\gamma} W_n$ with D_n and $W_n = F_n V_n F_n^{-1}$ as defined in* (3.18) *and* (3.20)*, and with V_n as defined in* (3.19) *using the potential function v.*

If the anti-aliasing set $\mathcal{A}(\boldsymbol{z}, n) = \{\boldsymbol{h}_\xi \in \mathbb{Z}^d : \boldsymbol{h}_\xi \cdot \boldsymbol{z} \equiv \xi \pmod{n}$ for $\xi = 0, \ldots, n-1\}$, with full cardinality, is chosen such that it has minimal ℓ_2 norm, i.e.,

$$\|\boldsymbol{h}_\xi\|_2 = \min_{\boldsymbol{h}' \in A(\boldsymbol{z}, n, \xi)} \|\boldsymbol{h}'\|_2, \tag{4.5}$$

with

$$A(\boldsymbol{z}, n, \xi) := \left\{ \boldsymbol{h} \in \mathbb{Z}^d : \boldsymbol{h} \cdot \boldsymbol{z} \equiv \xi \pmod{n} \right\},$$

then for all $\boldsymbol{y} \in \mathbb{R}^n$ we have the following bound for the p-th commutator:

$$\|\mathrm{ad}_D^p(W) \, \boldsymbol{y}\|_2 \leq c \, \|(D + I)^p \, \boldsymbol{y}\|_2,$$

where c is a constant independent of n and $\boldsymbol{y}$.
This commutator condition and [89, Theorem 1] directly give us the total error bound for (3.17):

$$\|\widehat{\boldsymbol{u}}_t - \widehat{\boldsymbol{u}}_a^m\|_2 \leq C \|\widehat{\boldsymbol{u}}_0 - \widehat{\boldsymbol{u}}_a^0\|_2 + C'(\Delta t)^p \|(D + I)^p \widehat{\boldsymbol{u}}_0\|_2,$$

where $m\Delta t = t$ and the constants depend on t but not on m or Δt.

Proof. Let $M := \mathrm{ad}_D^p(W) \, (D + I)^{-p}$. Since $(D + I)^p$ is is non-singular, the claim of the theorem is equivalent to the assertion that the induced ℓ_2 norm

of the matrix $\|M\|_2 := \sup_{\mathbf{0}\neq \boldsymbol{y}\in\mathbb{R}^n} \|M\,\boldsymbol{y}\|_2/\|\boldsymbol{y}\|_2$ is bounded independent of n. Each element of the matrix M is given by,

$$M = \left(\frac{(\|\boldsymbol{h}_\xi\|_2^2 - \|\boldsymbol{h}_{\xi'}\|_2^2)^p}{\gamma(\|\boldsymbol{h}_{\xi'}\|_2^2 + c_1)^p}\, w_{\xi,\xi'} \right)_{\xi,\xi'=0,\ldots,n-1},$$

where the constant $c_1 = 1/(2\pi\gamma)^p > 0$. Now we bound $\|M\|_2$ by using $\|M\|_2 \leq \sqrt{\|M\|_1\|M\|_\infty}$. First we bound $\|M\|_1$:

$$\begin{aligned}
\|M\|_1 &= \frac{1}{\gamma} \max_{\xi'\in\mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi\neq\xi'}}^{n-1} \left| \frac{(\|\boldsymbol{h}_\xi\|_2^2 - \|\boldsymbol{h}_{\xi'}\|_2^2)^p}{(\|\boldsymbol{h}_{\xi'}\|_2^2 + c_1)^p}\, w_{\xi,\xi'} \right| \\
&\leq \frac{1}{\gamma} \max_{\xi'\in\mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi\neq\xi'}}^{n-1} \left| \frac{(\max(\|\boldsymbol{h}_\xi\|_2^{2p}, \|\boldsymbol{h}_{\xi'}\|_2^{2p})}{(\|\boldsymbol{h}_{\xi'}\|_2^2 + c_1)^p}\, w_{\xi,\xi'} \right|.
\end{aligned}$$

We notice that the diagonal components of M ($\xi = \xi'$) are always 0, hence we exclude such cases in the following argument. Because we construct the anti-aliasing set by minimizing the ℓ_2 norm (4.5), we have $\|\boldsymbol{h}_\xi\|_2 \leq \|\boldsymbol{h}'_\xi\|_2$ for any $\boldsymbol{h}'_\xi \in A(\boldsymbol{z}, n, \xi)$. In particular, this holds for $\boldsymbol{h}'_\xi = \boldsymbol{h}_{\xi-\xi'} + \boldsymbol{h}_{\xi'}$ since $(\boldsymbol{h}_{\xi-\xi'} + \boldsymbol{h}_{\xi'}) \cdot \boldsymbol{z} \equiv \xi \pmod n$ for any choice of $\xi' = 0, \ldots, n-1$. This gives us the connection between $\|\boldsymbol{h}_\xi\|_2$ and $\|\boldsymbol{h}_{\xi'}\|_2$ using $\|\boldsymbol{h}_{\xi-\xi'}\|_2$:

$$\frac{\|\boldsymbol{h}_\xi\|_2^2}{\|\boldsymbol{h}_{\xi'}\|_2^2 + c_1} \leq \frac{\|\boldsymbol{h}_{\xi'} + \boldsymbol{h}_{\xi-\xi'}\|_2^2}{\|\boldsymbol{h}_{\xi'}\|_2^2 + c_1} \leq 4\|\boldsymbol{h}_{\xi-\xi'}\|_2^2,$$

for $\xi \neq \xi'$. We continue from the above bound of $\|M\|_1$,

$$\begin{aligned}
\|M\|_1 &\leq \frac{1}{\gamma} \max_{\xi'\in\mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi\neq\xi'}}^{n-1} \left| \frac{(\max(\|\boldsymbol{h}_\xi\|_2^{2p}, \|\boldsymbol{h}_{\xi'}\|_2^{2p})}{(\|\boldsymbol{h}_{\xi'}\|_2^2 + c_1)^p}\, w_{\xi,\xi'} \right| \\
&\leq \frac{1}{\gamma} \max_{\xi'\in\mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi\neq\xi'}}^{n-1} \left| \max\left(\frac{\|\boldsymbol{h}_\xi\|_2^2}{\|\boldsymbol{h}_{\xi'}\|_2^2 + c_1}, 1 \right)^p w_{\xi,\xi'} \right| \\
&\leq \frac{1}{\gamma} \max_{\xi'\in\mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi\neq\xi'}}^{n-1} \left| \max\left(4^p \|\boldsymbol{h}_{\xi-\xi'}\|_2^{2p}, 1 \right) w_{\xi,\xi'} \right|
\end{aligned}$$

$$= \frac{1}{\gamma} \max_{\xi' \in \mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi \neq \xi'}}^{n-1} \left| \left(4^p \|\boldsymbol{h}_{\xi-\xi'}\|_2^{2p} \right) w_{\xi,\xi'} \right|$$

$$\leq \frac{4^p}{\gamma} \max_{\xi' \in \mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi \neq \xi'}}^{n-1} \|\boldsymbol{h}_{\xi-\xi'}\|_2^{2p} \left| \sum_{\boldsymbol{h} \in A(\boldsymbol{z},n,\xi-\xi')} \widehat{v}(\boldsymbol{h}) \right|$$

$$\leq \frac{4^p}{\gamma} \max_{\xi' \in \mathbb{Z}_n} \sum_{\substack{\xi=0 \\ \xi \neq \xi'}}^{n-1} \sum_{\boldsymbol{h} \in A(\boldsymbol{z},n,\xi-\xi')} \|\boldsymbol{h}\|_2^{2p} \, |\widehat{v}(\boldsymbol{h})|$$

$$\leq \frac{4^p}{\gamma} \sum_{\boldsymbol{h} \in \mathbb{Z}^d} \|\boldsymbol{h}\|_2^{2p} \, |\widehat{v}(\boldsymbol{h})|.$$

For the last inequality, we used the conjugacy decomposition (3.2). By using Cauchy–Schwarz inequality and multiplying and dividing by r_α, we have

$$\sum_{\boldsymbol{h} \in \mathbb{Z}^d} \|\boldsymbol{h}\|_2^{2p} \, |\widehat{v}(\boldsymbol{h})| \leq \left(\sum_{\boldsymbol{h} \in \mathbb{Z}^d} r_\alpha^2(\boldsymbol{h}) \, |\widehat{v}(\boldsymbol{h})|^2 \right)^{1/2} \left(\sum_{\boldsymbol{h} \in \mathbb{Z}^d} \frac{\|\boldsymbol{h}\|_2^{4p}}{r_\alpha^2(\boldsymbol{h})} \right)^{1/2}$$

$$\leq \|v\|_{E_\alpha(\mathbb{T}^d)} \left(\sum_{\boldsymbol{h} \in \mathbb{Z}^d} \frac{(\sqrt{d}\,\|\boldsymbol{h}\|_\infty)^{4p}}{r_\alpha^2(\boldsymbol{h})} \right)^{1/2}$$

$$\leq \|v\|_{E_\alpha(\mathbb{T}^d)} \left(\sum_{\boldsymbol{h} \in \mathbb{Z}^d} \frac{d^{2p}}{r_{\alpha-2p}^2(\boldsymbol{h})} \right)^{1/2}$$

$$\leq \|v\|_{E_\alpha(\mathbb{T}^d)} \left(d^{2p} \left(1 + 2\,\zeta(2\alpha - 4p)\right)^d \right)^{1/2} < \infty.$$

This means we have bounded $\|M\|_1$ independent of n. For $\|M\|_\infty$ we can proceed in a similar way to obtain

$$\|M\|_\infty = \max_{\xi \in \mathbb{Z}_n} \sum_{\substack{\xi'=0 \\ \xi' \neq \xi}}^{n-1} \left| \frac{\left(\|\boldsymbol{h}_\xi\|_2^2 - \|\boldsymbol{h}_{\xi'}\|_2^2 \right)^p}{\left(\|\boldsymbol{h}_{\xi'}\|_2^2 + c_1 \right)^p} \, w_{\xi,\xi'} \right|$$

$$\leq \frac{4^p}{\gamma} \|v\|_{E_\alpha(\mathbb{T}^d)} \left(d^{2p} \left(1 + 2\,\zeta(2\alpha - 4p)\right)^d \right)^{1/2} < \infty.$$

d	n	$\boldsymbol{z}^\top$
2	2^{16}	$(1, 100135)$
4	2^{20}	$(1, 443165, 95693, 34519)$
6	2^{24}	$(1, 6422017, 7370323, 2765761, 8055041, 2959639)$
8	2^{24}	$(1, 6422017, 7370323, 2765761, 8055041, 2959639, 7161203, 4074015)$

Table 4.1: Parameters of the rank-1 lattice points for our numerical results.

Therefore, we have $\|M\|_2 < \infty$ independent of n. The total error bound directly follows from this commutator bound and [89, Theorem 1]. □

4.2 Numerical results for the higher-order exponential splitting

We demonstrate our method by showing some numerical results in this section. We construct rank-1 lattices by using the component-by-component (CBC) construction [14, 68]. The code for producing the rank-1 lattice is available online [69], `fastrank1expt.m`. With the script, we choose n being a power of 2 and generate the vector $\boldsymbol{z}$ which is optimized for integration in (unweighted) Korobov space with first order mixed derivatives, i.e., $\alpha = 1$. In Table 4.2 we display the generating vector $\boldsymbol{z}$ and the number of points n for the following numerical results. Using given n and $\boldsymbol{z}$, we construct the anti-aliasing set in accordance with Theorem 4.1 in the following manner: (i) first we generate all integer vector $\boldsymbol{h} \in \mathbb{Z}^d$ in a bounded region $\|\boldsymbol{h}\| \leq R$ for a well chosen R; (ii) then we sort the obtained set according to the ℓ_2 distance in ascending order; (iii) we calculate the value $m_{\boldsymbol{h}} := \boldsymbol{h} \cdot \boldsymbol{z} \bmod n$ in the sorted order and store $\boldsymbol{h}$ in $\mathcal{A}(\boldsymbol{z}, n)$ if the value $m_{\boldsymbol{h}}$ has not appeared before. We repeat this step (iii) until we have the full cardinality $|\mathcal{A}(\boldsymbol{z}, n)| = n$.

4.2.1 Convergence with respect to time step size

We consider a common numerical setting as it is considered in [26, 33, 89] where Fourier pseudospectral methods are used. We calculate the error with different value of time steps against a reference solution. For the initial condition $g(\boldsymbol{x})$,

we choose the *Gaussian wave packet* given by:

$$g(\boldsymbol{x}) := \left(\frac{2}{\pi\gamma}\right)^{d/4} \exp\left(-\frac{\sum_{j=1}^{d}(2\pi x_j - \pi)^2}{\gamma}\right)\frac{1}{c},$$

where the constant c is a normalizing constant to make $\|g\|_{L_2} = 1$. For the potential function v, we consider a *smooth potential* function

$$v_1(\boldsymbol{x}) = \prod_{j=1}^{d}(1 - \cos(2\pi x_j)),$$

and a *harmonic potential* function

$$v_2(\boldsymbol{x}) = \frac{1}{2}\sum_{j=1}^{d}(2\pi x_j - \pi)^2.$$

Our aim is to show the temporal discretization error $\|u_a(\boldsymbol{x}, t) - u_a^m(\boldsymbol{x})\|_{L_2}$ at fixed time $t = m\,\Delta t = 1$, for that sake we calculate a reference solution $u_a^M(\boldsymbol{x})$ with the finest time step size $\Delta t = 1/M = 1/10000$, as an approximation of $u_a(\boldsymbol{x}, t)$. We then vary the time step size $\Delta t = 1/m = 1/5, \ldots, 1/1000$ and calculate $u_a^m(\boldsymbol{x})$ to see the convergence plot of $\|u_a^M(\boldsymbol{x}) - u_a^m(\boldsymbol{x})\|_{L_2}$.

4.2.2 Sixth-order splitting

We recall that the higher-order exponential splitting is written as

$$y_{k+1} = \mathrm{e}^{b_1 B\,\Delta t}\,\mathrm{e}^{a_1 A\,\Delta t}\,\cdots\mathrm{e}^{b_s B\,\Delta t}\mathrm{e}^{a_s A\,\Delta t}\,y_k.$$

For the sixth-order method, we employ the coefficients a_j and b_j from [35] denoted as "s9odr6a" therein. We exhibit the coefficients in Table 4.2. We plot the results for dimension 2 to 8 in Fig 4.1. The potential v_2 is not smooth enough on the boundary of $[0, 1]^d$ so it does not satisfy the required condition in the strict sense. The initial condition g and the potential v_1 meet all the required conditions. The expected sixth-order convergence is consistent in every plot. When the error reaches the machine precision, the plot becomes flat. For the 2-dimensional case with the potential v_2, we see the convergence happening when the time step size is very small. This can be explained by a phenomenon, called instability of exponential splitting; this is caused by negative coefficients of the exponential splitting a_j and b_j, and is discussed in e.g., [8]. Especially in [8], commutator-free quasi-Magnus exponential integrators are proposed to avoid the issue, however, this is out of the scope of the present chapter. The instability issue does not happen in a higher-dimensional settings.

	a_j		b_j
$j = 1, 9$	0.392161444007314	$j = 1, 10$	0.196080722003657
$j = 2, 8$	0.332599136789359	$j = 2, 9$	0.362380290398337
$j = 3, 7$	-0.706246172557639	$j = 3, 8$	-0.186823517884140
$j = 4, 6$	0.0822135962935508	$j = 4, 7$	-0.312016288132044
$j = 5$	0.798543990934830	$j = 5, 6$	0.440378793614190
$j = 10$	0		

Table 4.2: Coefficients for the sixth-order method, calculated based on [35].

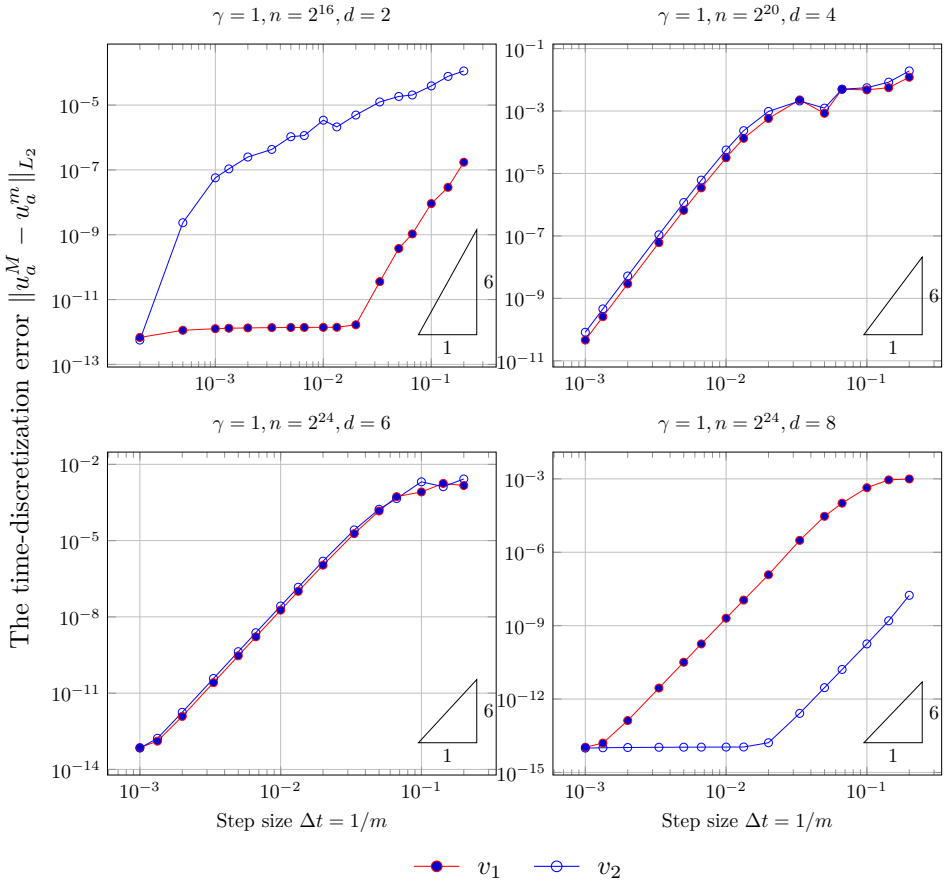

Figure 4.1: The time-discretization error with the sixth-order method.

	a_j		b_j
$j = 1, 17$	0.130202483088890	$j = 1, 18$	0.0651012415444450
$j = 2, 16$	0.561162981775108	$j = 2, 17$	0.345682732431999
$j = 3, 15$	-0.389474962644847	$j = 3, 16$	0.0858440095651306
$j = 4, 14$	0.158841906555156	$j = 4, 15$	-0.115316528044846
$j = 5, 13$	-0.395903894133238	$j = 5, 14$	-0.118530993789041
$j = 6, 12$	0.184539640978316	$j = 6, 13$	-0.105682126577461
$j = 7, 11$	0.258374387686322	$j = 7, 12$	0.221457014332319
$j = 8, 10$	0.295011723609310	$j = 8, 11$	0.276693055647816
$j = 9$	-0.605508533830035	$j = 9, 10$	-0.155248405110362
$j = 18$	0		

Table 4.3: Coefficients for the eighth-order method, calculated based on [35].

4.2.3 Eighth-order splitting

For the eighth-order method, we employ the coefficients again from [35] denoted as "s17odr8a". The coefficients are shown in Table 4.3. The results are shown in Fig 4.2 and we again see that the convergence rate is consistently eighth order in each plot. Most of the plot seems to be similar to Fig 4.1 but with faster convergence, therefore they reach to the machine precision more quickly.

4.3 Chapter overview

We proposed a numerical method to solve the TDSE. With our method using the time step size Δt, the temporal discretization error converges like $\mathcal{O}((\Delta t)^p)$ given that the potential function is in Korobov space of smoothness greater than $2p+1/2$. The numerical results (which are performed from 2 up to 8 dimensions) confirmed the theory and the rate of error convergence is consistent. By using rank-1 lattices, calculations of the time stepping operator and multiplications are efficiently done by only using one-dimensional FFTs.

Pseudospectral methods are widely used techniques for solving partial differential equations. It is a common choice to use regular grids, but the number of nodes increases exponential with d. We have shown an alternative, rank-1 lattice pseudospectral methods where the number of points can be chosen freely by the user. In combination with higher-order splitting methods, the proposed method solves the TDSE with higher-order convergence in time.

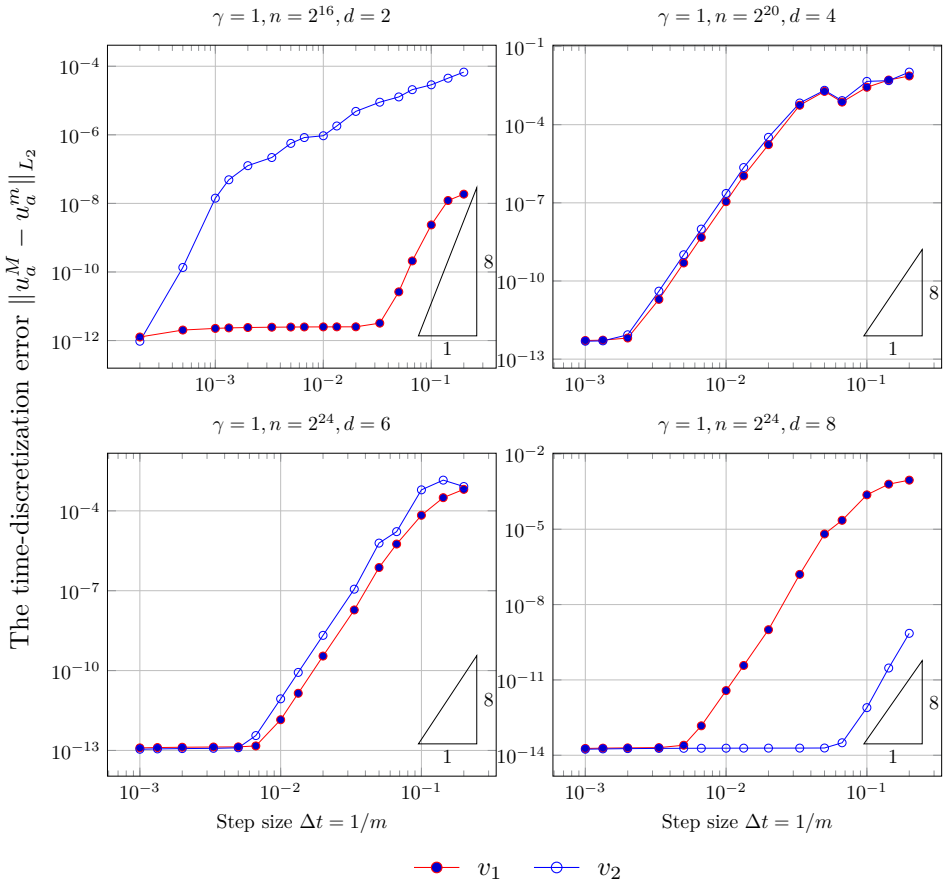

Figure 4.2: The time-discretization error with the eighth-order method.

Chapter 5

Higher-order operator splitting and rank-1 lattices for the time-dependent nonlinear Schrödinger equation

The considered problem in Chapters 3 and 4 is linear in terms of the solution; the equation (3.1) can be written as

$$\frac{\partial u}{\partial t}(\boldsymbol{x},t) = \mathcal{D}(\boldsymbol{x},t)\, u(\boldsymbol{x},t)$$

where the differential operator $\mathcal{D}(\boldsymbol{x},t)$ is independent of the solution $u(\boldsymbol{x},t)$. We have seen that lattice points can be used instead of the regular grid when Fourier pseudospectral methods are combined with exponential operator splitting. Fourier pseudospectral methods combined with exponential operator splitting are also used for solving nonlinear PDEs. Motivated by [38, 90, 91], we consider the nonlinear time-dependent Schrödinger equation in this chapter,

$$\begin{aligned} \mathrm{i}\,\frac{\partial u}{\partial t}(\boldsymbol{x},t) &= -\frac{1}{2}\,\nabla^2 u(\boldsymbol{x},t) + v(\boldsymbol{x})\, u(\boldsymbol{x},t) + \theta |u(\boldsymbol{x},t)|^2\, u(\boldsymbol{x},t), \qquad (5.1)\\ u(\boldsymbol{x},0) &= g(\boldsymbol{x}), \end{aligned}$$

where θ is a positive constant, ∇^2 is the Laplacian, $g(\boldsymbol{x})$ is the initial condition and $v(\boldsymbol{x})$ is the potential. Again, $\boldsymbol{x} \in \mathbb{T}^d$ is the spatial variable and t is time. As before, we construct an anti-aliasing set $\mathcal{A}(\boldsymbol{z}, n)$ according to a given rank-1 lattice $\Lambda(\boldsymbol{z}, n)$.

5.1 The higher-order splitting method for the nonlinear problem

We focus on the case $d = 3$ which is relevant for physical applications, i.e., Bose–Einstein condensates. By using the same arguments as in Section 3.2.4, we can discretize the space and obtain the following dynamics of the Fourier pseudospectral method:

$$\mathrm{i}\,\widehat{\boldsymbol{u}}_t' = \frac{1}{2} D_n\, \widehat{\boldsymbol{u}}_t + W_n(u_a(\cdot, t))\, \widehat{\boldsymbol{u}}_t, \tag{5.2}$$

where the initial condition $\widehat{\boldsymbol{u}}_0 = \widehat{\boldsymbol{g}}_a := (\widehat{g}_a(\boldsymbol{h}^{(0)}), \ldots, \widehat{g}_a(\boldsymbol{h}^{(n-1)}))$ with the same notation of (3.10),

$$D_n := \operatorname{diag}\left((4\pi^2 \|\boldsymbol{h}_\xi\|_2^2)_{\xi=0,\ldots,n-1}\right), \tag{5.3}$$

the operator $W_n(u_a(\cdot, t)) := F_n V_n(u_a(\cdot, t)) F_n^{-1}$ with

$$V_n(u_a(\cdot, t)) := \operatorname{diag}\left(\left(v(\boldsymbol{p}_k) + \theta |u_a(\boldsymbol{p}_k, t)|^2\right)_{k=0,\ldots,n-1}\right), \tag{5.4}$$

and F_n is the unitary Fourier matrix.

Following [91, Equation (16)] we perform higher-order exponential splitting with the pseudospectral method on lattice points for the nonlinear TDSE:

$$\begin{aligned}
\widehat{\boldsymbol{u}}_a^{k+1} &= \mathrm{e}^{b_s \Delta t W_n(\boldsymbol{U}_a^{k,s})} \boldsymbol{U}_a^{k,s}, \\
\boldsymbol{U}_a^{k,1} &= \mathrm{e}^{a_1 \Delta t D_n} \widehat{\boldsymbol{u}}_a^k, \\
\boldsymbol{U}_a^{k,j} &= \mathrm{e}^{a_j \Delta t D_n} \mathrm{e}^{b_{j-1} \Delta t W_n(\boldsymbol{U}_a^{k,j-1})} \boldsymbol{U}_a^{k,j-1}, \quad 2 \le j \le s.
\end{aligned} \tag{5.5}$$

5.2 Numerical results

In this section, we perform our method and show numerical results. We apply the sixth and eighth order operator splitting. The coefficients a_j and

b_j are again calculated from [35], the same ones from the previous chapter; see Table 4.2 and 4.3. The parameters for the rank-1 lattice are $n = 2^{16}$ and $\boldsymbol{z}^\top = (1, 19463, 8279)$. We construct the anti-aliasing set with the full cardinality n by minimizing the ℓ_2 distance, such that $\|\boldsymbol{h}_\xi\|_2 = \min_{\boldsymbol{h}' \in A(\boldsymbol{z},n,\xi)} \|\boldsymbol{h}'\|_2$ with $A(\boldsymbol{z}, n, \xi) := \left\{\boldsymbol{h} \in \mathbb{Z}^d : \boldsymbol{h} \cdot \boldsymbol{z} \equiv \xi \pmod{n}\right\}$. We choose the same initial condition and potential functions as used in Chapter 4: the Gaussian initial condition

$$g(\boldsymbol{x}) := \left(\frac{2}{\pi}\right)^{d/4} \exp\left(-\sum_{j=1}^{d} (2\pi x_j - \pi)^2\right) \frac{1}{c},$$

here the constant c is a normalizing constant to make $\|g\|_{L_2} = 1$; and we choose two potential functions, a smooth potential function

$$v_1(\boldsymbol{x}) = \prod_{j=1}^{d} (1 - \cos(2\pi x_j)),$$

and a harmonic potential function

$$v_2(\boldsymbol{x}) = \frac{1}{2} \sum_{j=1}^{d} (2\pi x_j - \pi)^2.$$

We calculate the reference solution u_a^M with the finest time step size $\Delta t = 1/M = 1/10000$. With different time step sizes Δt, we calculate the time discretization error $\|u_a^M - u_a^m\|_2$. We plot the error of the sixth order method in Figure 5.1 and the eighth order method in Figure 5.2. As expected, we see sixth order convergence in Figure 5.1 and eighth order convergence in Figure 5.2 respectively.

5.3 Discussion for theoretical error bounds

We are going to follow the analysis done by Thalhammer [90]. The only difference from the method therein is that we discretize the space by a rank-1 lattice whereas a regular grid is used in [90]. To apply the same analysis as [90], first we introduce their notation.

Let $\boldsymbol{M} = (M_1, \ldots, M_d) \in \mathbb{N}^d$ be a multiindex comprising of even integers, the index set of the regular grid $\mathcal{M}_{\boldsymbol{M}} := \prod_{j=1}^{d} \{-M_j/2, \ldots, M_j/2 - 1\}$, and the magnitude of eigenvalues of the Laplacian are denoted by $\lambda_{\boldsymbol{h}} = 4\pi^2 \|\boldsymbol{h}\|_2^2$ for the basis function $\phi_{\boldsymbol{h}}(\boldsymbol{x}) = \exp(2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{x})$ with $\boldsymbol{h} \in \mathbb{Z}^d$. In the analysis, one of the

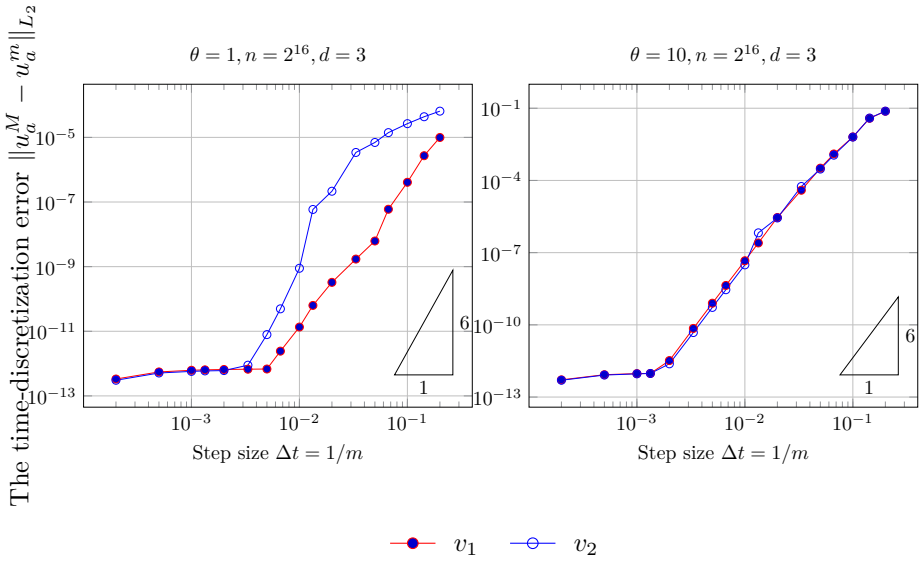

Figure 5.1: The time-discretization error with the sixth-order method for the nonlinear TDSE.

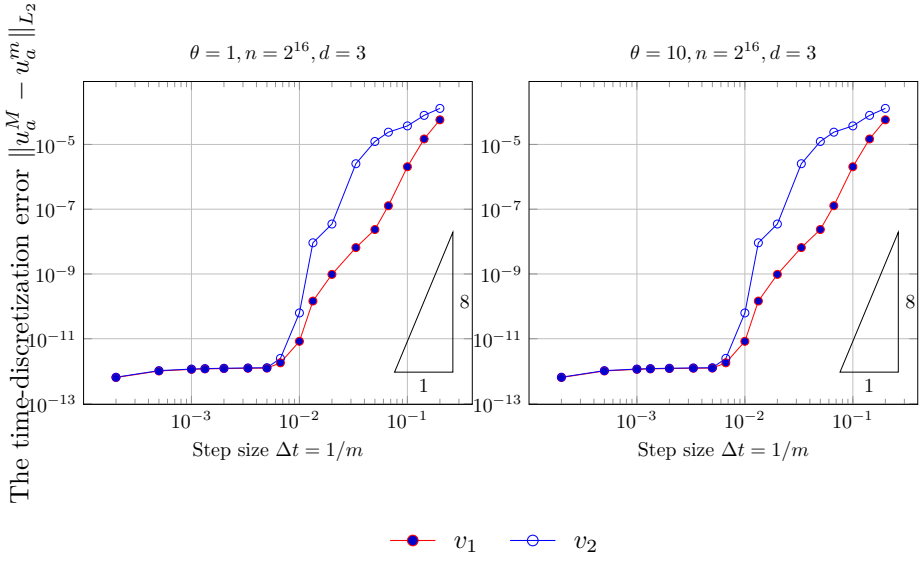

Figure 5.2: The time-discretization error with the eighth-order method for the nonlinear TDSE.

important quantities is

$$\frac{\lambda_{\max}}{\lambda_{\min}} := \frac{\max_{\boldsymbol{h}\in\mathcal{M}_{\boldsymbol{M}}} \lambda_{\boldsymbol{h}}}{\min_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{M}_{\boldsymbol{M}}} \lambda_{\boldsymbol{h}}}.$$

To obtain higher-order convergence in time, this quantity needs to be bounded by a constant independent of the number of grid points n.

In [90, p.3237 lines 13-14], it is erroneously stated that

$$\lambda_{\max} = \sum_{j=1}^{d} \pi^2 M_j^2 \leq \lambda_{\boldsymbol{h}} \quad \forall \boldsymbol{h} \in \mathbb{Z}^d \setminus \mathcal{M}_{\boldsymbol{M}}.$$

However, considering $\boldsymbol{h} = (0,\ldots,0,M_{\min}/2,0,\ldots,0)$, where $M_{\min}$ is the smallest element in $\boldsymbol{M}$, this bound needs to be corrected to

$$\lambda_{\max} \leq d\lambda_{\boldsymbol{h}} \quad \forall \boldsymbol{h} \in \mathbb{Z}^d \setminus \mathcal{M}_{\boldsymbol{M}},$$

with the additional constant d. This does not affect the main results of their analysis, only the constants of the derived error expression. To apply our method using lattice points, we also need the similar bound,

$$\max_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)} \lambda_{\boldsymbol{h}} \leq C_d \min_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{z},n)} \lambda_{\boldsymbol{h}},$$

with some constant C_d which does not depend on n but depends on d. To describe this problem we give an example in Figure 5.3 where blue points are our choice of the anti-aliasing set and the red point is the shortest index outside the anti-aliasing set. We remark that there can be multiple $\boldsymbol{h} \in \mathbb{Z}^d \setminus \mathcal{A}(\boldsymbol{z},n)$ with $\lambda_{\boldsymbol{h}} = \lambda_{\min}$, but this is not the case for the presented example of Figure 5.3. We wish to have a bound on the ratio of the longest distance in the anti-aliasing set divided by the shortest distance outside the set.

For the theoretical analysis, we assume the following conjecture holds where numerical justification is going to be given later in this section.

Conjecture 5.1. *For any d and n there exists rank-1 lattice points $\Lambda(\boldsymbol{z},n)$ satisfying*

$$\max_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)} \lambda_{\boldsymbol{h}} \leq C_d \min_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{z},n)} \lambda_{\boldsymbol{h}},$$

where the anti-aliasing set $\mathcal{A}(\boldsymbol{z},n) = \{\boldsymbol{h}_\xi \in \mathbb{Z}^d : \boldsymbol{h}_\xi \cdot \boldsymbol{z} \equiv \xi \pmod n$ for $\xi = 0,\ldots,n-1\}$, with full cardinality, is chosen such that it has minimal ℓ_2 norm, i.e.,

$$\|\boldsymbol{h}_\xi\|_2 = \min_{\boldsymbol{h}'\in A(\boldsymbol{z},n,\xi)} \|\boldsymbol{h}'\|_2,$$

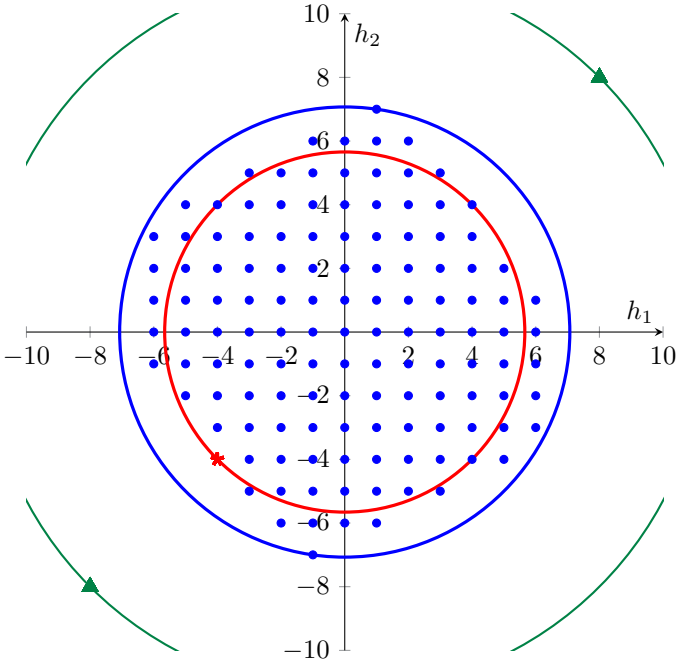

Figure 5.3: An example of an anti-aliasing set (blue points), the shortest index outside the set (red star $(-4,-4)^\top$), and the shortest dual lattice points (green triangles $(-8,-8)^\top$ and $(8,8)^\top$) where $n = 128$ and $\boldsymbol{z}^\top = (1, 47)$.

with

$$A(\boldsymbol{z}, n, \xi) := \big\{\boldsymbol{h} \in \mathbb{Z}^d : \boldsymbol{h} \cdot \boldsymbol{z} \equiv \xi \pmod{n}\big\},$$

and the constant C_d only depends on d.

We note that there is a related quantity called *the spectral index*,

$$\sigma(\boldsymbol{z}, n) := \frac{1}{\min_{\boldsymbol{0}\neq\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{z},n)} \|\boldsymbol{h}\|_2},$$

which is intensely studied; see [30, 48, 71]. This quantity is defined for

$$\boldsymbol{h} \in (\Lambda^\perp(\boldsymbol{z}, n) \setminus \{\boldsymbol{0}\}) \subseteq \mathbb{Z}^d \setminus \mathcal{A}(\boldsymbol{z}, n),$$

since we always have $\boldsymbol{0} \in \mathcal{A}(\boldsymbol{z}, n)$, hence $\mathcal{A}(\boldsymbol{z}, n) \cap (\Lambda^\perp(\boldsymbol{z}, n) \setminus \{\boldsymbol{0}\}) = \emptyset$. Therefore we have

$$\sigma(\boldsymbol{z}, n) = \frac{1}{\min_{\boldsymbol{0}\neq\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{z},n)} \|\boldsymbol{h}\|_2} \leq \frac{1}{\min_{\boldsymbol{0}\neq\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{z},n)} \|\boldsymbol{h}\|_2}.$$

We can also show that the upper bound on $\min_{\mathbf{0}\neq\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{z},n)}\|\boldsymbol{h}\|_2$ can be bounded by using the spectral index.

Lemma 5.2. *For any rank-1 lattice point set $\Lambda(\boldsymbol{z},n)$ and corresponding anti-aliasing set $\mathcal{A}(\boldsymbol{z},n)=\{\boldsymbol{h}_\xi\in\mathbb{Z}^d:\boldsymbol{h}_\xi\cdot\boldsymbol{z}\equiv\xi \pmod n$ for $\xi=0,\dots,n-1\}$, with full cardinality, which is chosen such that it has minimal ℓ_2 norm, i.e.,*

$$\|\boldsymbol{h}_\xi\|_2=\min_{\boldsymbol{h}'\in A(\boldsymbol{z},n,\xi)}\|\boldsymbol{h}'\|_2,$$

with

$$A(\boldsymbol{z},n,\xi):=\big\{\boldsymbol{h}\in\mathbb{Z}^d:\boldsymbol{h}\cdot\boldsymbol{z}\equiv\xi \pmod n\big\}.$$

Then the following bound holds:

$$\min_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{z},n)}\|\boldsymbol{h}\|_2\geq\frac{\min_{\mathbf{0}\neq\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{z},n)}\|\boldsymbol{h}\|_2}{2}.$$

Proof. Let $\boldsymbol{\ell}_{\min}:=\arg\min_{\mathbf{0}\neq\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{z},n)}\|\boldsymbol{h}\|_2$. We prove that

$$\forall\boldsymbol{h}\in\mathbb{Z}^d,\|\boldsymbol{h}\|_2<\frac{\|\boldsymbol{\ell}_{\min}\|_2}{2}\implies\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n),$$

by contradiction. Assume that there exists $\boldsymbol{h}\notin\mathcal{A}(\boldsymbol{z},n)$ which satisfies $\|\boldsymbol{h}\|_2<\|\boldsymbol{\ell}_{\min}\|_2/2$. Then due to the construction of $\mathcal{A}(\boldsymbol{z},n)$,

$$\exists\boldsymbol{h}'\in\mathcal{A}(\boldsymbol{z},n),\ \boldsymbol{h}'-\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{z},n)\setminus\{\mathbf{0}\},\text{ and }\|\boldsymbol{h}'\|_2\leq\|\boldsymbol{h}\|_2.$$

Then, by letting $\boldsymbol{\ell}:=\boldsymbol{h}'-\boldsymbol{h}$, we have

$$\begin{aligned}\|\boldsymbol{h}'\|_2^2=\|\boldsymbol{h}+\boldsymbol{\ell}\|_2^2&=\|\boldsymbol{h}\|_2^2+\|\boldsymbol{\ell}\|_2^2+2\boldsymbol{\ell}\cdot\boldsymbol{h}\\&\geq\|\boldsymbol{h}\|_2^2+\|\boldsymbol{\ell}\|_2^2-2\|\boldsymbol{\ell}\|_2\|\boldsymbol{h}\|_2\\&>\|\boldsymbol{h}\|_2^2+\|\boldsymbol{\ell}\|_2^2-\|\boldsymbol{\ell}\|_2\|\boldsymbol{\ell}_{\min}\|_2\\&\geq\|\boldsymbol{h}\|_2^2,\end{aligned}$$

therefore $\|\boldsymbol{h}'\|_2>\|\boldsymbol{h}\|_2$ but this contradicts to $\|\boldsymbol{h}'\|_2\leq\|\boldsymbol{h}\|_2$. Therefore the claim of the lemma is proved. □

The spectral index itself has lower and upper bounds for a good lattice point set $\Lambda(\boldsymbol{z},n)$, shown by [71, Proposition 3, 4] using Minkowski's fundamental theorem,

$$\frac{\sqrt{\pi}}{2}\left(\Gamma\left(\frac{d}{2}+1\right)\right)^{-1/d}N^{-1/d}\leq\sigma(\boldsymbol{z},n)\leq C_d'\,N^{-1/d},$$

where the constant C_d' only depends on d.

With the spectral index we can bound

$$\frac{\max_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)}\lambda_{\boldsymbol{h}}}{\min_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{z},n)}\lambda_{\boldsymbol{h}}} \le 4\sigma^2(\boldsymbol{z},n)\max_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)}\|\boldsymbol{h}\|_2^2.$$

However, due to the presence of $\max_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)}\|\boldsymbol{h}\|_2^2$, the bound on C_d in Conjecture 5.1 is still not known.

As we mentioned earlier, we numerically test this conjecture by calculating the quantity

$$q(\boldsymbol{z},n) := \frac{\max_{\boldsymbol{h}\in\mathcal{A}(\boldsymbol{z},n)}\lambda_{\boldsymbol{h}}}{\min_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\mathcal{A}(\boldsymbol{z},n)}\lambda_{\boldsymbol{h}}}$$

for fixed $\Lambda(\boldsymbol{z},n)$ which are generated by using `fastrank1expt.m` in the same way as in Section 3.3, optimized for integration in the Korobov space with smoothness $\alpha = 1$. We remark that this quantity do not depend on the anti-aliasing set $\mathcal{A}(\boldsymbol{z},n)$ as far as we construct the anti-aliasing set by minimizing ℓ_2 distance. We display the parameters $\boldsymbol{z}$ and n in Table 5.3. The calculated quantity q is plotted in Fig 5.4.

In Figure 5.4, we see that for relatively bigger n the quantity q is not affected by the number of points n. The bad behavior of q for small number of n in high-dimension d can be explained from the quality of rank-1 lattice point sets; with Table 5.3, it can be seen that the bigger value of q is caused by $\boldsymbol{z}$ that have the same value in different components. This is simply because of the difficulty of choosing good lattice rule for small number of points in high-dimension. For such $\boldsymbol{z}$, the spectral index becomes $1/\sqrt{2}$.

Given that Conjecture 5.1 holds, we have the following error bound.

Theorem 5.3. *Assume that Conjecture 5.1 holds. If the potential v and the analytical solution of (5.1) are in $E_\alpha(\mathbb{T}^d)$ for $\alpha > p + 1/2$. Then by using the p-th order splitting scheme, we have*

$$\|u(\cdot,t) - \widehat{u}_a^k(\cdot)\|_{L_2} \le C_1\|g(\cdot) - \mathcal{I}_n g(\cdot)\|_{L_2} + C_2\tau^p + C_3\|I - \mathcal{I}_n\|_{E_\alpha\to L_2},$$

where the constants C_1, C_2, C_3 does not depend on n or τ, I is the identity operator $L_2 \mapsto L_2$, and $\mathcal{I}_n$ is the interpolation operator on a rank-1 lattice $\Lambda(\boldsymbol{z},n)$, see (3.27).

Proof. We refer to [90, Theorem 3]. The term $C_3\|I - \mathcal{I}_n\|_{E_\alpha\to L_2}$ is equivalent to their error estimate denoted as $CM_0^{\beta-1}$ for the regular grid case. □

n	$\boldsymbol{z}^\top$
2^4	$(1, 7, 5, 7, 7, 7, 7, 7, 7)$
2^5	$(1, 7, 11, 9, 7, 7, 7, 7, 7)$
2^6	$(1, 19, 29, 31, 27, 31, 31, 31, 31)$
2^7	$(1, 47, 19, 3, 37, 1, 1, 1, 1)$
2^8	$(1, 75, 23, 57, 31, 15, 75, 75, 75)$
2^9	$(1, 149, 113, 207, 219, 77, 75, 1, 1)$
2^{10}	$(1, 283, 157, 387, 173, 491, 31, 283, 283)$
2^{11}	$(1, 791, 213, 89, 927, 557, 737, 557, 557)$
2^{12}	$(1, 1557, 1741, 1031, 363, 1705, 1349, 1667, 1705)$
2^{13}	$(1, 2433, 1715, 131, 3829, 255, 3317, 893, 3973)$
2^{14}	$(1, 6915, 3959, 7595, 6297, 1183, 2821, 8191, 1683)$
2^{15}	$(1, 12031, 7247, 2267, 14787, 7663, 4679, 11655, 4323)$
2^{16}	$(1, 19463, 8279, 14631, 10281, 23539, 4021, 11895, 9047)$
2^{17}	$(1, 50687, 7637, 38773, 16093, 13037, 27383, 30479, 43445)$
2^{18}	$(1, 100135, 28235, 39865, 25937, 127279, 60735, 85271, 111087)$
2^{19}	$(1, 154805, 242105, 171449, 24087, 51345, 240415, 179881, 166955)$
2^{20}	$(1, 443165, 95693, 34519, 235147, 317205, 59465, 515889, 497465)$

Table 5.1: Parameters of the numerical tests for the Conjecture 5.1. For $d \leq 8$, $\boldsymbol{z}$ is chosen to be the first d components, e.g., for $d = 4, n = 2^{13}$, we choose $\boldsymbol{z}^\top = (1, 2433, 1715, 131)$.

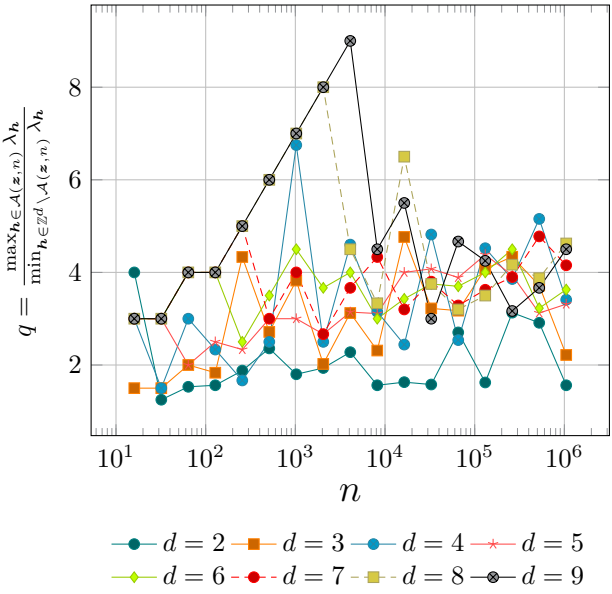

Figure 5.4: The quantity q for different n and d, the generating vector $\boldsymbol{z}$ is displayed in Table 5.3.

5.4 Chapter overview

We applied the pseudospectral Fourier method on lattices combined with higher-order exponential operator splitting to the nonlinear time-dependent Schrödinger equation. Numerical results show the higher-order convergence in the time discretization error. Conditional on Conjecture 5.1, we can prove the error bound of the whole scheme for the problem (Theorem 5.3). Depending on the potential function and the initial condition, we can choose suitable lattice points $\Lambda(\boldsymbol{z}, n)$ which approximate the solution better than the regular grid. Another merit of using lattice points for this problem is that one can choose the number of points more flexibly.

Chapter 6

Lattice rules for integration over $\mathbb{R}^d$

6.1 Introduction

In this chapter we study how to apply lattice rules in $\mathbb{R}^d$ with the aim of obtaining higher order convergence, i.e., to obtain error bounds which show that the error can be bounded by $O(n^{-\alpha})$ for $\alpha > 1$, with n the number of integrand evaluations. Typically the smoothness of the integrand is related to α by the number of derivatives that exist (in all directions) and are, e.g., L_2 integrable (see Chapter 2 and further sections for details). More specifically, we are interested in using lattice rules to approximate the integral by first truncating to a box $[\boldsymbol{a}, \boldsymbol{b}] := \prod_{j=1}^{d}[a_j, b_j]$, with all $a_j \le b_j$, and then mapping the lattice rule to the box:

$$\int_{\mathbb{R}^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \approx \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \approx \frac{\prod_{j=1}^{d}(b_j - a_j)}{n} \sum_{i=1}^{n} f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}).$$

The quadrature nodes $\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}$ are obtained from a rank-1 lattice point set $\Lambda(\boldsymbol{z}, n)$ which we defined in Section 2.2 by using two parameters, the number of points n and the generating vector $\boldsymbol{z} \in \mathbb{Z}_n^d$,

$$\Lambda(\boldsymbol{z}, n) := \left\{ \boldsymbol{p}_i := \frac{i\boldsymbol{z}}{n} \bmod 1 \right\}_{i=1}^{n} \subset [0,1]^d,$$

and the nodes are mapped from the unit cube $[0,1]^d$ to the box $[\boldsymbol{a},\boldsymbol{b}]$ by

$$p_{i,j}^{[\boldsymbol{a},\boldsymbol{b}]} := (b_j - a_j)\, p_{i,j} + a_j$$

where $p_{i,j}^{[\boldsymbol{a},\boldsymbol{b}]}$ is the j-th component of $\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}$, and likewise, $p_{i,j}$ is the j-th component of $\boldsymbol{p}_i$. By the triangle inequality we have

$$\left| \int_{\mathbb{R}^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) \right|$$

$$\le \left| \int_{\mathbb{R}^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \right| + \left| \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) \right|,$$

i.e., the total error is bouned by the sum of the truncation error and the integration error. We will select the box $[\boldsymbol{a},\boldsymbol{b}]$ and the number of lattice points n to balance both errors. Our major concern however will be how to analyse the error of integrating the non-periodic integrand on the box $[\boldsymbol{a},\boldsymbol{b}]$ by a lattice rule.

Different approaches exist in the literature to apply quasi-Monte Carlo methods designed for integration over the unit cube $[0,1]^d$ to integration over $\mathbb{R}^d$, e.g., [22, 32, 47, 56–58]. In [57] analytic functions with exponential decay in the physical and in the Fourier space are studied and exponential convergence is obtained using a regular grid with appropriate scaling in each direction. In [22] integration with respect to the Gaussian measure is considered where integrand functions are in some unanchored Sobolev spaces. Therein, higher-order convergence is achieved using linearly transformed higher-order digital nets on a box. Similarly, in [56] higher-order convergence is achieved with interlaced polynomial lattice rules where a different type of function spaces, anchored Sobolev spaces are considered. In [58] randomly shifted lattice rules are studied in combination with an inverse mapping of the considered cumulative density function, and first order convergence is achieved.

In this chapter, we consider linearly transformed lattice rules on a box of increasing size, without random shifting or inverse mapping, and we derive explicit conditions to obtain higher-order convergence.

Apart from the problem of adjusting from the unit cube $[0,1]^d$ to $\mathbb{R}^d$, we also need to handle that the integrand function f on the box $[\boldsymbol{a},\boldsymbol{b}]$ is non-periodic. The analysis of lattice rules for numerical integration is typically done in a periodic setting with absolutely converging Fourier series expansions. In this setting the smoothness α as previously defined, means that the function values and its derivative values (up to a certain order) will match on the boundaries of

the unit cube. To apply lattice rules in the non-periodic setting on the unit cube $[0,1]^d$ one can apply so-called "periodizing transformations" [46, 78, 80]. This method applies a variable substitution to the integral, dimension-by-dimension, to end up with a periodic integrand function while preserving the value of the integral. The problem with this strategy is that this "blows up the norm", as noted by many researchers [24, 31, 40, 94]. A second approach is known as the "method of Bernoulli polynomials" [40, 80, 94]. In this method we add Bernoulli polynomials to the original integrand in such a way that function values (and derivative values) match at the boundaries of the unit cube, making the integrand periodic, while preserving the value of the integral, since those Bernoulli polynomials integrate to zero. As pointed out by many researchers [7, 40, 80, 94], this method does not blow up the norm, but the amount of terms that need to be added increases exponentially (by correcting the value of the original integrand on all possible boundaries of the hypercube, and, by the need to analytically calculate the derivatives of the original integrand, which might also grow exponentially by, e.g., the need to use the chain rule). This is probably the reason why this last method appears to be not used as much. Both the periodizing transformation and the method of Bernoulli polynomials can transform a non-periodic function of smoothness α into a periodic function with the same smoothness, hence achieving $O(n^{-\alpha})$. Since the periodizing transformation might blow up the norm and the method of Bernoulli polynomials could increase the cost of evaluating the integrand, the recent literature on lattice rules mostly applies two other techniques: "random shifting" and the "tent-transform", which can also be combined [16, 20, 24, 31]. This method is much easier and does not have the previously mentioned defects, but the convergence is limited to $O(n^{-1})$ and $O(n^{-2})$.

Here we take yet another approach. First we analyse what the effect is of integrating a non-periodic function with a lattice rule in terms of a measure of how non-periodic the function is. We can measure the non-periodicity of a function by quantifying how much the function and its derivatives (up to a certain order) differ at the boundaries of the domain. We derive an explicit error bound taking this difference into account. Since the method of Bernoulli polynomials actually corrects this non-periodicity, we are in fact applying this method in the theoretical analysis. The advantage of this approach is that we never have to actually construct the modified integrand. Of course, since, this defect is constant, such an approach would not work on the unit cube, or any fixed box. However, in the setting of integration over $\mathbb{R}^d$ we can use the fact that as we increase the boundaries of our box $[\boldsymbol{a}, \boldsymbol{b}]$ to cover more and more of $\mathbb{R}^d$, then the function will differ less and less on the boundaries of our box. If we demand a certain decay of the function values and its derivatives up to a certain order, then our integrand function becomes more and more periodic as our box increases. It turns out that in our setting, where the integrand

functions are from a certain reproducing kernel Hilbert space which we call an unanchored Sobolev space, the modified integrand which can be obtained by applying the Bernoulli polynomial method, is actually the orthogonal projection of the original function from the unanchored Sobolev space to an associated Korobov space. Using the orthogonal projection we can thus decompose the integrand into a periodic part and a non-periodic part where those two parts are orthogonal in the unanchored Sobolev space. Hence, for integration over $\mathbb{R}^d$, by using this orthogonal projection we can divide the problem into three parts: truncation error, integration error of the periodic part, and approximation error of the non-periodic part:

$$\begin{aligned}
&\left|\int_{\mathbb{R}^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \frac{\prod_{j=1}^d (b_j-a_j)}{n}\sum_{i=1}^n f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]})\right| \\
&\quad\le \left|\int_{\mathbb{R}^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| \\
&\qquad + \left|\int_{[\boldsymbol{a},\boldsymbol{b}]} F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \frac{\prod_{j=1}^d (b_j-a_j)}{n}\sum_{i=1}^n F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]})\right| \\
&\qquad\quad + \left|\frac{\prod_{j=1}^d (b_j-a_j)}{n}\sum_{i=1}^n \left(F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]})\right)\right|,
\end{aligned}$$

where $F^{[\boldsymbol{a},\boldsymbol{b}]}$ is the projected part of f onto the Korobov space on the box $[\boldsymbol{a},\boldsymbol{b}]$ and we have $\int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{[\boldsymbol{a},\boldsymbol{b}]} F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$. To control the approximation error of the non-periodic part, we will ask the integrand to become more and more periodic as the box increases. The speed of becoming periodic will be carefully chosen such that the convergence of the total error is not deteriorated by the non-periodicity.

Current research on QMC often considers high-dimensional or even infinite-dimensional problems given the information of which dimension is how much important. This importance is represented by sequence of positive numbers, and such spaces are called to have weight structure, see e.g., [21, 81, 82] and [18, Section 4]. Our current method does not use weights and therefore suffers from the curse of dimensionality, but since we are integrating over $\mathbb{R}^d$ it is essential that we first see how the unweighted case goes; as in our setting there are more parameters to play with: the dimension weights (product weights) and the decay of the function. The study of weighted spaces and decay is remaining for future work.

6.2 Reproducing kernel Hilbert spaces

We consider integration for functions in unanchored Sobolev spaces which we have seen in Section 2.1. In the context of high-dimensional integration and QMC methods these are basically tensor products of one-dimensional spaces where derivatives up to a certain order α are L_2 integrable; and where in the multi-variate setting these derivatives can be taken up to order α in each dimension simultaneously. In different contexts, such spaces are often said to have "dominating mixed-smoothness". In our analysis, we connect this space to the Korobov spaces (see Section 2.1), which are similar function spaces of "periodic functions" and where it is known that lattice rules can achieve the optimal order of convergence [61, 80]. We first introduce those spaces on the unit cube $[0,1]^d$ and then on a box $[\boldsymbol{a},\boldsymbol{b}] \subset \mathbb{R}^d$.

In the classical references the function spaces are mostly Banach spaces. Here we want to make use of the connection between the Hilbert spaces L_2 and ℓ_2 with respect to Fourier coefficients by making use of the Parseval theorem to be able to express L_2 norms of derivatives as ℓ_2 norms of the Fourier coefficients of the original function. Secondarily, since we are now in a Hilbert space setting, we can later use the idea of an orthogonal projection and its complement to analyze the effect of integrating a non-periodic function by a lattice rule.

These function spaces are tensor-products of one dimensional spaces. For the one-dimensional spaces we will make use of the fact that if f is *absolutely continuous* on $[a,b]$ then f has a derivative f' almost everywhere that is Lebesgue integrable on $[a,b]$, $-\infty < a < b < \infty$, and for which

$$f(x) = f(a) + \int_a^x f'(t)\,\mathrm{d}t \qquad \text{for all } x \in [a,b],$$

and in particular $\int_a^b f'(t)\,\mathrm{d}t = f(b) - f(a)$ [75]. As a shorthand we sometimes write $\int_a^b f'(t)\,\mathrm{d}t = f(t)|_{t=a}^b = [f(t)]_{t=a}^b = f(b) - f(a)$. The one-dimensional Sobolev spaces we describe here will consist of f for which $f^{(\tau)}$ will be absolutely continuous for all $\tau = 0,\dots,\alpha-1$ and $f^{(\alpha)}$ will be $L_2([a,b])$ integrable. What remains is to choose an inner product, which we will do in the next sections. See, e.g., [64] for the standard $[0,1]^d$ domain.

6.2.1 Some properties of Bernoulli polynomials

We recall some well known properties of Bernoulli polynomials, see, e.g., [1, 63], which will be denoted by B_τ, where τ is the degree of the polynomial. The first few Bernoulli polynomials are $B_0(x) = 1$, $B_1(x) = x - 1/2$, $B_2(x) = x^2 - x + 1/6$,

$B_3(x) = x^3 - (3/2)x^2 + (1/2)x, \ldots$ We are interested in their properties on the interval $[0,1]$. We also define scaled Bernoulli polynomials for usage on arbitrary intervals $[a,b]$, with $a < b$, with equivalent properties. Those will be denoted by $B_\tau^{[a,b]}$.

The following properties are well known, for $\tau, \tau' \in \{0,1,2,\ldots\}$,

$$\int_0^1 B_\tau(x)\,\mathrm{d}x = \begin{cases} 1, & \text{if } \tau = 0, \\ 0, & \text{if } \tau \neq 0, \end{cases}$$

$$\frac{\mathrm{d}^{\tau'}}{\mathrm{d}x^{\tau'}} \frac{B_\tau(x)}{\tau!} = \frac{B_\tau^{(\tau')}(x)}{\tau!} = \begin{cases} 0, & \text{if } \tau' > \tau, \\ \dfrac{B_{\tau-\tau'}(x)}{(\tau-\tau')!}, & \text{if } \tau' \le \tau. \end{cases}$$

On the interval $[0,1]$ the Bernoulli polynomials have the Fourier expansion which we will use to define the "periodic Bernoulli polynomials" denoted by $\widetilde{B}_\tau$ as follows

$$\frac{\widetilde{B}_\tau(x)}{\tau!} := \frac{-1}{(2\pi\mathrm{i})^\tau} \sum_{0 \neq h \in \mathbb{Z}} \frac{\exp(2\pi\mathrm{i}\,hx)}{h^\tau} \quad \begin{cases} \text{for } \tau = 2,3,\ldots \text{ and } (x \bmod 1) \in [0,1], \text{ or,} \\ \text{for } \tau = 1 \text{ and } (x \bmod 1) \in (0,1). \end{cases}$$

For $\tau = 2,3,\ldots$ we have $\widetilde{B}_\tau(x) = B_\tau(x)$ when $x \in [0,1]$ and for $\tau = 1$ we have $\widetilde{B}_1(x) = B_1(x)$ when $x \in (0,1)$. From here it follows that for $\tau = 1,2,3,\ldots$,

$$\int \frac{\widetilde{B}_\tau(x)}{\tau!}\,\mathrm{d}x = \frac{-1}{(2\pi\mathrm{i})^{\tau+1}} \sum_{0 \neq h \in \mathbb{Z}} \frac{\exp(2\pi\mathrm{i}\,hx)}{h^{\tau+1}} = \frac{\widetilde{B}_{\tau+1}(x)}{(\tau+1)!} \tag{6.1}$$

and, again for $\tau = 1,2,3,\ldots$,

$$\int \frac{\widetilde{B}_\tau(z-y)}{\tau!}\,\mathrm{d}y = \frac{1}{(2\pi\mathrm{i})^{\tau+1}} \sum_{0 \neq h \in \mathbb{Z}} \frac{\exp(2\pi\mathrm{i}\,h(z-y))}{h^{\tau+1}} = \frac{-\widetilde{B}_{\tau+1}(z-y)}{(\tau+1)!}.$$

We also have the following property, for $\tau = 0,1,2,\ldots$,

$$\widetilde{B}_\tau(x-y) = (-1)^\tau\, \widetilde{B}_\tau(y-x),$$

which follows immediately from the symmetry $B_\tau(x) = (-1)^\tau\, B_\tau(1-x)$ for $x \in [0,1]$ and $(1-(x-y)) \bmod 1 = (y-x) \bmod 1$.

We now define scaled Bernoulli polynomials on the interval $[a,b]$, $a < b$, making sure that the same properties as for the standard Bernoulli polynomials on the interval $[0,1]$ hold. Define, for $\tau = 0,1,2,\ldots$,

$$B_\tau^{[a,b]}(x) := (b-a)^{\tau-1}\, B_\tau((x-a)/(b-a)).$$

Note that $B_0^{[a,b]}(x) = 1/(b-a)$. It can be easily verified that

$$\int_a^b B_\tau^{[a,b]}(x)\,\mathrm{d}x = \begin{cases} 1, & \text{if } \tau = 0, \\ 0, & \text{if } \tau \neq 0, \end{cases}$$

and

$$\frac{\mathrm{d}^{\tau'}}{\mathrm{d}x^{\tau'}} \frac{B_\tau^{[a,b]}(x)}{\tau!} = \frac{B_\tau^{[a,b]\,(\tau')}(x)}{\tau!} = \begin{cases} 0, & \text{if } \tau' > \tau, \\ \dfrac{B_{\tau-\tau'}^{[a,b]}(x)}{(\tau-\tau')!}, & \text{if } \tau' \leq \tau. \end{cases}$$

For the Fourier series on the interval $[a,b]$ we fix the basis functions, for $h \in \mathbb{Z}$, to be

$$\varphi_h^{[a,b]}(x) := \frac{\exp(2\pi\mathrm{i}\,h(x-a)/(b-a))}{\sqrt{b-a}}, \tag{6.2}$$

such that they are orthogonal and normalized with respect to the standard L_2-inner product on $[a,b]$, i.e., for $h, h' \in \mathbb{Z}$, we have

$$\int_a^b \varphi_h^{[a,b]}(x)\,\overline{\varphi_{h'}^{[a,b]}(x)}\,\mathrm{d}x = \begin{cases} 1, & \text{if } h = h', \\ 0, & \text{if } h \neq h'. \end{cases}$$

We can now also define periodic Bernoulli polynomials on the interval $[a,b]$ as follows

$$\frac{\widetilde{B}_\tau^{[a,b]}(x)}{\tau!} := \frac{-(b-a)^{\tau-1/2}}{(2\pi\mathrm{i})^\tau} \sum_{0\neq h\in\mathbb{Z}} \frac{\varphi_h^{[a,b]}(x)}{h^\tau} \quad \begin{cases} \text{for } \tau = 2,3,\ldots \text{ and } x \in \mathbb{T}[a,b], \text{ or,} \\ \text{for } \tau = 1 \text{ and } x \in \mathbb{T}(a,b), \end{cases}$$

in which the notation $x \in \mathbb{T}[a,b]$ means that the value of x is identified with its value on the closed torus $[a,b]$, likewise for $x \in \mathbb{T}(a,b)$ where x is identified with its value on the open torus (a,b). Obviously for $\tau = 2,3,\ldots$ we have $\widetilde{B}_\tau^{[a,b]}(x) = B_\tau^{[a,b]}(x)$ when $x \in [a,b]$ and for $\tau = 1$ we have $\widetilde{B}_1^{[a,b]}(x) = B_1^{[a,b]}(x)$ when $x \in (a,b)$. The periodicity implies that

$$\widetilde{B}_\tau^{[a,b]}(x) = \widetilde{B}_\tau^{[a,b]}(x + k\,(b-a)) \qquad \forall x \in \mathbb{R},\ \forall k \in \mathbb{Z}.$$

It follows that for $\tau = 1,2,3,\ldots$,

$$\int \frac{\widetilde{B}_\tau^{[a,b]}(x)}{\tau!}\,\mathrm{d}x = \frac{-(b-a)^{\tau-1/2+1}}{(2\pi\mathrm{i})^{\tau+1}} \sum_{0\neq h\in\mathbb{Z}} \frac{\varphi_h^{[a,b]}(x)}{h^{\tau+1}} = \frac{\widetilde{B}_{\tau+1}^{[a,b]}(x)}{(\tau+1)!}$$

and, again for $\tau = 1, 2, 3, \ldots$,

$$\int \frac{\widetilde{B}_\tau^{[a,b]}(z-y)}{\tau!}\,\mathrm{d}y = \frac{-(b-a)^{\tau-1/2+1}}{(2\pi\mathrm{i})^{\tau+1}} \sum_{0\neq h\in\mathbb{Z}} \frac{\varphi_h^{[a,b]}(z-y)}{h^{\tau+1}} = \frac{-\widetilde{B}_{\tau+1}^{[a,b]}(z-y)}{(\tau+1)!}.$$

We also have the following property, for $\tau = 0, 1, 2, \ldots$,

$$\widetilde{B}_\tau^{[a,b]}(x-y+a) = (-1)^\tau\, \widetilde{B}_\tau^{[a,b]}(b-(x-y)-(b-a)) = (-1)^\tau\, \widetilde{B}_\tau^{[a,b]}(y-x+a),$$

which followed immediately from the symmetry $B_\tau^{[a,b]}(a+t) = (-1)^\tau\, B_\tau^{[a,b]}(b-t)$ for $t \in [0, b-a]$ and the periodicity with period $(b-a)$.

The maximal magnitude of the scaled Bernoulli polynomials on the interval $[a,b]$ can be bounded by

$$\forall x \in [a,b]: \qquad |B_\tau^{[a,b]}(x)| \le (b-a)^{\tau-1}\frac{2\tau!}{(2\pi)^\tau}\begin{cases}\zeta(\tau), & \text{if } \tau \equiv 0 \pmod 2,\\ 1, & \text{otherwise,}\end{cases} \tag{6.3}$$

for $\tau \ge 2$. For derivation of the bound we refer to [49]. Note that this means that

$$\forall x \in [a,b],\ \forall \tau \ge 1: \qquad |B_\tau^{[a,b]}(x)| \le (b-a)^{\tau-1}\,\frac{\tau!}{2}.$$

6.2.2 RKHSs on the unit cube

To recall the unanchored Sobolev space and the Korobov space from Section 2.1, and also to see the deeper relation of those two spaces, we formally define those two spaces in the following.

Definition 6.1. *For $\alpha \in \mathbb{N}$ the space $\mathcal{H}(K_\alpha^{\mathrm{Sob}})$ is an unanchored Sobolev space on the unit cube $[0,1]^d$, which is a reproducing kernel Hilbert space with inner product*

$$\langle f, g\rangle_{K_{\alpha,d}^{\mathrm{Sob}}} := \sum_{\substack{\boldsymbol{\tau}\in\{0,\ldots,\alpha\}^d\\ \mathfrak{w}:=\{j:\tau_j=\alpha\}}} \int_{[\mathbf{0}_\mathfrak{w},\mathbf{1}_\mathfrak{w}]} \left(\int_{[\mathbf{0}_{-\mathfrak{w}},\mathbf{1}_{-\mathfrak{w}}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{w}}\right) \left(\int_{[\mathbf{0}_{-\mathfrak{w}},\mathbf{1}_{-\mathfrak{w}}]} g^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{w}}\right)\mathrm{d}\boldsymbol{x}_\mathfrak{w},$$

and reproducing kernel

$$K^{\mathrm{Sob}}_{\alpha,d}(\boldsymbol{x},\boldsymbol{y}) := \prod_{j=1}^{d}\left(1+\sum_{\tau_j=1}^{\alpha}\frac{B_{\tau_j}(x_j)}{\tau_j!}\,\frac{B_{\tau_j}(y_j)}{\tau_j!}+(-1)^{\alpha+1}\frac{\widetilde{B}_{2\alpha}(x_j-y_j)}{(2\alpha)!}\right).$$

Functions in this space have mixed partial derivatives up to order $\alpha-1$ in each variable which are absolutely continuous and of which the highest derivatives of order α are L_2 integrable. Next, we will consider a function space with the same properties, but in addition ask that the functions are "periodic".

All functions in the periodic space will allow to be expressed as absolutely converging Fourier series (which implies pointwise equality of the Fourier series)

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{h}\in\mathbb{Z}^d}\widehat{f}(\boldsymbol{h})\,\exp(2\pi\mathrm{i}\,\boldsymbol{h}\cdot\boldsymbol{x}),\quad \widehat{f}(\boldsymbol{h}) := \int_{[0,1]^d} f(\boldsymbol{x})\,\exp(-2\pi\mathrm{i}\,\boldsymbol{h}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

We define the periodic function space in such a way that the norms of the periodic functions are equal in both spaces for the same integer smoothness.

Definition 6.2. *For $\alpha > 1/2$ the space $\mathcal{H}(K^{\mathrm{Kor}}_{\alpha,d})$ is a Korobov space on the unit cube $[0,1]^d$, which is a reproducing kernel Hilbert space with inner product*

$$\langle f,g\rangle_{K^{\mathrm{Kor}}_{\alpha,d}} := \sum_{\boldsymbol{h}\in\mathbb{Z}^d}\widehat{f}(\boldsymbol{h})\,\overline{\widehat{g}(\boldsymbol{h})}\,\left[r_{\alpha,d}(\boldsymbol{h})\right]^2,$$

and reproducing kernel

$$K^{\mathrm{Kor}}_{\alpha,d}(\boldsymbol{x},\boldsymbol{y}) := \sum_{\boldsymbol{h}\in\mathbb{Z}^d}\frac{\exp(2\pi\mathrm{i}\,\boldsymbol{h}\cdot(\boldsymbol{x}-\boldsymbol{y}))}{\left[r_{\alpha,d}(\boldsymbol{h})\right]^2} = \prod_{j=1}^{d}\left(1+(-1)^{\alpha+1}\frac{\widetilde{B}_{2\alpha}(x_j-y_j)}{(2\alpha)!}\right), \tag{6.4}$$

where

$$r_{\alpha,d}(\boldsymbol{h}) := \prod_{j=1}^{d} r_\alpha(h_j), \qquad r_\alpha(h_j) := \begin{cases} 1, & \text{for } h_j = 0,\\ |2\pi\,h_j|^\alpha, & \text{for } h_j \neq 0,\end{cases}$$

For $\alpha\in\mathbb{N}$ the normalization in $r_{\alpha,d}$ is chosen in such a way that the inner product and norm coincides with that of the unanchored Sobolev space $\mathcal{H}(K^{\mathrm{Sob}}_{\alpha,d})$ of smoothness α for functions in $\mathcal{H}(K^{\mathrm{Kor}}_{\alpha,d})$. Function in the Korobov space $\mathcal{H}(K^{\mathrm{Kor}}_{\alpha,d})$ have mixed partial derivatives up to order $\alpha-1$ in each variable which are absolutely continuous, and of which derivatives of order α are L_2 integrable,

and, additionally, the derivatives up to order $\alpha - 1$ have matching values on the boundaries, i.e., periodic boundary conditions,

$$\forall \boldsymbol{\tau} \in \{0, \dots, \alpha - 1\}^d, \forall j = 1, \dots, d : f^{(\boldsymbol{\tau})}(\boldsymbol{x})\big|_{x_j=1} = f^{(\boldsymbol{\tau})}(\boldsymbol{x})\big|_{x_j=0},$$

and, as a consequence,

$$\forall \boldsymbol{\tau} \in \{1, \dots, \alpha\}^d, \forall \mathfrak{v} \subseteq \{1, \dots, d\} \setminus \{j : \tau_j = \alpha\}, \mathfrak{v} \neq \emptyset, \forall j \in \mathfrak{v} :$$

$$\int_{[0,1]^{|\mathfrak{v}|}} f^{(\boldsymbol{\tau})}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_{\mathfrak{v}} = \int_{[0,1]^{|\mathfrak{v}|-1}} \left[f^{(\boldsymbol{\tau}-1)}(\boldsymbol{x})\big|_{x_j=1} - f^{(\boldsymbol{\tau}-1)}(\boldsymbol{x})\big|_{x_j=0} \right] \mathrm{d}\boldsymbol{x}_{\mathfrak{v}\setminus\{j\}} = 0. \tag{6.5}$$

For $\boldsymbol{\tau} \in \{0, \dots, \alpha\}^d$ the mixed partial derivatives can be written in terms of Fourier series with modified coefficients:

$$f^{(\boldsymbol{\tau})}(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{h} \in \mathbb{Z}^d \\ h_j \neq 0 \text{ for } \tau_j \neq 0}} \left[\prod_{j=1}^{d} (2\pi \mathrm{i} \, h_j)^{\tau_j} \right] \widehat{f}(\boldsymbol{h}) \, \exp(2\pi \mathrm{i} \, \boldsymbol{h} \cdot \boldsymbol{x}),$$

which is L_2 integrable since $\|f\|^2_{K^{\mathrm{Kor}}_{\alpha,d}} < \infty$ for $f \in \mathcal{H}(K^{\mathrm{Kor}}_{\alpha,d})$. It is obvious that the Korobov space is a subspace of the Sobolev space and using the previous periodicity properties in the integrals in the inner product of the Sobolev space (for the $\tau_j \in \{1, \dots, \alpha - 1\}$) it can be seen that for $f, g \in \mathcal{H}(K^{\mathrm{Kor}}_{\alpha,d})$ the inner product for $\mathcal{H}(K^{\mathrm{Kor}}_{\alpha,d})$ with $\alpha \in \mathbb{N}$ can be written in two ways: using Fourier series or using derivatives

$$\begin{aligned}
\langle f, g \rangle_{K^{\mathrm{Kor}}_{\alpha,d}} &= \sum_{\boldsymbol{h} \in \mathbb{Z}^d} \widehat{f}(\boldsymbol{h}) \, \overline{\widehat{g}(\boldsymbol{h})} \, \left[r_{\alpha,d}(\boldsymbol{h}) \right]^2 \\
&= \sum_{\substack{\boldsymbol{\tau} \in \{0,\alpha\}^d \\ \mathfrak{w} := \{j : \tau_j = \alpha\}}} \int_{[\mathbf{0}_{\mathfrak{w}}, \mathbf{1}_{\mathfrak{w}}]} \left(\int_{[\mathbf{0}_{-\mathfrak{w}}, \mathbf{1}_{-\mathfrak{w}}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_{-\mathfrak{w}} \right) \\
&\qquad\qquad \left(\int_{[\mathbf{0}_{-\mathfrak{w}}, \mathbf{1}_{-\mathfrak{w}}]} g^{(\boldsymbol{\tau})}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_{-\mathfrak{w}} \right) \mathrm{d}\boldsymbol{x}_{\mathfrak{w}} \\
&= \langle f, g \rangle_{K^{\mathrm{Sob}}_{\alpha,d}},
\end{aligned}$$

where the sum in the second line is now only over $\boldsymbol{\tau}$ with components 0 or α because of the periodicity of the functions f and g.

6.2.3 RKHSs on a box

We now want to define a function space for non-periodic functions on a box $[\boldsymbol{a}, \boldsymbol{b}]$ where the inner product is a direct extension of the one over the unit cube from the previous section.

Definition 6.3. *For $\alpha \in \mathbb{N}$ the space $\mathcal{H}(K_\alpha^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]})$ is an unanchored Sobolev space on the box $[\boldsymbol{a}, \boldsymbol{b}]$, which is a reproducing kernel Hilbert space with inner product*

$$\langle f, g \rangle_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}} := \sum_{\substack{\boldsymbol{\tau} \in \{0,\dots,\alpha\}^d \\ \mathfrak{w} := \{j : \tau_j = \alpha\}}} \int_{[\boldsymbol{a}_{\mathfrak{w}}, \boldsymbol{b}_{\mathfrak{w}}]} \left(\int_{[\boldsymbol{a}_{-\mathfrak{w}}, \boldsymbol{b}_{-\mathfrak{w}}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_{-\mathfrak{w}} \right) \left(\int_{[\boldsymbol{a}_{-\mathfrak{w}}, \boldsymbol{b}_{-\mathfrak{w}}]} g^{(\boldsymbol{\tau})}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_{-\mathfrak{w}} \right) \mathrm{d}\boldsymbol{x}_{\mathfrak{w}}, \quad (6.6)$$

and reproducing kernel

$$K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x}, \boldsymbol{y}) := \prod_{j=1}^{d} \left(\frac{1}{(b_j - a_j)^2} + \sum_{\tau_j=1}^{\alpha-1} \frac{B_{\tau_j}^{[a_j,b_j]}(x_j)}{\tau_j!} \frac{B_{\tau_j}^{[a_j,b_j]}(y_j)}{\tau_j!} \right.$$

$$\left. + (b_j - a_j) \frac{B_\alpha^{[a_j,b_j]}(x_j)}{\alpha!} \frac{B_\alpha^{[a_j,b_j]}(y_j)}{\alpha!} + (-1)^{\alpha+1} \frac{\widetilde{B}_{2\alpha}^{[a_j,b_j]}(x_j - y_j + a_j)}{(2\alpha)!} \right). \quad (6.7)$$

We will show in Proposition 6.9 that $K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}$ is indeed the kernel corresponding to the inner product above.

For the periodic space we define scaled periodic basis functions which are just the tensor product of the one-dimensional Fourier basis on $[a, b]$ given in (6.2), that is, for $\boldsymbol{h} \in \mathbb{Z}^d$, we define

$$\varphi_{\boldsymbol{h}}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x}) := \prod_{j=1}^{d} \varphi_{h_j}^{[a_j,b_j]}(x_j) = \prod_{j=1}^{d} \frac{\exp(2\pi \mathrm{i}\, h_j (x_j - a_j)/(b_j - a_j))}{\sqrt{b_j - a_j}}.$$

They form an orthogonal and normalized set of basis functions against the standard L_2 inner product on the box $[\boldsymbol{a}, \boldsymbol{b}]$. Therefore, our periodic functions are expressed as absolutely converging Fourier series with respect to the basis

$\varphi_{\boldsymbol{h}}^{[\boldsymbol{a},\boldsymbol{b}]}$ as follows

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{h}\in\mathbb{Z}^d} \widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}\, \varphi_{\boldsymbol{h}}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x}), \qquad \widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h}) := \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\, \overline{\varphi_{\boldsymbol{h}}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x})}\, \mathrm{d}\boldsymbol{x}.$$

Definition 6.4. *For $\alpha > 1/2$ the space $\mathcal{H}(K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]})$ is a Korobov space on the box $[\boldsymbol{a},\boldsymbol{b}]$, which is a reproducing kernel Hilbert space with inner product*

$$\langle f, g\rangle_{K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}} := \sum_{\boldsymbol{h}\in\mathbb{Z}^d} \widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h})\, \overline{\widehat{g}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h})}\, \left[r_{\alpha,d}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h})\right]^2,$$

and reproducing kernel

$$\begin{aligned} K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x},\boldsymbol{y}) &:= \sum_{\boldsymbol{h}\in\mathbb{Z}^d} [r_{\alpha,d}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h})]^{-2}\, \varphi_{\boldsymbol{h}}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x})\, \overline{\varphi_{\boldsymbol{h}}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{y})} \\ &= \prod_{j=1}^{d} \left(\frac{1}{(b_j-a_j)^2} + (-1)^{\alpha+1} \frac{\widetilde{B}_{2\alpha}^{[a_j,b_j]}(x_j - y_j + a_j)}{(2\alpha)!} \right), \end{aligned}$$

where

$$r_{\alpha,d}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h}) := \prod_{j=1}^{d} r_{\alpha}^{[a_j,b_j]}(h_j), \qquad r_{\alpha}^{[a_j,b_j]}(h_j) := \begin{cases} \sqrt{b_j - a_j}, & \text{for } h_j = 0, \\ \dfrac{|2\pi\, h_j|^{\alpha}}{(b_j-a_j)^{\alpha}}, & \text{for } h_j \neq 0, \end{cases}$$

The choice of $r_{\alpha}^{[a,b]}$ made in the definition is chosen such that functions in the Sobolev space of smoothness $\alpha \in \mathbb{N}$ which are periodic will have the same norm in this Korobov space. This is shown in the next lemma.

Lemma 6.1. *For $\alpha \in \mathbb{N}$ and $f \in \mathcal{H}(K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}) \subset \mathcal{H}(K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]})$ we have*

$$\|f\|_{K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}} = \|f\|_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}}.$$

Proof. As is the case in the unit cube we only need to consider partial derivatives of order 0 and α in the norm of the Sobolev space when the function is periodic as all the interior integrals vanish due to the periodicity, see (6.5). For

$f, g \in \mathcal{H}(K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]})$ we can thus write

$$\langle f, g \rangle_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}} = \sum_{\substack{\boldsymbol{\tau} \in \{0,\alpha\}^d \\ \mathfrak{w} := \{j : \tau_j = \alpha\}}} \int_{[\boldsymbol{a}_{\mathfrak{w}}, \boldsymbol{b}_{\mathfrak{w}}]} \left(\int_{[\boldsymbol{a}_{-\mathfrak{w}}, \boldsymbol{b}_{-\mathfrak{w}}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_{-\mathfrak{w}} \right) \left(\int_{[\boldsymbol{a}_{-\mathfrak{w}}, \boldsymbol{b}_{-\mathfrak{w}}]} g^{(\boldsymbol{\tau})}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_{-\mathfrak{w}} \right) \mathrm{d}\boldsymbol{x}_{\mathfrak{w}}.$$

Therefore, for $f \in \mathcal{H}(K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]})$ and $\boldsymbol{\tau} \in \{0,\alpha\}^d$ we have, with $\mathfrak{w} = \{j : \tau_j = \alpha\}$,

$$\begin{aligned}
&\int_{[\boldsymbol{a}_{\mathfrak{w}},\boldsymbol{b}_{\mathfrak{w}}]} \left(\int_{[\boldsymbol{a}_{-\mathfrak{w}},\boldsymbol{b}_{-\mathfrak{w}}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{w}} \right)^2 \mathrm{d}\boldsymbol{x}_{\mathfrak{w}} \\
&\quad = \int_{[\boldsymbol{a}_{\mathfrak{w}},\boldsymbol{b}_{\mathfrak{w}}]} \left(\sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^d \\ h_j\neq 0 \text{ for } j\in\mathfrak{w}}} \widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h}) \left[\prod_{j\in\mathfrak{w}} \frac{(2\pi\mathrm{i}\,h_j)^\alpha}{(b_j-a_j)^\alpha}\, \varphi_{h_j}^{[a_j,b_j]}(x_j) \right] \right. \\
&\qquad\qquad \left. \left[\prod_{j\notin\mathfrak{w}} \underbrace{\int_{a_j}^{b_j} \varphi_{h_j}^{[a_j,b_j]}(x_j)\,\mathrm{d}x_j}_{=0 \text{ if } h_j\neq 0} \right] \right)^2 \mathrm{d}\boldsymbol{x}_{\mathfrak{w}} \\
&\quad = \int_{[\boldsymbol{a}_{\mathfrak{w}},\boldsymbol{b}_{\mathfrak{w}}]} \left(\sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^d \\ h_j\neq 0 \text{ for } j\in\mathfrak{w} \\ h_j=0 \text{ for } j\notin\mathfrak{w}}} \widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h}) \left[\prod_{j\in\mathfrak{w}} \frac{(2\pi\mathrm{i}\,h_j)^\alpha}{(b_j-a_j)^\alpha}\, \varphi_{h_j}^{[a_j,b_j]}(x_j) \right] \right. \\
&\qquad\qquad \left. \left[\prod_{j\notin\mathfrak{w}} (b_j-a_j)^{1-1/2} \right] \right)^2 \mathrm{d}\boldsymbol{x}_{\mathfrak{w}} \\
&\quad = \left(\prod_{j\notin\mathfrak{w}} \sqrt{b_j-a_j} \right)^2 \int_{[\boldsymbol{a}_{\mathfrak{w}},\boldsymbol{b}_{\mathfrak{w}}]} \left(\sum_{\boldsymbol{h}_{\mathfrak{w}}\in(\mathbb{Z}\setminus\{0\})^{|\mathfrak{w}|}} \widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h}_{\mathfrak{w}};\boldsymbol{0}_{-\mathfrak{w}}) \right. \\
&\qquad\qquad \left. \left[\prod_{j\in\mathfrak{w}} \frac{(2\pi\mathrm{i}\,h_j)^\alpha}{(b_j-a_j)^\alpha}\, \varphi_{h_j}^{[a_j,b_j]}(x_j) \right] \right)^2 \mathrm{d}\boldsymbol{x}_{\mathfrak{w}} \\
&\quad = \sum_{\boldsymbol{h}_{\mathfrak{w}}\in(\mathbb{Z}\setminus\{0\})^{|\mathfrak{w}|}} \left| \widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h}) \right|^2 [r_{\alpha,d}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h})]^2,
\end{aligned}$$

and thus for $f \in \mathcal{H}(K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]})$

$$\begin{aligned}\|f\|^2_{K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}} &= \sum_{\boldsymbol{h}\in\mathbb{Z}^d} |\widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h})|^2 \, [r_{\alpha,d}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h})]^2 \\ &= \sum_{\mathfrak{w}\subseteq\{1,\ldots,d\}} \sum_{\boldsymbol{h}_{\mathfrak{w}}\in(\mathbb{Z}\setminus\{0\})^{|\mathfrak{w}|}} \left|\widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h})\right|^2 \, [r_{\alpha,d}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h})]^2 = \|f\|^2_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}}.\end{aligned}$$

□

Since we will apply a scaled lattice rule to a function in $\mathcal{H}(K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]})$ we show how to bound the error.

Proposition 6.2. *For $f \in \mathcal{H}(K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}) \subset \mathcal{H}(K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]})$ the error of using an n-point lattice rule with generating vector $\boldsymbol{z} \in \mathbb{Z}^d$ and scaled nodes $\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}$ can be bounded as*

$$\begin{aligned}&\left|\frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| \\ &\qquad \le \|f\|_{K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}} \left(\sum_{\boldsymbol{0}\neq\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{z},n)} [r_{\alpha,d}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h})]^{-2} \prod_{j=1}^d (b_j - a_j)\right)^{1/2} \\ &\le \|f\|_{K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}} \left(\prod_{j=1}^d \max(1, b_j - a_j)^{\alpha+1/2}\right) \left(\sum_{\boldsymbol{0}\neq\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{z},n)} [r_{\alpha,d}(\boldsymbol{h})]^{-2}\right)^{1/2},\end{aligned}$$

where $\Lambda^\perp(\boldsymbol{z},n) := \{\boldsymbol{h}\in\mathbb{Z}^d : \boldsymbol{h}\cdot\boldsymbol{z} \equiv 0 \pmod n\}$ is the dual of the lattice. The norm $\|f\|_{K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}}$ can be replaced by the norm in the associated Sobolev space since $\|f\|_{K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}} = \|f\|_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}}$ for the spaces here defined.

Proof. First we note that

$$\widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{0}) \prod_{j=1}^d \sqrt{b_j - a_j} = \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Therefore

$$\begin{aligned}
&\frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \\
&\quad = \sum_{\boldsymbol{h}\in\mathbb{Z}^d} \widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h}) \frac{\prod_{j=1}^d \sqrt{b_j - a_j}}{n} \sum_{i=1}^n \exp(2\pi \mathrm{i}\, \boldsymbol{h}\cdot\boldsymbol{z}\, i/n) - \widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{0}) \prod_{j=1}^d \sqrt{b_j - a_j} \\
&\quad = \left[\prod_{j=1}^d \sqrt{b_j - a_j}\right] \sum_{\boldsymbol{0}\neq\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{z},n)} \widehat{f}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h}),
\end{aligned}$$

where we made use of the character property, for $t \in \mathbb{Z}$,

$$\frac{1}{n}\sum_{i=1}^n \exp(2\pi \mathrm{i}\, t\, i/n) = \begin{cases} 1, & \text{if } t \equiv 0 \pmod n, \\ 0, & \text{otherwise,} \end{cases}$$

with $t = \boldsymbol{h}\cdot\boldsymbol{z}$. The result now follows by multiplying and dividing with $r_{\alpha,d}^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h})$ and applying the Cauchy–Schwarz inequality to arrive at the norm of f in the Korobov space $\mathcal{H}(K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]})$. □

6.3 The Bernoulli polynomial method as a projection

In this section, we are going to introduce the Bernoulli polynomial method introduced by Korobov [40], and also considered by Zaremba [94]. This method connects two different spaces: the unanchored Sobolev space and Korobov space.

6.3.1 The Bernoulli polynomial method on the unit cube

In this section, we review the method and conclude that this method coincides with an orthogonal projection from an unanchored Sobolev space $\mathcal{H}(K_{\alpha,d}^{\mathrm{Sob}})$ to a Korobov space $\mathcal{H}(K_{\alpha,d}^{\mathrm{Kor}})$ where those two spaces have exactly the same norms for a function $f \in \mathcal{H}(K_{\alpha,d}^{\mathrm{Kor}}) \subset \mathcal{H}(K_{\alpha,d}^{\mathrm{Sob}})$. In Section 6.4, this projection will be used for measuring the non-periodicity of a function.

In [40], the Bernoulli polynomial method was originally introduced. In the present paper, to introduce the method we follow the definition by [7, 80] for

the sake of notational simplicity. Define the recursion

$$\begin{cases} F_{[0]}(\boldsymbol{x}) & := f(\boldsymbol{x}), \\ F_{[j]}(\boldsymbol{x}) & := F_{[j-1]}(\boldsymbol{x}) - \displaystyle\sum_{\tau_j=1}^{\alpha} \frac{B_{\tau_j}(x_j)}{\tau_j!} \left[\frac{\partial^{\tau_j-1} F_{[j-1]}(\boldsymbol{x})}{\partial x_j^{\tau_j-1}} \right]_{x_j=0}^{1} \\ & = F_{[j-1]}(\boldsymbol{x}) - \displaystyle\sum_{\tau_j=1}^{\alpha} \frac{B_{\tau_j}(x_j)}{\tau_j!} \int_0^1 \frac{\partial^{\tau_j} F_{[j-1]}(\boldsymbol{x})}{\partial x_j^{\tau_j}} \,\mathrm{d}x_j \qquad \text{for } j = 1, \ldots, d, \end{cases} \tag{6.8}$$

where we used our assumption that for our function f all derivatives up to order $\alpha - 1$ are absolutely continuous such that we can replace the difference by an integral. The final function

$$F := F_{[d]}$$

now has the property of the Korobov space in that it satisfies that all derivatives up to order $\alpha - 1$, in all dimensions, have matching boundary values. Furthermore, since the Bernoulli polynomials integrate to zero, the integral of F is equal to the integral of the original function f.

From the recursive definition above we can reach the following expression (see [94, p. 61]):

$$F(\boldsymbol{x}) = f(\boldsymbol{x}) + \sum_{\substack{\mathbf{0} \neq \boldsymbol{\tau} \in \{0,\ldots,\alpha\}^d \\ \mathfrak{u} := \mathrm{supp}(\boldsymbol{\tau})}} (-1)^{|\mathfrak{u}|} \left[\prod_{j \in \mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!} \right] \int_{[0,1]^{|\mathfrak{u}|}} f^{(\boldsymbol{\tau})}(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}_{\mathfrak{u}}. \tag{6.9}$$

We remark that the product over $j \in \mathfrak{u}$ can be replaced by the full product over $j = 1, \ldots, d$ since $B_0(x)/0! = 1$. In the proof of the next lemma we demonstrate that (6.8) and (6.9) are the same by an inductive proof, as this formal method returns later in the proof of Lemma 6.4 and Proposition 6.5. We remark that (6.9) can also be written in terms of differences as in the first form of (6.8), but we write the integral form for conciseness, using our assumption that all derivatives up to order $\alpha - 1$ are absolutely continuous.

Lemma 6.3. *The recursive definition* (6.8) *and* (6.9) *are the same.*

Proof. We prove this inductively. We start with the base case and obtain immediately from (6.8) that

$$F_{[1]}(\boldsymbol{x}) = \sum_{\mathfrak{u} \subseteq \{1\}} (-1)^{|\mathfrak{u}|} \sum_{\boldsymbol{\tau}_{\mathfrak{u}} \in \{1,\ldots,\alpha\}^{|\mathfrak{u}|}} \left[\prod_{j \in \mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!} \right] \int_{[0,1]^{|\mathfrak{u}|}} f^{(\tau_1,0,\ldots,0)}(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}_{\mathfrak{u}},$$

since for $\mathfrak{u} = \emptyset$ we have $\tau_1 = 0$ and obtain $f(\boldsymbol{x}) = F_{[0]}(\boldsymbol{x})$, while for $\mathfrak{u} = \{1\}$ we obtain the sum over $\tau_1 = \tau_j$ from (6.8) with $F_{[0]}(\boldsymbol{x}) = f(\boldsymbol{x})$. Now assume the first $j-1$ dimensions are handled and $F_{[j-1]}$ can be written as follows (by the induction hypothesis)

$$F_{[j-1]}(\boldsymbol{x}) = \sum_{\mathfrak{u}\subseteq\{1,\ldots,j-1\}} (-1)^{|\mathfrak{u}|} \sum_{\boldsymbol{\tau}_\mathfrak{u}\in\{1,\ldots,\alpha\}^{|\mathfrak{u}|}} \left[\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right] \int_{[0,1]^{|\mathfrak{u}|}} f^{(\tau_1,\ldots,\tau_{j-1},0,\ldots,0)}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_\mathfrak{u},$$

then, by the recursion formula and the same reasoning as for the expression of the base case,

$$F_{[j]}(\boldsymbol{x}) = \sum_{\mathfrak{v}\subseteq\{j\}} (-1)^{|\mathfrak{v}|} \sum_{\boldsymbol{\tau}_\mathfrak{v}\in\{1,\ldots,\alpha\}^{|\mathfrak{v}|}} \sum_{\mathfrak{u}\subseteq\{1,\ldots,j-1\}} (-1)^{|\mathfrak{u}|} \sum_{\boldsymbol{\tau}_\mathfrak{u}\in\{1,\ldots,\alpha\}^{|\mathfrak{u}|}} \left[\prod_{j\in\mathfrak{v}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right] \left[\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right] \int_{[0,1]^{|\mathfrak{v}|}} \int_{[0,1]^{|\mathfrak{u}|}} f^{(\tau_1,\ldots,\tau_{j-1},\tau_j,0,\ldots,0)}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_\mathfrak{u}\,\mathrm{d}\boldsymbol{x}_\mathfrak{v}.$$

From this expression it is clear that we can merge the disjunctive sets $\mathfrak{u}$ and $\mathfrak{v}$ to obtain

$$F_{[j]}(\boldsymbol{x}) = \sum_{\mathfrak{u}\subseteq\{1,\ldots,j\}} (-1)^{|\mathfrak{u}|} \sum_{\boldsymbol{\tau}_\mathfrak{u}\in\{1,\ldots,\alpha\}^{|\mathfrak{u}|}} \left[\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right] \int_{[0,1]^{|\mathfrak{u}|}} f^{(\tau_1,\ldots,\tau_j,0,\ldots,0)}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_\mathfrak{u}.$$

The expression for $F_{[d]} = F$ can then be rewritten in the form (6.9) by pulling out $\mathfrak{u} = \emptyset$ to obtain $f(\boldsymbol{x})$ and joining the sums over $\mathfrak{u}$ and $\boldsymbol{\tau}_\mathfrak{u}$. □

Lemma 6.4. *For a function f having mixed partial derivatives up to order $\alpha-1$ in each variable which are absolutely continuous and of which the derivatives with the highest order α are integrable, the following holds for all $\boldsymbol{x} \in [0,1]^d$:*

$$\sum_{\mathfrak{u}\subseteq\{1,\ldots,d\}} (-1)^{d-|\mathfrak{u}|} \sum_{\substack{\boldsymbol{\tau}_\mathfrak{u}\in\{0,\ldots,\alpha\}^{|\mathfrak{u}|} \\ \tau_j=0 \text{ for } j\notin\mathfrak{u}}} \left[\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right] \int_{[0,1]^{|\mathfrak{u}|}} f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_\mathfrak{u} = \int_{[0,1]^d} \left[\prod_{j=1}^{d} \frac{\widetilde{B}_\alpha(x_j - y_j)}{\alpha!}\right] f^{(\alpha,\ldots,\alpha)}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y},$$

where the integral over the empty domain $[0,1]^{|\mathfrak{u}|}$ *when* $\mathfrak{u}=\emptyset$ *is the identity operation. In this way the term for* $\mathfrak{u}=\emptyset$ *in the left hand side corresponds to* $(-1)^d f(\boldsymbol{x})$.

Proof. First we prove the one-dimensional case. For $\tau = 1,\dots,\alpha-1$ define

$$F_\tau(x) := \int_0^1 f^{(\tau)}(y)\frac{\widetilde{B}_\tau(x-y)}{\tau!}\,\mathrm{d}y.$$

Making use of the properties of the Bernoulli polynomials from Section 6.2.1 we obtain the following results. For $\tau = 1,\dots,\alpha-1$, using integration by parts on $F_\tau(x)$ gives

$$\begin{aligned}
&\int_0^1 f^{(\tau)}(y)\frac{\widetilde{B}_\tau(x-y)}{\tau!}\,\mathrm{d}y \\
&= \left[f^{(\tau)}(y)\frac{-\widetilde{B}_{\tau+1}(x-y)}{(\tau+1)!}\right]_{y=0}^{1} - \int_0^1 f^{(\tau+1)}(y)\frac{-\widetilde{B}_{\tau+1}(x-y)}{(\tau+1)!}\,\mathrm{d}y \\
&\Leftrightarrow F_{\tau+1}(x) - F_\tau(x) = \frac{\widetilde{B}_{\tau+1}(x)}{(\tau+1)!}\left(f^{(\tau)}(1)-f^{(\tau)}(0)\right) = \frac{\widetilde{B}_{\tau+1}(x)}{(\tau+1)!}\int_0^1 f^{(\tau+1)}(y)\,\mathrm{d}y.
\end{aligned}$$

We remark that the integration by parts, $\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u$, is valid since $u = f^{(\tau)}(y)$ is absolutely continuous for all $\tau = 0,\dots,\alpha-1$ and $\mathrm{d}v = (\widetilde{B}_\tau(x-y)/\tau!)\,\mathrm{d}y$ is Lebesgue integrable [75]. Therefore, using a telescoping sum, we obtain

$$\begin{aligned}
F_\alpha(x) - F_1(x) = \sum_{\tau=1}^{\alpha-1} F_{\tau+1}(x) - F_\tau(x) &= \sum_{\tau=1}^{\alpha-1}\frac{\widetilde{B}_{\tau+1}(x)}{(\tau+1)!}\int_0^1 f^{(\tau+1)}(y)\,\mathrm{d}y \\
&= \sum_{\tau=2}^{\alpha}\frac{B_\tau(x)}{\tau!}\int_0^1 f^{(\tau)}(y)\,\mathrm{d}y \quad \text{for } x\in[0,1].
\end{aligned} \tag{6.10}$$

Note that the last line is obtained by shifting the index and since $\widetilde{B}_\tau(x) = B_\tau(x)$ for $\tau \geq 2$ and $x \in [0,1]$. Furthermore, for $\tau = 1$, we can also use integration by parts with the roles of u and $\mathrm{d}v$ interchanged, carefully splitting the integral at

the discontinuity of $\widetilde{B}_1$, such that for $x \in [0,1]$ we have

$$\begin{aligned}
F_1(x) &= \int_0^{x^-} \widetilde{B}_1(x-y)f'(y)\,\mathrm{d}y + \int_{x^+}^1 \widetilde{B}_1(x-y)f'(y)\,\mathrm{d}y \\
&= \int_0^{x^-} B_1(x-y)f'(y)\,\mathrm{d}y + \int_{x^+}^1 B_1(x-y+1)f'(y)\,\mathrm{d}y \\
&= \Big[B_1(x-y)f(y)\Big]_{y=0}^{x^-} + \int_0^{x^-} f(y)\,\mathrm{d}y \\
&\quad + \Big[B_1(x-y+1)f(y)\Big]_{y=x^+}^{1} + \int_{x^+}^1 f(y)\,\mathrm{d}y \\
&= -f(x) + B_1(x)(f(1)-f(0)) + \int_0^1 f(y)\,\mathrm{d}y.
\end{aligned}$$

Substituting this into (6.10) then gives, for $x \in [0,1]$,

$$\begin{aligned}
F_\alpha(x) &= -f(x) + \sum_{\tau=1}^{\alpha} \frac{B_\tau(x)}{\tau!} \int_0^1 f^{(\tau)}(y)\,\mathrm{d}y + \int_0^1 f(y)\,\mathrm{d}y \\
&= -f(x) + \sum_{\tau=0}^{\alpha} \frac{B_\tau(x)}{\tau!} \int_0^1 f^{(\tau)}(y)\,\mathrm{d}y \qquad (6.11) \\
&= \sum_{\substack{\mathfrak{u}\subseteq\{1\} \\ \overline{\mathfrak{u}}:=\{1\}\setminus\mathfrak{u}}} (-1)^{|\overline{\mathfrak{u}}|} \sum_{\substack{\boldsymbol{\tau}_\mathfrak{u}\in\{0,\ldots,\alpha\}^{|\mathfrak{u}|} \\ \tau_j=0 \text{ for } j\in\overline{\mathfrak{u}}}} \left[\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right] \int_{[0,1]^{|\mathfrak{u}|}} f^{(\tau_1)}(x_1)\,\mathrm{d}x_\mathfrak{u}, \qquad (6.12)
\end{aligned}$$

where the last line is already written in the notation of the multivariate statement since this is the form we want to use further down. In other words we have obtained the following series expansion with remainder term, for $x \in [0,1]$,

$$f(x) = \sum_{\tau=0}^{\alpha} \frac{B_\tau(x)}{\tau!} \int_0^1 f^{(\tau)}(y)\,\mathrm{d}y - \int_0^1 \frac{\widetilde{B}_\alpha(x-y)}{\alpha!} f^{(\alpha)}(y)\,\mathrm{d}y. \qquad (6.13)$$

To derive the multivariate claim we will apply (6.11) recursively, starting from the following expression, where we apply (6.11) for the first dimension in the

form of (6.12) such that the sums over $\mathfrak{u}$ and $\boldsymbol{\tau}$ can be moved to the front:

$$\int_{[0,1]^d} \left[\prod_{j=1}^{d} \frac{\widetilde{B}_\alpha(x_j - y_j)}{\alpha!}\right] f^{(\alpha,\dots,\alpha)}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}$$

$$= \int_{[0,1]^{d-1}} \left[\prod_{j=2}^{d} \frac{\widetilde{B}_\alpha(x_j - y_j)}{\alpha!}\right] \left[\int_0^1 \frac{\widetilde{B}_\alpha(x_1 - y_1)}{\alpha!} f^{(\alpha,\dots,\alpha)}(\boldsymbol{y})\,\mathrm{d}y_1\right] \mathrm{d}\boldsymbol{y}_{-\{1\}}$$

$$= \int_{[0,1]^{d-1}} \left[\prod_{j=2}^{d} \frac{\widetilde{B}_\alpha(x_j - y_j)}{\alpha!}\right] \Bigg[- f^{(0,\alpha,\dots,\alpha)}(x_1, \boldsymbol{y}_{-\{1\}}) + \sum_{\tau_1=0}^{\alpha} \frac{B_{\tau_1}(x_1)}{\tau_1!} \int_0^1 f^{(\tau_1,\alpha,\dots,\alpha)}(\boldsymbol{y})\,\mathrm{d}y_1\Bigg] \mathrm{d}\boldsymbol{y}_{-\{1\}}$$

$$= \sum_{\substack{\mathfrak{u}\subseteq\{1\}\\ \overline{\mathfrak{u}}:=\{1\}\setminus\mathfrak{u}}} (-1)^{|\overline{\mathfrak{u}}|} \sum_{\substack{\boldsymbol{\tau}_\mathfrak{u}\in\{0,\dots,\alpha\}^{|\mathfrak{u}|}\\ \tau_j=0 \text{ for } j\in\overline{\mathfrak{u}}}} \left[\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right]$$

$$\int_{[0,1]^{|\mathfrak{u}|}} \left[\int_{[0,1]^{d-1}} \left[\prod_{j=2}^{d} \frac{\widetilde{B}_\alpha(x_j - y_j)}{\alpha!}\right] f^{(\tau_1,\alpha,\dots,\alpha)}(x_1, \boldsymbol{y}_{-\{1\}})\,\mathrm{d}\boldsymbol{y}_{-\{1\}}\right] \mathrm{d}x_\mathfrak{u},$$

where $\boldsymbol{y}_{-\{1\}} = (y_2, \ldots, y_d)$ is the complement w.r.t. the full dimensional set, i.e., $\{1, \ldots, d\} \setminus \{1\}$; at this point the part under the outside integral on the second line of the previous expression is of the form we started from and we repeat the procedure now for the second dimension to obtain

$$
\begin{aligned}
&= \sum_{\substack{\mathfrak{u} \subseteq \{1\} \\ \overline{\mathfrak{u}} := \{1\} \setminus \mathfrak{u}}} (-1)^{|\overline{\mathfrak{u}}|} \sum_{\substack{\boldsymbol{\tau}_{\mathfrak{u}} \in \{0,\ldots,\alpha\}^{|\mathfrak{u}|} \\ \tau_j = 0 \text{ for } j \in \overline{\mathfrak{u}}}} \left[\prod_{j \in \mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!} \right] \int_{[0,1]^{|\mathfrak{u}|}} \left[\int_{[0,1]^{d-2}} \left[\prod_{j=3}^{d} \frac{\widetilde{B}_\alpha(x_j - y_j)}{\alpha!} \right] \right. \\
&\quad \left. \left[\int_0^1 \frac{\widetilde{B}_\alpha(x_2 - y_2)}{\alpha!} f^{(\tau_1, \alpha, \ldots, \alpha)}(x_1, y_2, \boldsymbol{y}_{-\{1,2\}}) \, \mathrm{d}y_2 \right] \mathrm{d}\boldsymbol{y}_{-\{1,2\}} \right] \mathrm{d}\boldsymbol{x}_{\mathfrak{u}} \\
&= \sum_{\substack{\mathfrak{u} \subseteq \{1\} \\ \overline{\mathfrak{u}} := \{1\} \setminus \mathfrak{u}}} (-1)^{|\overline{\mathfrak{u}}|} \sum_{\substack{\boldsymbol{\tau}_{\mathfrak{u}} \in \{0,\ldots,\alpha\}^{|\mathfrak{u}|} \\ \tau_j = 0 \text{ for } j \in \overline{\mathfrak{u}}}} \left[\prod_{j \in \mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!} \right] \\
&\quad \sum_{\substack{\mathfrak{v} \subseteq \{2\} \\ \overline{\mathfrak{v}} := \{2\} \setminus \mathfrak{v}}} (-1)^{|\overline{\mathfrak{v}}|} \sum_{\substack{\boldsymbol{\tau}_{\mathfrak{v}} \in \{0,\ldots,\alpha\}^{|\mathfrak{v}|} \\ \tau_j = 0 \text{ for } j \in \overline{\mathfrak{v}}}} \left[\prod_{j \in \mathfrak{v}} \frac{B_{\tau_j}(x_j)}{\tau_j!} \right] \\
&\quad \int_{[0,1]^{|\mathfrak{u}|}} \int_{[0,1]^{|\mathfrak{v}|}} \left[\int_{[0,1]^{d-2}} \left[\prod_{j=3}^{d} \frac{\widetilde{B}_\alpha(x_j - y_j)}{\alpha!} \right] f^{(\tau_1, \tau_2, \alpha, \ldots, \alpha)}(x_1, x_2, \boldsymbol{y}_{-\{1,2\}}) \, \mathrm{d}\boldsymbol{y}_{-\{1,2\}} \right] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \mathrm{d}\boldsymbol{x}_{\mathfrak{v}} \, \mathrm{d}\boldsymbol{x}_{\mathfrak{u}} \\
&= \sum_{\substack{\mathfrak{w} \subseteq \{1,2\} \\ \overline{\mathfrak{w}} := \{1,2\} \setminus \mathfrak{w}}} (-1)^{|\overline{\mathfrak{w}}|} \sum_{\substack{\boldsymbol{\tau}_{\mathfrak{w}} \in \{0,\ldots,\alpha\}^{|\mathfrak{w}|} \\ \tau_j = 0 \text{ for } j \in \overline{\mathfrak{w}}}} \left[\prod_{j \in \mathfrak{w}} \frac{B_{\tau_j}(x_j)}{\tau_j!} \right] \\
&\quad \int_{[0,1]^{|\mathfrak{w}|}} \left[\int_{[0,1]^{d-2}} \left[\prod_{j=3}^{d} \frac{\widetilde{B}_\alpha(x_j - y_j)}{\alpha!} \right] f^{(\tau_1, \tau_2, \alpha, \ldots, \alpha)}(x_1, x_2, \boldsymbol{y}_{-\{1,2\}}) \, \mathrm{d}\boldsymbol{y}_{-\{1,2\}} \right] \mathrm{d}\boldsymbol{x}_{\mathfrak{w}},
\end{aligned}
$$

and therefore

$$= \sum_{\mathfrak{u}\subseteq\{1,\dots,d\}} (-1)^{d-|\mathfrak{u}|} \sum_{\substack{\boldsymbol{\tau}_\mathfrak{u}\in\{0,\dots,\alpha\}^{|\mathfrak{u}|}\\ \tau_j=0 \text{ for } j\notin\mathfrak{u}}} \left[\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right] \int_{[0,1]^{|\mathfrak{u}|}} f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_\mathfrak{u}.$$

This concludes the proof. □

It is useful to compare the representation of a function f as in the previous statement to other possible representations.

Proposition 6.5. *For a function f having mixed partial derivatives up to order $\alpha-1$ in each variable which are absolutely continuous and of which the derivatives with the highest order α are integrable, we have the following equivalent representations:*
(i) series representation in terms of Bernoulli polynomials with remainder term

$$f(\boldsymbol{x}) = \sum_{\emptyset\neq\mathfrak{u}\subseteq\{1,\dots,d\}} (-1)^{|\mathfrak{u}|+1} \sum_{\substack{\boldsymbol{\tau}_\mathfrak{u}\in\{0,\dots,\alpha\}^{|\mathfrak{u}|}\\ \tau_j=0 \text{ for } j\notin\mathfrak{u}}} \left[\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right] \int_{[0,1]^{|\mathfrak{u}|}} f^{(\boldsymbol{\tau})}(\boldsymbol{x}_{-\mathfrak{u}},\boldsymbol{y}_\mathfrak{u})\,\mathrm{d}\boldsymbol{y}_\mathfrak{u}$$
$$+ (-1)^d \int_{[0,1]^d} \left[\prod_{j=1}^{d} \frac{\widetilde{B}_\alpha(x_j-y_j)}{\alpha!}\right] f^{(\alpha,\dots,\alpha)}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y},$$

(ii) series representation in terms of Bernoulli polynomials with mixed remainder terms

$$f(\boldsymbol{x}) = \sum_{\mathfrak{u}\subseteq\{1,\dots,d\}} (-1)^{d-|\mathfrak{u}|} \sum_{\substack{\boldsymbol{\tau}_\mathfrak{u}\in\{0,\dots,\alpha\}^{|\mathfrak{u}|}\\ \tau_j=\alpha \text{ for } j\notin\mathfrak{u}}} \left[\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right]$$
$$\int_{[0,1]^d} \left[\prod_{j\notin\mathfrak{u}} \frac{\widetilde{B}_\alpha(x_j-y_j)}{\alpha!}\right] f^{(\boldsymbol{\tau})}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}.$$

Proof. Representation *(i)* follows immediately from Lemma 6.4. For representation *(ii)* we start from the one-dimensional form (6.13), which, says that for all

such f and for all $x \in [0,1]$ we have the representation

$$f(x) = \sum_{\tau=0}^{\alpha} \frac{B_\tau(x)}{\tau!} \int_0^1 f^{(\tau)}(y)\,\mathrm{d}y - \int_0^1 \frac{\widetilde{B}_\alpha(x-y)}{\alpha!} f^{(\alpha)}(y)\,\mathrm{d}y$$

$$= \sum_{\substack{\mathfrak{u}\subseteq\{1\} \\ \overline{\mathfrak{u}}:=\{1\}\setminus\mathfrak{u}}} (-1)^{|\overline{\mathfrak{u}}|} \sum_{\substack{\boldsymbol{\tau}_{\mathfrak{u}}\in\{0,\ldots,\alpha\}^{|\mathfrak{u}|} \\ \tau_j=\alpha \text{ for } j\in\overline{\mathfrak{u}}}} \left[\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right] \int_0^1 \left[\prod_{j\in\overline{\mathfrak{u}}} \frac{\widetilde{B}_\alpha(x_j-y_j)}{\alpha!}\right] f^{(\tau_1)}(y_1)\,\mathrm{d}y_1.$$

If we apply this expansion dimension by dimension then we arrive at the claimed form with the proof technique of Lemma 6.3. □

We note that representation *(ii)* from Proposition 6.5 is equivalent to the following form which was used in [19, Equation (3.13) in Theorem 3.5]

$$f(\boldsymbol{x}) = \sum_{\mathfrak{w}\subseteq\{1,\ldots,d\}} \sum_{\mathfrak{v}\subseteq\mathfrak{w}} (-1)^{|\mathfrak{v}|} \sum_{\substack{\boldsymbol{\tau}_{\mathfrak{w}\setminus\mathfrak{v}}\in\{1,\ldots,\alpha\}^{|\mathfrak{w}\setminus\mathfrak{v}|} \\ \tau_j=\alpha \text{ for } j\in\mathfrak{v} \\ \tau_j=0 \text{ for } j\notin\mathfrak{w}}} \left[\prod_{j\in\mathfrak{w}\setminus\mathfrak{v}} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right]$$

$$\int_{[0,1]^d} \left[\prod_{j\in\mathfrak{v}} \frac{\widetilde{B}_\alpha(x_j-y_j)}{\alpha!}\right] f^{(\boldsymbol{\tau})}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y},$$

by considering $j \notin \mathfrak{u}$ in the formula for *(ii)* to be $j \in \mathfrak{v}$ in this formula, and using $B_0(x)/0! = 1$ and we already used $\widetilde{B}_\alpha(x-y) = (-1)^\alpha \widetilde{B}_\alpha(y-x)$ in comparison with the original formula in [19]. The latter formula is advantageous if considering weighted spaces or when access to the ANOVA decomposition of f is required. The advantage of representation *(i)* over *(ii)* and the equivalent form just discussed is that *(i)* does not contain products with other Bernoulli polynomials when considering the term with $\prod_{j=1}^d \widetilde{B}_\alpha(x_j-y_j)/\alpha!$. This is useful when one needs to evaluate or bound each term of the expression, e.g., see Section 6.4.

Now we are ready to state the main result of this section.

Proposition 6.6. *Let $f \in \mathcal{H}(K^{\mathrm{Sob}}_{\alpha,d})$, then we have*

$$F(\boldsymbol{y}) = \left\langle f, K^{\mathrm{Kor}}_{\alpha,d}(\cdot,\boldsymbol{y})\right\rangle_{K^{\mathrm{Sob}}_{\alpha,d}},$$

where F is the function obtained by applying the Bernoulli polynomial method (6.9) *to f. Therefore, the transformed function F does not have a*

bigger norm than the original function f,

$$\|F\|_{K^{\mathrm{Kor}}_{\alpha,d}} = \|F\|_{K^{\mathrm{Sob}}_{\alpha,d}} \le \|f\|_{K^{\mathrm{Sob}}_{\alpha,d}}.$$

Proof. For this proof, we introduce a shorthand notation for the result of Lemma 6.4. Instead of writing $\boldsymbol{\tau}_{\mathfrak{u}} \in \{0,\dots,\alpha\}^{|\mathfrak{u}|}$ we write the multiindex $\boldsymbol{\tau} \in \{0,\dots,\alpha,\overline{0}\}^d$ where now whenever $\tau_j = \overline{0}$ this would be $j \notin \mathfrak{u}$ in the original notation. In other words we use

$$\sum_{\boldsymbol{\tau}\in\{0,\dots,\alpha,\overline{0}\}^d} A(\boldsymbol{\tau}) = \sum_{\mathfrak{u}\subseteq\{1,\dots,d\}} \sum_{\substack{\boldsymbol{\tau}_{\mathfrak{u}}\in\{0,\dots,\alpha\}^{|\mathfrak{u}|}\\ \tau_j=\overline{0} \text{ for } j\notin\mathfrak{u}}} A(\boldsymbol{\tau}),$$

for any function A defined on $\{0,\dots,\alpha,\overline{0}\}^d$. Lemma 6.4 can then be written as

$$\sum_{\substack{\boldsymbol{\tau}\in\{0,\dots,\alpha,\overline{0}\}^d\\ \overline{\mathfrak{u}}:=\{j:\tau_j=\overline{0}\}}} (-1)^{|\overline{\mathfrak{u}}|} \left[\prod_{j=1}^{d} \frac{B_{\tau_j}(x_j)}{\tau_j!}\right] \int_{[0,1]^{d-|\overline{\mathfrak{u}}|}} f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\overline{\mathfrak{u}}}$$

$$= \int_{[0,1]^d} \left[\prod_{j=1}^{d} \frac{\widetilde{B}_\alpha(x_j - y_j)}{\alpha!}\right] f^{(\alpha,\dots,\alpha)}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y},$$

where for the derivatives and $B_\tau(x)/\tau!$ we interpret $\overline{0} = 0$. In the following we will use this with $\boldsymbol{x}$ and $\boldsymbol{y}$ interchanged and where the derivatives of order α in the right hand side appear for a certain subset $\mathfrak{w} \subseteq \{1,\dots,d\}$ such that the left hand side is modified accordingly. More precisely, for given $\mathfrak{w} \subseteq \{1,\dots,d\}$ and $\boldsymbol{\tau}_{-\mathfrak{w}} = \boldsymbol{\tau}_{\{1,\dots,d\}\setminus\mathfrak{w}} = \mathbf{0}$ we will use

$$\sum_{\substack{\boldsymbol{\tau}_{\mathfrak{w}}\in\{0,\dots,\alpha,\overline{0}\}^{|\mathfrak{w}|}\\ \overline{\mathfrak{u}}:=\{j\in\mathfrak{w}:\tau_j=\overline{0}\}\\ \tau_j=0 \text{ for } j\notin\mathfrak{w}}} (-1)^{|\overline{\mathfrak{u}}|} \left[\prod_{j\in\mathfrak{w}} \frac{B_{\tau_j}(y_j)}{\tau_j!}\right] \int_{[0,1]^{|\mathfrak{w}|-|\overline{\mathfrak{u}}|}} f^{(\boldsymbol{\tau})}(\boldsymbol{y}_{\mathfrak{w}\setminus\overline{\mathfrak{u}}}, \boldsymbol{y}_{\overline{\mathfrak{u}}}, \boldsymbol{x}_{-\mathfrak{w}})\,\mathrm{d}\boldsymbol{y}_{\mathfrak{w}\setminus\overline{\mathfrak{u}}}$$

$$= \int_{[0,1]^{|\mathfrak{w}|}} \left[\prod_{j\in\mathfrak{w}} \frac{\widetilde{B}_\alpha(y_j - x_j)}{\alpha!}\right] f^{(\boldsymbol{\alpha}_{\mathfrak{w}},\mathbf{0}_{-\mathfrak{w}})}(\boldsymbol{x}_{\mathfrak{w}}, \boldsymbol{x}_{-\mathfrak{w}})\,\mathrm{d}\boldsymbol{x}_{\mathfrak{w}},$$

which, after integrating on both sides over $\boldsymbol{x}_{-\mathfrak{w}}$ gives

$$\sum_{\substack{\boldsymbol{\tau}_{\mathfrak{w}} \in \{0,\ldots,\alpha,\overline{0}\}^{|\mathfrak{w}|} \\ \overline{\mathfrak{u}} := \{j \in \mathfrak{w} : \tau_j = \overline{0}\} \\ \tau_j = 0 \text{ for } j \notin \mathfrak{w}}} (-1)^{|\overline{\mathfrak{u}}|} \left[\prod_{j \in \mathfrak{w}} \frac{B_{\tau_j}(y_j)}{\tau_j!}\right] \int_{[0,1]^{d-|\overline{\mathfrak{u}}|}} f^{(\boldsymbol{\tau})}(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}_{-\overline{\mathfrak{u}}}$$

$$= \int_{[0,1]^d} \left[\prod_{j \in \mathfrak{w}} \frac{\widetilde{B}_\alpha(y_j - x_j)}{\alpha!}\right] f^{(\boldsymbol{\alpha}_{\mathfrak{w}}, \boldsymbol{0}_{-\mathfrak{w}})}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x},$$

where $-\overline{\mathfrak{u}} = \{1, \ldots, d\} \setminus \overline{\mathfrak{u}}$. Note that the products over $j \in \mathfrak{w}$ can as well be written as $j \in \{1, \ldots, d\}$.

We now start by expanding the inner product of the Sobolev space $K^{\mathrm{Sob}}_{\alpha,d}$ for f and the kernel of the Korobov space $K^{\mathrm{Kor}}_{\alpha,d}$ with one input fixed to $\boldsymbol{y}$:

$$\left\langle f, K^{\mathrm{Kor}}_{\alpha,d}(\cdot, \boldsymbol{y}) \right\rangle_{K^{\mathrm{Sob}}_{\alpha,d}} = \sum_{\substack{\boldsymbol{\tau} \in \{0,\ldots,\alpha\}^d \\ \mathfrak{w} := \{j : \tau_j = \alpha\}}} \int_{[0,1]^{|\mathfrak{w}|}} \left(\int_{[0,1]^{d-|\mathfrak{w}|}} f^{(\boldsymbol{\tau})}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_{-\mathfrak{w}}\right)$$

$$\left(\int_{[0,1]^{d-|\mathfrak{w}|}} K^{\mathrm{Kor}\,(\boldsymbol{\tau})}_{\alpha,d}(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x}_{-\mathfrak{w}}\right) \mathrm{d}\boldsymbol{x}_{\mathfrak{w}},$$

here we have, cf. (6.4) and (6.1),

$$K^{\mathrm{Kor}\,(\boldsymbol{\tau})}_{\alpha,d}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j=1}^{d} \left(\delta_{\tau_j,0} + (-1)^{\alpha+1} \frac{\widetilde{B}_{2\alpha-\tau_j}(x_j - y_j)}{(2\alpha - \tau_j)!}\right)$$

hence integrating with respect to x_j with $j \in -\mathfrak{w}$, i.e., $\tau_j \neq \alpha$, vanishes unless $\tau_j = 0$. Therefore only $\tau_j = 0$ or $\tau_j = \alpha$ remain and we obtain

$$
\begin{aligned}
&\left\langle f, K_{\alpha,d}^{\mathrm{Kor}}(\cdot,\boldsymbol{y})\right\rangle_{K_{\alpha,d}^{\mathrm{Sob}}} \\
&\quad= \sum_{\substack{\boldsymbol{\tau}\in\{0,\alpha\}^d \\ \mathfrak{w}:=\{j:\tau_j=\alpha\}}} \int_{[0,1]^d} f^{(\boldsymbol{\tau})}(\boldsymbol{x}) \prod_{j\in\mathfrak{w}} \left((-1)^{\alpha+1}\frac{\widetilde{B}_\alpha(x_j-y_j)}{\alpha!}\right) \mathrm{d}\boldsymbol{x} \\
&\quad= \sum_{\mathfrak{w}\subseteq\{1,\ldots,d\}} (-1)^{|\mathfrak{w}|} \int_{[0,1]^d} \left[\prod_{j\in\mathfrak{w}} \frac{\widetilde{B}_\alpha(y_j-x_j)}{\alpha!}\right] f^{(\boldsymbol{\alpha}_{\mathfrak{w}},\mathbf{0}_{-\mathfrak{w}})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&\quad= \sum_{\mathfrak{w}\subseteq\{1,\ldots,d\}} (-1)^{|\mathfrak{w}|} \sum_{\substack{\boldsymbol{\tau}_{\mathfrak{w}}\in\{0,\ldots,\alpha,\overline{0}\}^{|\mathfrak{w}|} \\ \overline{\mathfrak{u}}:=\{j\in\mathfrak{w}:\tau_j=\overline{0}\} \\ \tau_j=0 \text{ for } j\notin\mathfrak{w}}} (-1)^{|\overline{\mathfrak{u}}|} \left[\prod_{j\in\mathfrak{w}} \frac{B_{\tau_j}(y_j)}{\tau_j!}\right] \int_{[0,1]^{d-|\overline{\mathfrak{u}}|}} f^{(\boldsymbol{\tau})}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}_{-\overline{\mathfrak{u}}} \\
&\quad= \sum_{\substack{\boldsymbol{\tau}\in\{1,\ldots,\alpha,\overline{0}\}^d \\ \overline{\mathfrak{u}}:=\{j:\tau_j=\overline{0}\}}} (-1)^{d-|\overline{\mathfrak{u}}|} \left[\prod_{j=1}^{d} \frac{B_{\tau_j}(y_j)}{\tau_j!}\right] \int_{[0,1]^{d-|\overline{\mathfrak{u}}|}} f^{(\boldsymbol{\tau})}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}_{-\overline{\mathfrak{u}}}, \qquad (6.14)
\end{aligned}
$$

where we applied Lemma 6.4 as laid out in the beginning of the proof and where in the last step we used the inclusion-exclusion principle

$$
\begin{aligned}
\sum_{\boldsymbol{\tau}\in\{1,\ldots,\alpha,\overline{0}\}^d} A(\boldsymbol{\tau}) &= \sum_{\boldsymbol{\tau}\in\{0,\ldots,\alpha,\overline{0}\}^d} A(\boldsymbol{\tau}) - \sum_{j=1}^{d} \sum_{\substack{\boldsymbol{\tau}_{-\{j\}}\in\{0,\ldots,\alpha,\overline{0}\}^{d-1} \\ \tau_j=0}} A(\boldsymbol{\tau}) + \cdots + (-1)^d \sum_{\boldsymbol{\tau}=\mathbf{0}} A(\boldsymbol{\tau}) \\
&= \sum_{\mathfrak{w}\subseteq\{1,\ldots,d\}} (-1)^{d-|\mathfrak{w}|} \sum_{\substack{\boldsymbol{\tau}_{\mathfrak{w}}\in\{0,\ldots,\alpha,\overline{0}\}^{|\mathfrak{w}|} \\ \tau_j=0 \text{ for } j\notin\mathfrak{w}}} A(\boldsymbol{\tau}),
\end{aligned}
$$

for any function $A(\boldsymbol{\tau})$ defined for multiindices $\boldsymbol{\tau} \in \{0,\ldots,\alpha,\overline{0}\}^d$ and we took

$$
A(\boldsymbol{\tau}) = (-1)^{|\overline{\mathfrak{u}}|} \left[\prod_{j\in\mathfrak{w}} \frac{B_{\tau_j}(y_j)}{\tau_j!}\right] \int_{[0,1]^{d-|\overline{\mathfrak{u}}|}} f^{(\boldsymbol{\tau})}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}_{-\overline{\mathfrak{u}}},
$$

with $\overline{\mathfrak{u}} = \{j : \tau_j = \overline{0}\}$. We conclude the proof by noting that (6.14) and (6.9) are the same by our definition of the symbol $\overline{0}$. □

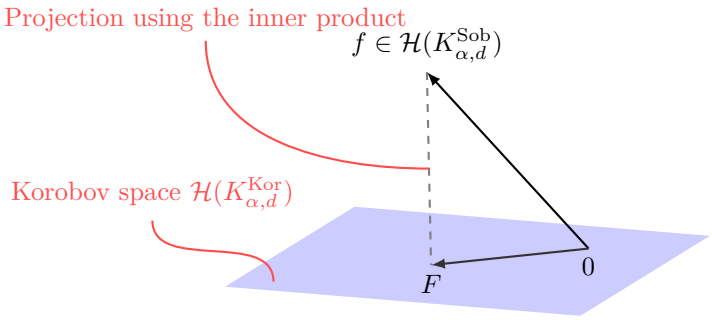

Figure 6.1: Interpretation of the Bernoulli polynomial method in terms of reproducing kernels.

We remark that in [40, 94] it is proved that the norm of the integrand does not increase by this Bernoulli polynomial method but in a different space setting:

$$E_\alpha := \left\{ f \in L_2([0,1]^d) : |\widehat{f}(\boldsymbol{h})| \le \prod_{j=1}^{d} \frac{C}{\max(1, |h_j|^{\alpha+1})} \right\},$$

where C is a positive constant. Those spaces are embedded in our Korobov spaces with the same smoothness α, which means we have more general results.

6.3.2 The Bernoulli polynomial method on a box

We first generalize Lemma 6.4 to a general box $[\boldsymbol{a}, \boldsymbol{b}]$.

Lemma 6.7. *For a function f having mixed partial derivatives up to order $\alpha-1$ in each variable which are absolutely continuous and of which the derivatives with the highest order α are integrable, the following holds for all $\boldsymbol{x} \in [\boldsymbol{a}, \boldsymbol{b}]$:*

$$\sum_{\mathfrak{u} \subseteq \{1,\dots,d\}} (-1)^{d-|\mathfrak{u}|} \sum_{\substack{\boldsymbol{\tau}_\mathfrak{u} \in \{0,\dots,\alpha\}^{|\mathfrak{u}|} \\ \tau_j = 0 \text{ for } j \notin \mathfrak{u}}} \left[\prod_{j \in \mathfrak{u}} \frac{B_{\tau_j}^{[a_j,b_j]}(x_j)}{\tau_j!} \right] \int_{[\boldsymbol{a}_\mathfrak{u}, \boldsymbol{b}_\mathfrak{u}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_\mathfrak{u}$$

$$= \int_{[\boldsymbol{a},\boldsymbol{b}]} \left[\prod_{j=1}^{d} \frac{\widetilde{B}_\alpha^{[a_j,b_j]}(x_j - y_j + a_j)}{\alpha!} \right] f^{(\alpha,\dots,\alpha)}(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}.$$

Proof. The proof is similar to the proof of Lemma 6.4. We show the one-dimensional results which follow from the properties of the scaled Bernoulli

polynomials from Section 6.2.1. Define

$$F_\tau^{[a,b]}(x) := \int_a^b f^{(\tau)}(y) \frac{\widetilde{B}_\tau^{[a,b]}(x-y+a)}{\tau!}\,\mathrm{d}y,$$

then, as is the case for the interval $[0,1]$, we here obtain, for $\tau = 1,\dots,\alpha-1$, using integration by parts,

$$F_{\tau+1}^{[a,b]}(x) - F_\tau^{[a,b]}(x) = \frac{\widetilde{B}_{\tau+1}^{[a,b]}(x)}{(\tau+1)!}\int_a^b f^{(\tau+1)}(y)\,\mathrm{d}y.$$

We also obtain for $F_1^{[a,b]}$ using integration by parts with the terms interchanged and for $x \in [a,b]$

$$\begin{aligned}
&\int_a^b \widetilde{B}_1^{[a,b]}(x-y+a) f'(y)\,\mathrm{d}y \\
&= \int_a^{x^-} \widetilde{B}_1^{[a,b]}(x-y+a) f'(y)\,\mathrm{d}y + \int_{x^+}^b \widetilde{B}_1^{[a,b]}(x-y+a) f'(y)\,\mathrm{d}y \\
&= \int_a^{x^-} B_1^{[a,b]}(x-y+a) f'(y)\,\mathrm{d}y + \int_{x^+}^b B_1^{[a,b]}(x-y+b) f'(y)\,\mathrm{d}y \\
&= \Big[B_1^{[a,b]}(x-y+a) f(y)\Big]_{y=a}^{x^-} + \int_a^{x^-} \frac{f(y)}{b-a}\,\mathrm{d}y \\
&\quad + \Big[B_1^{[a,b]}(x-y+b) f(y)\Big]_{y=x^+}^{b} + \int_{x^+}^b \frac{f(y)}{b-a}\,\mathrm{d}y \\
&= -f(x) + B_1^{[a,b]}(x)(f(b)-f(a)) + B_0^{[a,b]}(x)\int_a^b f(y)\,\mathrm{d}y.
\end{aligned}$$

Therefore we have the following representation for one-dimensional functions on $[a,b]$

$$f(x) = \sum_{\tau=0}^{\alpha} \frac{B_\tau^{[a,b]}(x)}{\tau!}\int_a^b f^{(\tau)}(y)\,\mathrm{d}y - \int_a^b \frac{\widetilde{B}_\alpha^{[a,b]}(x-y+a)}{\alpha!} f^{(\alpha)}(y)\,\mathrm{d}y. \quad (6.15)$$

Using the same induction argument as Lemma 6.4 the result for the multivariate case now follows. □

Using this result we can represent functions in $\mathcal{H}(K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]})$ similarly as those in $\mathcal{H}(K_{\alpha,d}^{\mathrm{Sob}})$, cf. Proposition 6.5.

Proposition 6.8. *For a function f on the box $[\boldsymbol{a},\boldsymbol{b}]$ having mixed partial derivatives up to order $\alpha-1$ in each variable which are absolutely continuous and of which the derivatives with the highest order α are integrable, we have the following equivalent representations*
(i) series representation in terms of Bernoulli polynomials with remainder term

$$f(\boldsymbol{x}) = \sum_{\emptyset\ne\mathfrak{u}\subseteq\{1,\dots,d\}} (-1)^{|\mathfrak{u}|+1} \sum_{\substack{\boldsymbol{\tau}_\mathfrak{u}\in\{0,\dots,\alpha\}^{|\mathfrak{u}|}\\ \tau_j=0 \text{ for } j\notin\mathfrak{u}}} \left[\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}^{[a_j,b_j]}(x_j)}{\tau_j!}\right] \int_{[\boldsymbol{a}_\mathfrak{u},\boldsymbol{b}_\mathfrak{u}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x}_{-\mathfrak{u}},\boldsymbol{y}_\mathfrak{u})\,\mathrm{d}\boldsymbol{y}_\mathfrak{u}$$

$$+ (-1)^d \int_{[\boldsymbol{a},\boldsymbol{b}]} \left[\prod_{j=1}^d \frac{\widetilde{B}_\alpha^{[a_j,b_j]}(x_j-y_j+a_j)}{\alpha!}\right] f^{(\alpha,\dots,\alpha)}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y},$$

(ii) series representation in terms of Bernoulli polynomials with mixed remainder terms

$$f(\boldsymbol{x}) = \sum_{\mathfrak{u}\subseteq\{1,\dots,d\}} (-1)^{d-|\mathfrak{u}|} \sum_{\substack{\boldsymbol{\tau}_\mathfrak{u}\in\{0,\dots,\alpha\}^{|\mathfrak{u}|}\\ \tau_j=\alpha \text{ for } j\notin\mathfrak{u}}}$$

$$\left[\prod_{j\in\mathfrak{u}} \frac{B_{\tau_j}^{[a_j,b_j]}(x_j)}{\tau_j!}\right] \int_{[\boldsymbol{a},\boldsymbol{b}]} \left[\prod_{j\notin\mathfrak{u}} \frac{\widetilde{B}_\alpha^{[a_j,b_j]}(x_j-y_j+a_j)}{\alpha!}\right] f^{(\boldsymbol{\tau})}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}.$$

Proof. The proof is similar to the proof of Proposition 6.5, making use of the one-dimensional representation (6.15). □

We are now finally in a position to show that $K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}$ is in fact the reproducing kernel of $\mathcal{H}(K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]})$ by making use of the representations just given.

Proposition 6.9. *Then function $K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}$, given in* (6.7)*, is the reproducing kernel corresponding to the inner product* (6.6) *of the space $\mathcal{H}(K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]})$. That is*

$$\langle f, K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}(\cdot,\boldsymbol{y})\rangle_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}} = f(\boldsymbol{y}) \qquad \forall \boldsymbol{y}\in[\boldsymbol{a},\boldsymbol{b}] \text{ and } \forall f\in\mathcal{H}(K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}).$$

Proof. We need to calculate

$$\langle f, K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}(\cdot,\boldsymbol{y})\rangle_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}} = \sum_{\substack{\boldsymbol{\tau}\in\{0,\dots,\alpha\}^d \\ \mathfrak{w}:=\{j:\tau_j=\alpha\}}} \int_{[\boldsymbol{a}_{\mathfrak{w}},\boldsymbol{b}_{\mathfrak{w}}]} \left(\int_{[\boldsymbol{a}_{-\mathfrak{w}},\boldsymbol{b}_{-\mathfrak{w}}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{w}} \right)$$

$$\left(\int_{[\boldsymbol{a}_{-\mathfrak{w}},\boldsymbol{b}_{-\mathfrak{w}}]} K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]^{(\boldsymbol{\tau})}}(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{w}} \right) \mathrm{d}\boldsymbol{x}_{\mathfrak{w}}.$$

For $\boldsymbol{\tau} \in \{0,\dots,\alpha\}^d$, we have

$$K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]^{(\boldsymbol{\tau})}}(\boldsymbol{x},\boldsymbol{y}) = \prod_{j=1}^{d} \left(\sum_{\tau'=\tau_j}^{\alpha-1} \frac{B_{\tau'-\tau_j}^{[a_j,b_j]}(x_j)}{(\tau'-\tau_j)!} \frac{B_{\tau'}^{[a_j,b_j]}(y_j)}{\tau'!} \right.$$

$$\left. + (b_j-a_j)\frac{B_{\alpha-\tau_j}^{[a_j,b_j]}(x_j)}{(\alpha-\tau_j)!}\frac{B_{\alpha}^{[a_j,b_j]}(y_j)}{\alpha!} + (-1)^{\alpha+1}\frac{\widetilde{B}_{2\alpha-\tau_j}^{[a_j,b_j]}(x_j-y_j+a_j)}{(2\alpha-\tau_j)!} \right).$$

Therefore, integrating over x_j when $\tau_j \neq \alpha$ gives

$$\int_{[\boldsymbol{a}_{-\mathfrak{w}},\boldsymbol{b}_{-\mathfrak{w}}]} K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]^{(\boldsymbol{\tau})}}(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{w}}$$

$$= \left[\prod_{j\notin\mathfrak{w}} \frac{B_{\tau_j}^{[a_j,b_j]}(y_j)}{\tau_j!} \right] \prod_{j\in\mathfrak{w}} \left(\frac{B_{\alpha}^{[a_j,b_j]}(y_j)}{\alpha!} + (-1)^{\alpha+1}\frac{\widetilde{B}_{\alpha}^{[a_j,b_j]}(x_j-y_j+a_j)}{\alpha!} \right).$$

It follows that

$$\langle f, K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}(\cdot,\boldsymbol{y})\rangle_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}} = \sum_{\substack{\boldsymbol{\tau}\in\{0,\dots,\alpha\}^d \\ \mathfrak{w}:=\{j:\tau_j=\alpha\}}} \int_{[\boldsymbol{a}_{\mathfrak{w}},\boldsymbol{b}_{\mathfrak{w}}]} \left(\int_{[\boldsymbol{a}_{-\mathfrak{w}},\boldsymbol{b}_{-\mathfrak{w}}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{w}} \right)$$

$$\left[\prod_{j\notin\mathfrak{w}} \frac{B_{\tau_j}^{[a_j,b_j]}(y_j)}{\tau_j!} \right] \prod_{j\in\mathfrak{w}} \left(\frac{B_{\alpha}^{[a_j,b_j]}(y_j)}{\alpha!} + (-1)^{\alpha+1}\frac{\widetilde{B}_{\alpha}^{[a_j,b_j]}(x_j-y_j+a_j)}{\alpha!} \right) \mathrm{d}\boldsymbol{x}_{\mathfrak{w}}$$

which is equivalent to representation *(ii)* in Proposition 6.8. □

We can now state the Bernoulli polynomial method on a box $[\boldsymbol{a},\boldsymbol{b}]$ and also observe that this is an orthogonal projection, cf. Proposition 6.6.

Proposition 6.10. *Let $f \in \mathcal{H}(K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]})$, then the modified function $F^{[\boldsymbol{a},\boldsymbol{b}]} \in \mathcal{H}(K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]})$, obtained by the Bernoulli polynomial method on the box $[\boldsymbol{a},\boldsymbol{b}]$, is equivalent to the orthogonal projection on $\mathcal{H}(K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]})$, i.e.,*

$$F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x}) := f(\boldsymbol{x}) + \sum_{\substack{\mathbf{0}\neq\boldsymbol{\tau}\in\{0,\ldots,\alpha\}^d \\ \mathrm{u}:=\mathrm{supp}(\boldsymbol{\tau})}} (-1)^{|\mathrm{u}|} \left[\prod_{j\in\mathrm{u}} \frac{B_{\tau_j}^{[a_j,b_j]}(x_j)}{\tau_j!}\right] \int_{[\boldsymbol{a}_{\mathrm{u}},\boldsymbol{b}_{\mathrm{u}}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{\mathrm{u}}$$

$$= \langle f, K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}(\cdot,\boldsymbol{x})\rangle_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}},$$

with $\int_{[\boldsymbol{a},\boldsymbol{b}]} F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$. Furthermore

$$\|F^{[\boldsymbol{a},\boldsymbol{b}]}\|_{K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}} = \|F^{[\boldsymbol{a},\boldsymbol{b}]}\|_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}} \le \|f\|_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}}.$$

Proof. The proof is the equivalent of the proof of Proposition 6.6 but then making use of Lemma 6.7. □

6.4 Error analysis for integration over $\mathbb{R}^d$

In this section, the integration error for approximating an integral over $\mathbb{R}^d$ using a scaled lattice rule is shown. Our proposed algorithm is simple: first we truncate the domain to a box $[\boldsymbol{a},\boldsymbol{b}]$, and then we apply a lattice rule to compute the integral $\int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$. We assume that the function f and its derivatives decay fast enough, so that we can exploit the fact that the bigger box we truncate to is, the smaller the orthogonal complement of the projected function becomes. We start with the following lemma.

Lemma 6.11. *Let $\Lambda(\boldsymbol{z},n)$ be the nodes of an n-point lattice rule with generating vector $\boldsymbol{z}$. The integration error of using the scaled lattice rule, with nodes $\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}$,*

to integrate a function $f \in \mathcal{H}(K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]})$ *is bounded by*

$$\left|\frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right|$$

$$\le \|f\|_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}} \left(\prod_{j=1}^d \max(1, b_j - a_j)^{\alpha+1/2}\right) \left(\sum_{\mathbf{0}\ne\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{z},n)} [r_{\alpha,d}(\boldsymbol{h})]^{-2}\right)^{1/2}$$

$$+ \left|\frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n \left(F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]})\right)\right|,$$

where the term $(\sum_{\mathbf{0}\ne\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{z},n)} [r_{\alpha,d}(\boldsymbol{h})]^{-2})^{1/2}$ *is the worst-case error for a standard lattice rule on the unit cube* $[0,1]^d$*. There exists such lattice rule* $\Lambda(\boldsymbol{z}, n)$ *for prime* n*,*

$$\left(\sum_{\mathbf{0}\ne\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{z},n)} [r_{\alpha,d}(\boldsymbol{h})]^{-2}\right)^{1/2} \le C_{\alpha,d} \frac{(\log n)^{\alpha d}}{n^\alpha},$$

where $C_{\alpha,d}$ *is a constant only depends on* α *and* d*.*

Proof. We apply the triangle inequality and Proposition 6.10 to obtain

$$\left|\frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right|$$

$$\le \left|\frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - \int_{[\boldsymbol{a},\boldsymbol{b}]} F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right|$$

$$+ \left|\frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n \left(F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]})\right)\right|.$$

The result now follows from Proposition 6.2 and the fact that $F^{[\boldsymbol{a},\boldsymbol{b}]}$ is obtained by a projection such that $\|F^{[\boldsymbol{a},\boldsymbol{b}]}\|_{K_{\alpha,d}^{\mathrm{Kor},[\boldsymbol{a},\boldsymbol{b}]}} = \|F^{[\boldsymbol{a},\boldsymbol{b}]}\|_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}} \le \|f\|_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}}$. For the existence of such lattice rules, we refer to [80, Section 4]. □

The previous lemma tells us that the integration error of a non-periodic function on the box by a lattice rule is bounded by a sum of two terms: the integration

error of the projected function $F^{[\boldsymbol{a},\boldsymbol{b}]}$ and the approximation error of the difference $F^{[\boldsymbol{a},\boldsymbol{b}]} - f$. By asking the function f and the derivatives to decay, we will be able to obtain higher order convergence. We will show this in the next statement.

To make the statement for integration over $\mathbb{R}^d$ we need to know how the norm in $\mathcal{H}(K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]})$ of our integrand function behaves as the box grows larger. Since our inner product (6.6) is an "unanchored" one, meaning that we first integrate out all dimensions in which we take derivatives with lower degrees than the smoothness α of the space, we cannot guarantee that for certain boxes our norm becomes very small (e.g., some of these integrals might vanish for non-zero functions). This would be different with a more classical norm in which this inner integration is not present, and the outer L_2 integral integrates over all dimensions. In such a case the norm can only grow bigger when the box increases. To be able to state results for arbitrary boxes using our unanchored norm leads us to define a norm which captures the worst possible box as follows

$$\|f\|_{K_{\alpha,d}^{\mathrm{Sob},*}} := \sup_{[\boldsymbol{a},\boldsymbol{b}]\subset\mathbb{R}^d} \|f\|_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}}.$$

We can now make the main statement of the paper.

Theorem 6.12. *Suppose $f \in \mathcal{H}(K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]})$ for increasing boxes $[\boldsymbol{a},\boldsymbol{b}]$, i.e., $\|f\|_{K_{\alpha,d}^{\mathrm{Sob},*}} < \infty$. If additionally, f satisfies the following decay condition*

$$\|f\|_{\alpha,\beta} := \sup_{\substack{\boldsymbol{x}\in\mathbb{R}^d \\ \boldsymbol{\tau}\in\{0,\ldots,\alpha-1\}^d}} \left|\exp(\beta\|\boldsymbol{x}\|_\infty)\, f^{(\boldsymbol{\tau})}(\boldsymbol{x})\right| < \infty,$$

for some $\beta > 0$, then, by choosing the box proportional to n in the following way

$$[\boldsymbol{a},\boldsymbol{b}] = \left[-\frac{\alpha}{\beta}(\log n), \frac{\alpha}{\beta}(\log n)\right]^d,$$

and using a good lattice rule $\Lambda(\boldsymbol{z}, n)$ from Lemma 6.11, the integration error is bounded by

$$\left|\frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right|$$

$$\leq C_1 (\log n)^{\alpha d} \|f\|_{K_{\alpha,d}^{\mathrm{Sob},*}} \frac{(\log n)^{\alpha d}}{n^\alpha} + C_2 \|f\|_{\alpha,\beta} \frac{(\log n)^{\alpha d}}{n^\alpha},$$

where the constants C_1, C_2 are independent of n and f but depending on α, β and d. Also the truncation error is bounded by the smaller rate

$$\left| \int_{\mathbb{R}^d} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right| \le C_3 \|f\|_{\alpha,\beta} \frac{(\log n)^{d-1}}{n^\alpha},$$

where C_3 is a constant independent of n and f, such that the total error behaves like

$$\left| \frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - \int_{\mathbb{R}^d} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right| \le C(\|f\|_{K_{\alpha,d}^{\mathrm{Sob},*}} + \|f\|_{\alpha,\beta}) \frac{(\log n)^{2\alpha d}}{n^\alpha},$$

where the constant C is independent of n and f.

Proof. First we prove the bound on the integration error. Due to Lemma 6.11, the integration error of the projected function $F^{[\boldsymbol{a},\boldsymbol{b}]}$ is bounded by

$$\begin{aligned}
&\left| \frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - \int_{[\boldsymbol{a},\boldsymbol{b}]} F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right| \\
&\le \left(\sum_{\boldsymbol{0} \ne \boldsymbol{h} \in \Lambda^\perp(\boldsymbol{z},n)} [r_\alpha^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{h})]^{-2} \right)^{1/2} \left(\|f\|_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}} \right) \\
&\le \left(\sum_{\boldsymbol{0} \ne \boldsymbol{h} \in \Lambda^\perp(\boldsymbol{z},n)} [r_\alpha(\boldsymbol{h})]^{-2} \right)^{1/2} \left[\prod_{j=1}^d (b_j - a_j)^\alpha \right] \left(\|f\|_{K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]}} \right) \\
&\le \left(\sum_{\boldsymbol{0} \ne \boldsymbol{h} \in \Lambda^\perp(\boldsymbol{z},n)} [r_\alpha(\boldsymbol{h})]^{-2} \right)^{1/2} \left(\frac{2\alpha \log n}{\beta} \right)^{\alpha d} \left(\|f\|_{K_{\alpha,d}^{\mathrm{Sob},*}} \right).
\end{aligned}$$

For the approximation error on the non-periodic part, using Proposition 6.10 the following holds

$$\left|\frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n \left(F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) \right)\right|$$

$$= \frac{\prod_{j=1}^d (b_j - a_j)}{n}$$

$$\left| \sum_{i=1}^n \left(\sum_{\substack{\mathbf{0} \neq \boldsymbol{\tau} \in \{0,1,\ldots,\alpha\}^d \\ \mathrm{u} := \mathrm{supp}(\boldsymbol{\tau})}} \left((-1)^{\|\boldsymbol{\tau}\|_0} \left[\int_{[\boldsymbol{a}_\mathrm{u}, \boldsymbol{b}_\mathrm{u}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}_\mathrm{u} \right]_{\boldsymbol{x}_{-\mathrm{u}} = \boldsymbol{p}_{i,-\mathrm{u}}^{[\boldsymbol{a}_{-\mathrm{u}}, \boldsymbol{b}_{-\mathrm{u}}]}} \right.\right.\right.$$

$$\left.\left.\left. \left(\prod_{j=1}^d \frac{B_{\tau_j}^{[a_j,b_j]}(p_{i,j}^{[a_j,b_j]})}{\tau_j!} \right) \right) \right) \right|$$

then, by using the decay condition of the integrand f, we have $f^{(\boldsymbol{\tau} - \mathbf{1}_{-\mathrm{u}})}(\boldsymbol{x}) \leq \exp(-\beta \|\boldsymbol{x}\|_\infty) \|f\|_{\alpha,\beta} = n^{-\alpha} \, \|f\|_{\alpha,\beta}$ for $\boldsymbol{x}$ being at the boundary of $[\boldsymbol{a}, \boldsymbol{b}]$ with any $\boldsymbol{\tau}$ in the sum, and by using the upper bound on the Bernoulli polynomials (6.3), we obtain

$$\leq \frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n \left(\sum_{\substack{\mathbf{0} \neq \boldsymbol{\tau} \in \{0,1,\ldots,\alpha\}^d \\ \mathrm{u} := \mathrm{supp}(\boldsymbol{\tau})}} \left(2\|f\|_{\alpha,\beta} \frac{1}{n^\alpha} \right) \left(\prod_{j=1}^d (b_j - a_j)^{\tau_j - 1} \right) \right)$$

since these terms do not depend on the lattice points anymore,

$$\leq \sum_{\substack{\mathbf{0} \neq \boldsymbol{\tau} \in \{0,1,\ldots,\alpha\}^d \\ \mathrm{u} := \mathrm{supp}(\boldsymbol{\tau})}} \left(2\|f\|_{\alpha,\beta} \frac{1}{n^\alpha} \right) \left(\prod_{j=1}^d (b_j - a_j)^{\tau_j} \right)$$

by bounding $(b_j - a_j)^{\tau_j} \leq (b_j - a_j)^\alpha = (2\alpha(\log n)/\beta)^\alpha$, we have

$$\leq 2\|f\|_{\alpha,\beta} (\alpha + 1)^d \left(\frac{2\alpha}{\beta} \right)^{\alpha d} \frac{(\log n)^{\alpha d}}{n^\alpha}.$$

For the truncation error, we have

$$\begin{aligned}\left|\int_{\mathbb{R}^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| &\le \int_{\mathbb{R}^d\setminus[\boldsymbol{a},\boldsymbol{b}]} \|f\|_{\alpha,\beta}\exp(-\beta\|\boldsymbol{x}\|_\infty)\,\mathrm{d}\boldsymbol{x}\\ &\le 2^d\int_a^\infty \|f\|_{\alpha,\beta}\exp(-\beta r)\; d\; r^{d-1}\,\mathrm{d}r\\ &\le 2^d C''\; d\|f\|_{\alpha,\beta}\;\left(\frac{\alpha}{\beta}\right)^{d-1}\frac{(\log n)^{d-1}}{n^\alpha},\end{aligned}$$

where C'' is a constant only dependent on α and β. By summing up all three errors, we have proved the claim of the theorem. □

When the considered problem has slow decay, such as calculating expectations under heavy-tailed distributions (e.g., the Student-t distribution and the alpha-stable distribution with stable parameter smaller than 2), our method still works. Those problems especially occur in finance and extreme value theory. We note that for such distributions the classical Gaussian-type quadrature cannot be constructed for arbitrary number of nodes since such distributions do not have bounded moments for arbitrary order.

Theorem 6.13. *Suppose $f \in \mathcal{H}(K_{\alpha,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]})$ for increasing boxes $[\boldsymbol{a},\boldsymbol{b}]$, i.e., $\|f\|_{K_{\alpha,d}^{\mathrm{Sob},*}} < \infty$. If additionally, f satisfies the following decay condition*

$$\|f\|_{\alpha,\beta} := \sup_{\substack{\boldsymbol{x}\in\mathbb{R}^d\\ \boldsymbol{\tau}\in\{0,\ldots,\alpha-1\}^d}} \left|\|\boldsymbol{x}\|_\infty^\beta f^{(\boldsymbol{\tau})}(\boldsymbol{x})\right| < \infty,,$$

for some $\beta > \alpha d$, then, by choosing the box proportional to n in the following way

$$[\boldsymbol{a},\boldsymbol{b}] = \left[-n^{\alpha/\beta}, n^{\alpha/\beta}\right]^d,$$

and using a good lattice rule $\Lambda(\boldsymbol{z},n)$ from Lemma 6.11, the integration error is bounded by

$$\begin{aligned}&\left|\frac{\prod_{j=1}^d (b_j-a_j)}{n}\sum_{i=1}^n f(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) - \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right|\\ &\le C_1 n^{\alpha^2 d/\beta}\|f\|_{K_{\alpha,d}^{\mathrm{Sob},*}}\left(\sum_{\boldsymbol{0}\ne\boldsymbol{h}\in\Lambda^\perp(\boldsymbol{z},n)}[r_\alpha(\boldsymbol{h})]^{-2}\right)^{1/2} + C_2\|f\|_{\alpha,\beta}\; n^{-\alpha+(\alpha^2 d/\beta)},\end{aligned}$$

where the constants C_1, C_2 are independent of n and f but depending on α, β and d. Also the truncation error is bounded by the smaller rate

$$\left| \int_{\mathbb{R}^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{[\boldsymbol{a},\boldsymbol{b}]} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \right| \le C_3 \|f\|_{\alpha,\beta}\, n^{-\alpha+(\alpha d/\beta)},$$

where C_3 is a constant independent of n and f, such that the total error behaves like

$$\le C(\|f\|_{K^{\mathrm{Sob},*}_{\alpha,d}} + \|f\|_{\alpha,\beta})\, n^{-\alpha+(\alpha^2 d/\beta)} (\log n)^{\alpha d},$$

where the constant C is independent of n and f.

Proof. The proof is equivalent of the proof of Theorem 6.12 with polynomial decay and truncation region. □

6.5 Auxiliary results on the regular grid

If we use the regular grid, equispaced points on each coordinate, then we can calculate the integration error of the non-periodic part more explicitly. Let us denote the regular grid by

$$\Lambda_d^{\mathrm{reg}}(m) := \left\{0, \frac{1}{m}, \dots, \frac{m-1}{m}\right\}^d =: (\Lambda_1^{\mathrm{reg}}(m))^d,$$

where the number of points $n = m^d$.

Lemma 6.14. *Let $\|f\|_{K^{\mathrm{Sob},*}_{\alpha,d}} < \infty$. If we truncate the domain to $[\boldsymbol{a},\boldsymbol{b}]$ and apply the regular grid $\Lambda_d^{\mathrm{reg}}(m)$ for the quadrature, then the integration error for the non-periodic part has the following form*

$$\left| \frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{\boldsymbol{p} \in \Lambda_d^{\mathrm{reg}}(m)} \left(F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{p}^{[\boldsymbol{a},\boldsymbol{b}]}) - f(\boldsymbol{p}^{[\boldsymbol{a},\boldsymbol{b}]}) \right) \right|$$

$$= \left| \sum_{\substack{\boldsymbol{0} \ne \boldsymbol{\tau} \in \{0,1,\dots,\alpha\}^d \\ \mathfrak{u} := \{j : \tau_j = 0\}}} \left((-1)^{\|\boldsymbol{\tau}\|_0} \left(\frac{\prod_{j \in \mathfrak{u}} (b_j - a_j)}{m^{|\mathfrak{u}|}} \right.\right.\right.$$

$$\left.\left.\left. \sum_{\boldsymbol{p}_{\mathfrak{u}} \in \Lambda^{\mathrm{reg}}_{|\mathfrak{u}|}(m)} \left[\int_{[\boldsymbol{a}_{-\mathfrak{u}}, \boldsymbol{b}_{-\mathfrak{u}}]} f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{u}} \right]_{\boldsymbol{x}_{\mathfrak{u}} = \boldsymbol{p}_{\mathfrak{u}}} \right) \prod_{j \notin \mathfrak{u}} \left(\frac{(b_j - a_j)^{\tau_j} B_{\tau_j}(0)}{\tau_j!\, m^{\tau_j}} \right) \right) \right|.$$

Proof.

$$\left|\frac{\prod_{j=1}^{d}(b_j-a_j)}{n}\sum_{\boldsymbol{p}\in\Lambda_d^{\mathrm{reg}}(m)}\left(F^{[\boldsymbol{a},\boldsymbol{b}]}(\boldsymbol{p}^{[\boldsymbol{a},\boldsymbol{b}]})-f(\boldsymbol{p}^{[\boldsymbol{a},\boldsymbol{b}]})\right)\right|$$

$$=\left|\frac{\prod_{j=1}^{d}(b_j-a_j)}{n}\sum_{\boldsymbol{p}\in\Lambda_d^{\mathrm{reg}}(m)}\left(\sum_{\substack{\mathbf{0}\neq\boldsymbol{\tau}\in\{0,1,\ldots,\alpha\}^d\\ \mathfrak{u}:=\{j:\tau_j=0\}}}\left((-1)^{\|\boldsymbol{\tau}\|_0}\right.\right.\right.$$

$$\left.\left.\left.\left[\int_{[\boldsymbol{a}_{-\mathfrak{u}},\boldsymbol{b}_{-\mathfrak{u}}]}f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{u}}\right]_{\boldsymbol{x}_{\mathfrak{u}}=\boldsymbol{p}_{\mathfrak{u}}}\left(\prod_{j\notin\mathfrak{u}}\frac{B_{\tau_j}^{[a_j,b_j]}(p_j)}{\tau_j!}\right)\right)\right)\right|.$$

Now consider the terms for a fixed $\boldsymbol{\tau}$

$$\frac{\prod_{j=1}^{d}(b_j-a_j)}{n}\sum_{\boldsymbol{p}\in\Lambda_d^{\mathrm{reg}}(m)}\left[\int_{[\boldsymbol{a}_{-\mathfrak{u}},\boldsymbol{b}_{-\mathfrak{u}}]}f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{u}}\right]_{\boldsymbol{x}_{\mathfrak{u}}=\boldsymbol{p}_{\mathfrak{u}}}\left(\prod_{j\notin\mathfrak{u}}^{d}\frac{B_{\tau_j}^{[a_j,b_j]}(p_j)}{\tau_j!}\right)$$

$$=\left(\frac{\prod_{j\in\mathfrak{u}}(b_j-a_j)}{m^{|\mathfrak{u}|}}\sum_{\boldsymbol{p}_{\mathfrak{u}}\in\Lambda_{|\mathfrak{u}|}^{\mathrm{reg}}(m)}\left[\int_{[\boldsymbol{a}_{-\mathfrak{u}},\boldsymbol{b}_{-\mathfrak{u}}]}f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{u}}\right]_{\boldsymbol{x}_{\mathfrak{u}}=\boldsymbol{p}_{\mathfrak{u}}}\right)$$

$$\prod_{j\notin\mathfrak{u}}\left(\frac{b_j-a_j}{m}\sum_{p_j\in\Lambda_1^{\mathrm{reg}}(m)}\frac{B_{\tau_j}^{[a_j,b_j]}(p_j)}{\tau_j!}\right)$$

$$=\left(\frac{\prod_{j\in\mathfrak{u}}(b_j-a_j)}{m^{|\mathfrak{u}|}}\sum_{\boldsymbol{p}_{\mathfrak{u}}\in\Lambda_{|\mathfrak{u}|}^{\mathrm{reg}}(m)}\left[\int_{[\boldsymbol{a}_{-\mathfrak{u}},\boldsymbol{b}_{-\mathfrak{u}}]}f^{(\boldsymbol{\tau})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-\mathfrak{u}}\right]_{\boldsymbol{x}_{\mathfrak{u}}=\boldsymbol{p}_{\mathfrak{u}}}\right)$$

$$\prod_{j\notin\mathfrak{u}}\left(\frac{(b_j-a_j)^{\tau_j}B_{\tau_j}(0)}{\tau_j!\,m^{\tau_j}}\right),$$

where we used

$$\frac{1}{m}\sum_{p_j\in\Lambda_1^{\mathrm{reg}}(m)}B_{\tau_j}(p_j)=\frac{B_{\tau_j}(0)}{m^{\tau_j}},$$

which directly comes from the multiplication theorem of Bernoulli polynomials. □

By using this form, we can relax the condition of Theorem 6.12 to

$$\|f\|_{\boldsymbol{\alpha},\beta} := \sup_{\substack{\boldsymbol{x}\in\mathbb{R}^d \\ \boldsymbol{\tau}\in\{0,\dots,\alpha-1\}^d}} \left|\exp\left(\beta\|\boldsymbol{x}\|_\infty\left(1-\frac{\|\boldsymbol{\tau}\|_1+d-\|\boldsymbol{\tau}\|_0}{\alpha d}\right)\right) f^{(\boldsymbol{\tau})}(\boldsymbol{x})\right| < \infty.$$

6.6 Numerical experiments

In this section, we numerically demonstrate our method. Due to the decay condition which we require for the integrand, it is not possible to obtain the worst-case error which is often calculated numerically in QMC literature. First we consider an example against the logistic distribution for one and two-dimensional settings

$$g_1(\boldsymbol{x}) := \prod_{j=1}^{d} \frac{(4x_j + 10\cos^2(x_j))\ \exp((x_j-\mu_j)/s_j)}{(1+\exp((x_j-\mu_j)/s))^2 s_j}.$$

The exact values of $\int_{\mathbb{R}^d} g_1(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ and $\int_{[\boldsymbol{a},\boldsymbol{b}]} g_1(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ are possible to obtain by using symbolic computation. Using those values, we compute the truncation error, the cubature error in the truncated box, and the total error respectively in Figure 6.2. The parameters for the integrand function are chosen to be $\mu = 3, s = 4$ for the one-dimensional case and $(\mu_1,\mu_2) = (3,-3),\ (s_1,s_2) = (4,4)$ for the two-dimensional case. We choose $z = 1$ for the one-dimensional case and $\boldsymbol{z} = (1, 433461)$ for the two-dimensional case. We slightly take different numbers of points n_1 and n_2 between one and two-dimensional cases with keeping the same truncation domain $[-(\alpha/\beta)\log(n), (\alpha/\beta)\log(n)]^d$ where we set $\beta = 1/s$. For α we can take arbitrarily large and hence we fix it to be one of 1, 2, 3, 4. Then we set $n_2 = 2^\alpha n_1$ meaning we add more cubature points in the box. This is to see if the truncation error and the cubature error in the box converge with the same speed, because often it is difficult to see the asymptotic behavior for higher-order in multidimensional settings. Our aim here is to see if theoretical error behavior matches with this numerical example.

In Figure 6.2, we see the truncation error

$$\left|\int_{\mathbb{R}^d\setminus[\boldsymbol{a},\boldsymbol{b}]} g_1(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right|,$$

in the first row. This error converges with $n^{-\alpha}$ rate in $d = 1$ and $d = 2$ cases, which confirms our choice of the truncation box $[\boldsymbol{a},\boldsymbol{b}]$ is correct. In the second

row, we see the cubature error in the boxes

$$\left| \int_{[\boldsymbol{a},\boldsymbol{b}]} g_1(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n g_1(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) \right|.$$

For the case of $d = 1$, we see the expected order of convergence $n^{-\alpha}$. For the case of $d = 2$, we observe small "bumps" but soon we see the convergence with the same order. The third row shows the total error

$$\left| \int_{\mathbb{R}^d} g_1(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \frac{\prod_{j=1}^d (b_j - a_j)}{n} \sum_{i=1}^n g_1(\boldsymbol{p}_i^{[\boldsymbol{a},\boldsymbol{b}]}) \right|,$$

which is bounded by the sum of the truncation error and the cubature error. In any cases the total error converges with the rate of $n^{-\alpha}$.

Next, we consider an integrand that has only finite smoothness with known exact integral value. Particularly, we choose

$$g_2(\boldsymbol{x}) := \frac{\exp(-\|\boldsymbol{x}\|_2^2/2)}{\sqrt{2\pi}^d} \prod_{j=1}^d |x_j|^p.$$

Integrating this function means calculating a product of fractional p absolute moments of the standard normal distribution which have known exact values:

$$\int_{\mathbb{R}^d} g_2(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \mathbb{E}\left[\prod_{j=1}^d |X_j|^p\right] = \left(\frac{(\sqrt{2})^p \Gamma((p+1)/2)}{\sqrt{\pi}} M(-p/2, 1/2, 0)\right)^d,$$

where $\boldsymbol{X} = (X_1, \dots, X_d)$ follows the standard normal distribution, Γ is the gamma function and M is the Kummer's confluent hypergeometric function. We choose $d = 3$ and $p = 0.6, 1.6, 2.6$. For the parameters of lattice rules, we choose from $n = 2^4$ to $n = 2^{23}$ for the number of points, and the corresponding generating vectors are the same ones from Table 5.3 until $n = 2^{20}$, otherwise $(n, \boldsymbol{z}^\top) = (2^{21}, (1, 768165, 242667))$, $(2^{22}, (1, 1737355, 261247))$, $(2^{23}, (1, 3513381, 901073))$. We truncate the domain to $[\boldsymbol{a}, \boldsymbol{b}] = [-\sqrt{2p \log n}, \sqrt{2p \log n}]^d$ which can be obtained by the following argument similar to [22, Theorem 2]. Let

$$\|f\|_{\alpha,\beta} := \sup_{\substack{\boldsymbol{x} \in \mathbb{R}^d \\ \boldsymbol{\tau} \in \{0,\dots,\alpha-1\}^d}} \left| \exp(\beta \|\boldsymbol{x}\|_2^2) f^{(\boldsymbol{\tau})}(\boldsymbol{x}) \right| < \infty,$$

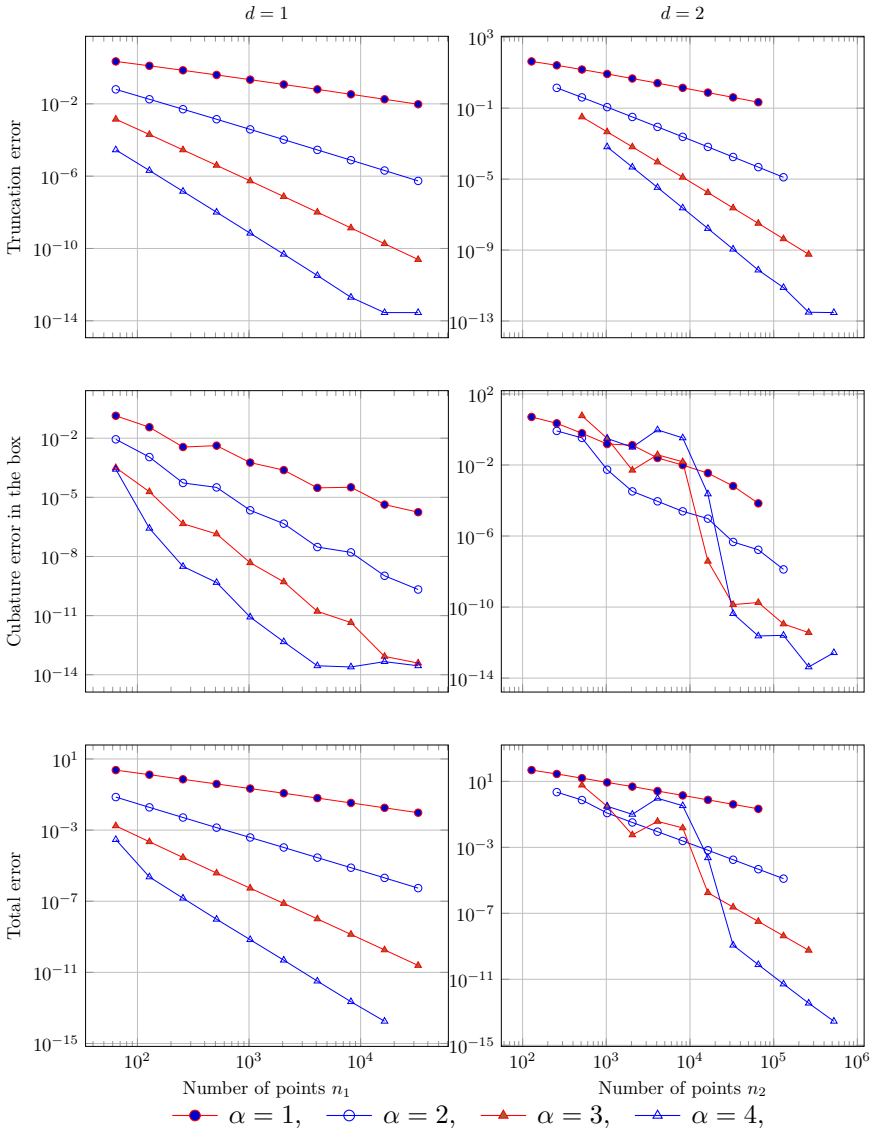

Figure 6.2: The error behavior for the integrand function g_1 with different smoothness α and dimension $d = 1, 2$.

for this normal distribution case. Then for any $\beta \le 1/2$, we have

$$\int_{\mathbb{R}^d \setminus [\boldsymbol{a},\boldsymbol{b}]} g_2(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \le \|g_2\|_{\alpha,1/2-\beta} \int_{\mathbb{R}^d \setminus [\boldsymbol{a},\boldsymbol{b}]} \frac{\exp(-\beta\|\boldsymbol{x}\|_2^2)}{\sqrt{2\pi}^d}\,\mathrm{d}\boldsymbol{x}$$
$$\le \|g_2\|_{\alpha,1/2-\beta} 2^d \int_{[0,\infty]^d \setminus [0,b]^d} \frac{\exp(-\beta\|\boldsymbol{x}\|_2^2)}{\sqrt{2\pi}^d}\,\mathrm{d}\boldsymbol{x}.$$

We have also

$$\int_{[0,\infty]^d \setminus [0,b]^d} \frac{\exp(-\beta\|\boldsymbol{x}\|_2^2)}{\sqrt{2\pi}^d}\,\mathrm{d}\boldsymbol{x}$$
$$= \int_{[0,\infty]^d} \frac{\exp(-\beta\|\boldsymbol{x}\|_2^2)}{\sqrt{2\pi}^d}\,\mathrm{d}\boldsymbol{x} - \int_{[0,b]^d} \frac{\exp(-\beta\|\boldsymbol{x}\|_2^2)}{\sqrt{2\pi}^d}\,\mathrm{d}\boldsymbol{x}$$
$$= \left(\int_{[0,\infty]} \frac{\exp(-\beta x^2)}{\sqrt{2\pi}}\,\mathrm{d}x\right)^d - \left(\int_{[0,b]} \frac{\exp(-\beta x^2)}{\sqrt{2\pi}}\,\mathrm{d}x\right)^d$$
$$= \left(\frac{1}{2}\sqrt{\frac{1}{2\beta}}\right)^d - \left(\frac{1}{2}\sqrt{\frac{1}{2\beta}}\right)^d \left(\frac{2\sqrt{\beta}}{\sqrt{\pi}} \int_0^b \exp(-\beta x^2)\,\mathrm{d}x\right)^d$$
$$= \left(\frac{1}{2}\sqrt{\frac{1}{2\beta}}\right)^d \left(1 - \left(\frac{2\sqrt{\beta}}{\sqrt{\pi}} \int_0^b \exp(-\beta x^2)\,\mathrm{d}x\right)^d\right)$$
$$\le \left(\frac{1}{2}\sqrt{\frac{1}{2\beta}}\right)^d \left(1 - (1 - \exp(-\beta b^2))^d\right)$$
$$\le \left(\frac{1}{2}\sqrt{\frac{1}{2\beta}}\right)^d d\ \exp(-\beta b^2).$$

Hence, by taking $b = \sqrt{(\alpha/\beta)\log n}$ for arbitrary β close to $1/2$, we can obtain the convergence rate of $n^{-\alpha}$ for the truncation error.

Using the exact values of the integrand, we plot the absolute error in Figure 6.3.

The integrand function g_2 is at least smooth enough for obtaining $\lfloor p \rfloor$ order convergence. As expected, we see convergence slower than first order for $p = 0.6$, which is not in $\mathcal{H}(K_{1,d}^{\mathrm{Sob},[\boldsymbol{a},\boldsymbol{b}]})$ for any box containing the origin. For $p = 1.6$ we see convergence faster than the first order but not more than the second order. In case of $p = 2.6$, we also see convergence faster than the second order. Making

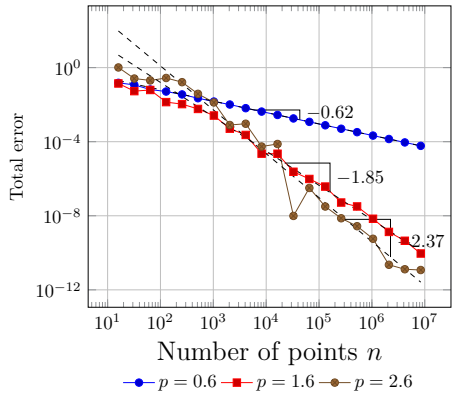

Figure 6.3: The absolute error of our cubature method for the function g.

use of fractional derivatives can possibly explain the order of convergence which we are seeing, because the integrand function can be p-smooth.

6.7 Chapter overview

Lattice rules for integration over $\mathbb{R}^d$ was considered. We derived explicit conditions where lattice rules can obtain higher order convergence. In our new strategy of the error analysis, the integrand is divided into a periodic part and a non-periodic part by orthogonal projection. The true value of the integral of the periodic part is same as that of the original integrand. Then the integration error of the periodic part is bounded by smoothness condition, and the error of the non-periodic part is bounded by the decay condition. This decay condition also helps to bound the truncation error and thus the total error of the cubature is bounded by $\mathcal{O}((\log n)^{2\alpha d}/n^{\alpha})$ with n function evaluations. The numerical examples confirmed the theory and showed different convergence rates depending on the smoothness of the integrand function.

We considered an unweighted setting, meaning that there is no information on the importance of a dimension for a given integrand function. It might be possible to consider such weight structure that allows us to avoid the curse of dimensionality even using the truncation strategy. However, this is out of the scope of the present paper and remains as future work.

Chapter 7

Conclusions

In this chapter, we conclude this thesis with a summary and an outlook on possible future works.

7.1 Summary

- In Chapter 3, we proposed a new numerical method which is a combination of pseudospectral Fourier methods on lattice point sets and exponential operator splitting methods for solving TDSEs. For the second order exponential splitting, which is called Strang splitting, we proved error bounds with explicit regularity conditions. We further compared our method with a sparse grid-based method [26], and showed that our method is more accurate and consistent with the theoretical results in high-dimensional settings. Compared with the method in [26], one notable advantage of our method is that Fourier transforms on lattices are always unitary whereas the sparse grid-based method is not. Since Strang splitting requires to use the Fourier transform on each time-step several times, this non-unitarity theoretically causes the exponential error growth.
- In Chapter 4, we obtained higher-order convergence of the time discretization error using higher-order operator splitting combined with the pseudospectral Fourier method on lattices. We numerically demonstrated our method with sixth and eighth order exponential splitting in two-dimensional to eight-dimensional settings, and we obtained the expected error convergence.

- In Chapter 5, we applied the above method for the nonlinear time-dependent Schrödinger equation by following [90, 91]. We conducted numerical experiments for a three-dimensional problem which is relevant to Bose–Einstein condensates in quantum physics. The numerical results confirmed the higher-order convergence in time. We gave a theoretical error bound given that a conjecture on lattice points holds. We validated this conjecture numerically.

- In Chapter 6, we proposed and analysed a new numerical integration method on $\mathbb{R}^d$. For analysing the cubature error, we gave a new interpretation of the Bernoulli polynomial method. By using this interpretation we decomposed the cubature error into two parts; a periodic part which is in the Korobov space, and a non-periodic part. To handle the error expression of the non-periodic part, we derived explicit decay conditions for obtaining higher-order error convergence for finitely smooth functions which are in unanchored Sobolev spaces.

7.2 Future works

Our contributions above can be applied to other problems. We list possible future work below.

- The pseudospectral method on lattices combined with operator splitting can be used for different problems such as time-dependent Dirac equations [6] and seismic wave propagation [37]. We expect that lattice-based methods work where the regular grid is applicable and one can make use of the flexibility of the lattice structure.

- Using lattice points for pseudospectral methods with different basis functions is another interesting idea. Using tent-transformed lattice points, function approximation with half-period cosine series is studied in [43]. Therein, approximating functions with multivariate Chebyshev series is also studied where an explicit CBC construction is given. Since the pseudospectral Chebyshev method is a strong tool for solving different types of PDEs, this method may lead us to solve such problems in high-dimensional settings.

- By using the analysis strategy used in Chapter 6, it can also be interesting to consider projection between different spaces where the reproducing kernel of the smaller space is explicitly known. Also, by using the orthogonal projection, function approximation over $\mathbb{R}^d$ can be considered.

Bibliography

[1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Vol. 55. National Bureau of Standards Applied Mathematics Series. For sale by the Superintendent of Documents, U.S. Government Printing Office Washington, D.C., 1964, pp. xiv+1046.

[2] N. Aronszajn. "Theory of reproducing kernels". *Transactions of the American Mathematical Society* 68 (1950), pp. 337–404.

[3] G. Avila, J. Oettershagen, and T. Carrington Jr. "Comparing nested sequences of Leja and pseudogauss points to interpolate in 1D and solve the Schroedinger equation in 9D". *Sparse grids and applications—Miami 2016*. Vol. 123. Lect. Notes Comput. Sci. Eng. Springer, Cham, 2018, pp. 1–17.

[4] J. Baldeaux and J. Dick. "QMC rules of arbitrary high order: reproducing kernel Hilbert space approach". *Constructive Approximation. An International Journal for Approximations and Expansions* 30.3 (2009), pp. 495–527.

[5] A. D. Bandrauk and H. Shen. "Higher order exponential split operator method for solving time-dependent Schrödinger equations". *Canadian Journal of Chemistry* 70.2 (1992), pp. 555–559.

[6] W. Bao and J. Yin. "A fourth-order compact time-splitting Fourier pseudospectral method for the Dirac equation". *Research in the Mathematical Sciences* 6.1 (2019), Paper No. 11, 35.

[7] M. Beckers and A. Haegemans. "Transformation of integrands for lattice rules". *Numerical integration (Bergen, 1991)*. Vol. 357. NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. Kluwer Acad. Publ., Dordrecht, 1992, pp. 329–340.

[8] S. Blanes, F. Casas, and M. Thalhammer. “High-order commutator-free quasi-Magnus exponential integrators for non-autonomous linear evolution equations”. *Computer Physics Communications* 220 (2017), pp. 243–262.

[9] G. Byrenheid, L. Kämmerer, T. Ullrich, and T. Volkmer. “Tight error bounds for rank-1 lattice sampling in spaces of hybrid mixed smoothness”. *Numerische Mathematik* 136.4 (2017), pp. 993–1034.

[10] F. Calogero. “Solution of the one-dimensional N-body problems with quadratic and/or inversely quadratic pair potentials”. *Journal of Mathematical Physics* 12 (1971), pp. 419–436.

[11] R. Cools. “Constructing cubature formulae: the science behind the art”. *Acta numerica, 1997*. Vol. 6. Acta Numer. Cambridge Univ. Press, Cambridge, 1997, pp. 1–54.

[12] R. Cools and D. Nuyens. “A Belgian view on lattice rules”. *Monte Carlo and quasi-Monte Carlo methods 2006*. Springer, Berlin, 2008, pp. 3–21.

[13] R. Cools and I. H. Sloan. “Minimal cubature formulae of trigonometric degree”. *Mathematics of Computation* 65.216 (1996), pp. 1583–1600.

[14] R. Cools, F. Y. Kuo, and D. Nuyens. “Constructing embedded lattice rules for multivariable integration”. *SIAM Journal on Scientific Computing* 28.6 (2006), pp. 2162–2188.

[15] R. Cools, F. Y. Kuo, and D. Nuyens. “Constructing lattice rules based on weighted degree of exactness and worst case error”. *Computing* 87.1 (2010), pp. 63–89.

[16] R. Cools, F. Y. Kuo, D. Nuyens, and G. Suryanarayana. “Tent-transformed lattice rules for integration and approximation of multivariate non-periodic functions”. *Journal of Complexity* 36 (2016), pp. 166–181.

[17] H. Dammertz, A. Keller, and S. Dammertz. “Simulation on rank-1 lattices”. *Monte Carlo and quasi-Monte Carlo methods 2006*. Springer, Berlin, 2008, pp. 205–216.

[18] J. Dick, F. Y. Kuo, and I. H. Sloan. “High-dimensional integration: the quasi-Monte Carlo way”. *Acta Numerica* 22 (2013), pp. 133–288.

[19] J. Dick, F. Y. Kuo, Q. T. Le Gia, D. Nuyens, and C. Schwab. “Higher order QMC Petrov-Galerkin discretization for affine parametric operator equations with random field inputs”. *SIAM Journal on Numerical Analysis* 52.6 (2014), pp. 2676–2702.

[20] J. Dick, D. Nuyens, and F. Pillichshammer. “Lattice rules for nonperiodic smooth integrands”. *Numerische Mathematik* 126.2 (2014), pp. 259–291.

[21] J. Dick, I. H. Sloan, X. Wang, and H. Woźniakowski. “Liberating the weights”. *Journal of Complexity* 20.5 (2004), pp. 593–623.

[22] J. Dick, C. Irrgeher, G. Leobacher, and F. Pillichshammer. “On the optimal order of integration in Hermite spaces with finite smoothness”. *SIAM Journal on Numerical Analysis* 56.2 (2018), pp. 684–707.

[23] S. Disney and I. H. Sloan. “Error bounds for the method of good lattice points”. *Mathematics of Computation* 56.193 (1991), pp. 257–266.

[24] T. Goda, K. Suzuki, and T. Yoshiki. “Lattice rules in non-periodic subspaces of Sobolev spaces”. *Numerische Mathematik* 141.2 (2019), pp. 399–427.

[25] V. Gradinaru. “Fourier transform on sparse grids: code design and the time dependent Schrödinger equation”. *Computing. Archives for Scientific Computing* 80.1 (2007), pp. 1–22.

[26] V. Gradinaru. “Strang splitting for the time-dependent Schrödinger equation on sparse grids”. *SIAM Journal on Numerical Analysis* 46.1 (2007/08), pp. 103–123.

[27] L. Grafakos. *Classical Fourier Analysis*. Third Edition. Vol. 249. Graduate Texts in Mathematics. Springer, New York, 2014, pp. xviii+638.

[28] M. Griebel and J. Hamaekers. “Sparse grids for the Schrödinger equation”. *M2AN. Mathematical Modelling and Numerical Analysis* 41.2 (2007), pp. 215–247.

[29] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*. Second. Vol. 31. Springer Series in Computational Mathematics. Structure-preserving algorithms for ordinary differential equations. Springer-Verlag, Berlin, 2006, pp. xviii+644.

[30] P. Hellekalek. “On the assessment of random and quasi-random point sets”. *Random and quasi-random point sets*. Vol. 138. Lect. Notes Stat. Springer, New York, 1998, pp. 49–108.

[31] F. J. Hickernell. “Obtaining $O(N^{-2+\epsilon})$ convergence for lattice quadrature rules”. *Monte Carlo and quasi-Monte Carlo methods, 2000 (Hong Kong)*. Springer, Berlin, 2002, pp. 274–289.

[32] C. Irrgeher and G. Leobacher. “High-dimensional integration on $\mathbb{R}^d$, weighted Hermite spaces, and orthogonal transforms”. *Journal of Complexity* 31.2 (2015), pp. 174–205.

[33] T. Jahnke and C. Lubich. “Error bounds for exponential operator splittings”. *BIT Numerical Mathematics* 40.4 (2000), pp. 735–744.

[34] S. Joe. “Randomization of lattice rules for numerical multiple integration”. *Journal of Computational and Applied Mathematics* 31.2 (1990), pp. 299–304.

[35] W. Kahan and R.-C. Li. "Composition constants for raising the orders of unconventional schemes for ordinary differential equations". *Mathematics of Computation* 66.219 (1997), pp. 1089–1099.

[36] L. Kämmerer. "Reconstructing multivariate trigonometric polynomials from samples along rank-1 lattices". *Approximation theory XIV: San Antonio 2013*. Vol. 83. Springer Proc. Math. Stat. Springer, Cham, 2014, pp. 255–271.

[37] P. Klin, E. Priolo, and G. Seriani. "Numerical simulation of seismic wave propagation in realistic 3-D geo-models with a Fourier pseudo-spectral method". *Geophysical Journal International* 183.2 (2010), pp. 905–922.

[38] O. Koch, C. Neuhauser, and M. Thalhammer. "Error analysis of high-order splitting methods for nonlinear evolutionary Schrödinger equations and application to the MCTDHF equations in electron dynamics". *ESAIM. Mathematical Modelling and Numerical Analysis* 47.5 (2013), pp. 1265–1286.

[39] N. M. Korobov. "Approximate evaluation of repeated integrals". *Doklady Akademii Nauk SSSR* 124 (1959), pp. 1207–1210.

[40] N. M. Korobov. *Teoretiko-chislovye metody v priblizhennom analize*. Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1963, p. 224.

[41] D. Kosloff and R. Kosloff. "A Fourier method solution for the time dependent Schrödinger equation as a tool in molecular dynamics". *Journal of Computational Physics* 52.1 (1983), pp. 35–53.

[42] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, and H. Woźniakowski. "On decompositions of multivariate functions". *Mathematics of Computation* 79.270 (2010), pp. 953–966.

[43] F. Y. Kuo, G. Migliorati, F. Nobile, and D. Nuyens. "Function integration, reconstruction and approximation using rank-1 lattices" (2019). arXiv: `1908.01178 [math.NA]`.

[44] F. Y. Kuo, G. W. Wasilkowski, and H. Woźniakowski. "Lattice algorithms for multivariate L_∞ approximation in the worst-case setting". *Constructive Approximation. An International Journal for Approximations and Expansions* 30.3 (2009), pp. 475–493.

[45] F. Y. Kuo, I. H. Sloan, and H. Woźniakowski. "Lattice rules for multivariate approximation in the worst case setting". *Monte Carlo and quasi-Monte Carlo methods 2004*. Springer, Berlin, 2006, pp. 289–330.

[46] F. Y. Kuo, I. H. Sloan, and H. Woźniakowski. "Periodization strategy may fail in high dimensions". *Numerical Algorithms* 46.4 (2007), pp. 369–391.

[47] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, and B. J. Waterhouse. "Randomly shifted lattice rules with the optimal rate of convergence for unbounded integrands". *Journal of Complexity* 26.2 (2010), pp. 135–160.

[48] P. L'Ecuyer and R. Couture. "An implementation of the lattice and spectral tests for multiple recursive linear random number generators". *INFORMS Journal on Computing* 9.2 (1997), pp. 206–217.

[49] D. H. Lehmer. "On the maxima and minima of Bernoulli polynomials". *American Mathematical Monthly* 47 (1940), pp. 533–538.

[50] D. Li and F. J. Hickernell. "Trigonometric spectral collocation methods on lattices". *Recent advances in scientific computing and partial differential equations (Hong Kong, 2002)*. Vol. 330. Contemp. Math. Amer. Math. Soc., Providence, RI, 2003, pp. 121–132.

[51] C. Lubich. *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2008, pp. x+144.

[52] J. N. Lyness. "An introduction to lattice rules and their generator matrices". *IMA Journal of Numerical Analysis* 9.3 (1989), pp. 405–419.

[53] H. Munthe-Kaas and T. Sørevik. "Multidimensional pseudo-spectral methods on lattice grids". *Applied Numerical Mathematics. An IMACS Journal* 62.3 (2012), pp. 155–165.

[54] C. Neuhauser and M. Thalhammer. "On the convergence of splitting methods for linear evolutionary Schrödinger equations involving an unbounded potential". *BIT. Numerical Mathematics* 49.1 (2009), pp. 199–215.

[55] D. T. P. Nguyen and D. Nuyens. "MDFEM: Multivariate decomposition finite element method for elliptic PDEs with uniform random diffusion coefficients using higher-order QMC and FEM" (2018). arXiv: `1810.13264 [math.NA]`.

[56] D. T. P. Nguyen and D. Nuyens. "MDFEM: Multivariate decomposition finite element methods for elliptic PDEs with lognormal diffusion coefficients using higher-order QMC and FEM" (2019). arXiv: `1904.13327 [math.NA]`.

[57] D. T. P. Nguyen and D. Nuyens. "Multivariate integration over $\mathbb{R}^s$ with exponential rate of convergence". *Journal of Computational and Applied Mathematics* 315 (2017), pp. 327–342.

[58] J. A. Nichols and F. Y. Kuo. "Fast CBC construction of randomly shifted lattice rules achieving $\mathcal{O}(n^{-1+\delta})$ convergence for unbounded integrands over $\mathbb{R}^s$ in weighted spaces with POD weights". *Journal of Complexity* 30.4 (2014), pp. 444–468.

[59] H. Niederreiter. “Existence theorems for efficient lattice rules”. *Numerical integration (Bergen, 1991)*. Vol. 357. NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. Kluwer Acad. Publ., Dordrecht, 1992, pp. 71–80.

[60] H. Niederreiter. “Improved error bounds for lattice rules”. Vol. 9. 1. Festschrift for Joseph F. Traub, Part I. 1993, pp. 60–75.

[61] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. Vol. 63. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992, pp. vi+241.

[62] H. Niederreiter. “The existence of efficient lattice rules for multidimensional numerical integration”. *Mathematics of Computation* 58.197 (1992), 305–314, S7–S16.

[63] *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.0.24 of 2019-09-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

[64] E. Novak and H. Woźniakowski. *Tractability of Multivariate Problems. Volume I: Linear Information*. Vol. 6. EMS Tracts in Mathematics. European Mathematical Society (EMS), Zürich, 2008, pp. xii+384.

[65] E. Novak and H. Woźniakowski. *Tractability of Multivariate Problems. Volume II: Standard Information for Functionals*. Vol. 12. EMS Tracts in Mathematics. European Mathematical Society (EMS), Zürich, 2010, pp. xviii+657.

[66] E. Novak and H. Woźniakowski. *Tractability of Multivariate Problems. Volume III: Standard Information for Operators*. Vol. 18. EMS Tracts in Mathematics. European Mathematical Society (EMS), Zürich, 2012, pp. xviii+586.

[67] E. Novak, I. H. Sloan, and H. Woźniakowski. “Tractability of approximation for weighted Korobov spaces on classical and quantum computers”. *Foundations of Computational Mathematics. The Journal of the Society for the Foundations of Computational Mathematics* 4.2 (2004), pp. 121–156.

[68] D. Nuyens. “The construction of good lattice rules and polynomial lattice rules”. *Uniform Distribution and Quasi-Monte Carlo Methods: Discrepancy, Integration and Applications*. Ed. by Kritzer, P., Niederreiter, H., Pillichshammer, F., and Winterhof, A. Vol. 15. Radon Series on Computational and Applied Mathematics. Berlin, Boston: De Gruyter, 2014, pp. 223–256.

[69] D. Nuyens and R. Cools. “Fast component-by-component construction of rank-1 lattice rules with a non-prime number of points”. *Journal of Complexity* 22.1 (2006), 4–28, Code available online at `https://people.cs.kuleuven.be/~dirk.nuyens/code/`.

[70] D. Nuyens, G. Suryanarayana, and M. Weimar. “Rank-1 lattice rules for multivariate integration in spaces of permutation-invariant functions: error bounds and tractability”. *Advances in Computational Mathematics* 42.1 (2016), pp. 55–84.

[71] F. Pillichshammer and M. Sonnleitner. “A note on isotropic discrepancy and spectral test of lattice point sets”. *Journal of Complexity* (2019).

[72] K. Poppe and R. Cools. “CHEBINT: a MATLAB/Octave toolbox for fast multivariate integration and interpolation based on Chebyshev approximations over hypercubes”. *Association for Computing Machinery. Transactions on Mathematical Software* 40.1 (2013), Art. 2, 13.

[73] D. Potts and T. Volkmer. “Fast and exact reconstruction of arbitrary multivariate algebraic polynomials in Chebyshev form”. *2015 International Conference on Sampling Theory and Applications (SampTA)*. 2015, pp. 392–396.

[74] M. Reed and B. Simon. *Methods of Modern Mathematical Physics. II. Fourier Analysis, Self-adjointness*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975, pp. xv+361.

[75] W. Rudin. *Principles of Mathematical Analysis*. Third. International Series in Pure and Applied Mathematics. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976, pp. x+342.

[76] W. Rudin. *Real and Complex Analysis*. Third. McGraw-Hill Book Co., New York, 1987, pp. xiv+416.

[77] K. Schmüdgen. *Unbounded self-adjoint operators on Hilbert space*. Vol. 265. Graduate Texts in Mathematics. Springer, Dordrecht, 2012, pp. xx+432.

[78] A. Sidi. “A new variable transformation for numerical integration”. *Numerical integration, IV (Oberwolfach, 1992)*. Vol. 112. Internat. Ser. Numer. Math. Birkhäuser, Basel, 1993, pp. 359–373.

[79] I. H. Sloan, F. Y. Kuo, and S. Joe. “Constructing randomly shifted lattice rules in weighted Sobolev spaces”. *SIAM Journal on Numerical Analysis* 40.5 (2002), pp. 1650–1665.

[80] I. H. Sloan and S. Joe. *Lattice Methods for Multiple Integration*. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1994, pp. xii+239.

[81] I. H. Sloan and H. Woźniakowski. “Tractability of multivariate integration for weighted Korobov classes”. Vol. 17. 4. Complexity of multivariate problems (Kowloon, 1999). 2001, pp. 697–721.

[82] I. H. Sloan and H. Woźniakowski. “When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals?” *Journal of Complexity* 14.1 (1998), pp. 1–33.

[83] G. Strang. “On the construction and comparison of difference schemes”. *SIAM Journal on Numerical Analysis* 5 (1968), pp. 506–517.

[84] G. Suryanarayana, D. Nuyens, and R. Cools. “Reconstruction and collocation of a class of non-periodic functions by sampling along tent-transformed rank-1 lattices”. *The Journal of Fourier Analysis and Applications* 22.1 (2016), pp. 187–214.

[85] M. Suzuki. “Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations”. *Physics Letters. A* 146.6 (1990), pp. 319–323.

[86] Y. Suzuki and D. Nuyens. “Lattice rules for integration over $\mathbb{R}^d$”. *In preparation* (2019).

[87] Y. Suzuki and D. Nuyens. “Rank-1 lattices and higher-order exponential splitting for the time-dependent Schrödinger equation”. *Proceedings of Monte Carlo and quasi-Monte Carlo methods 2018*. Accepted for publication. Springer, Berlin, 2019/2020.

[88] Y. Suzuki, G. Suryanarayana, and D. Nuyens. “Strang Splitting in Combination with Rank-1 and Rank-r Lattices for the Time-Dependent Schrödinger Equation”. *SIAM Journal on Scientific Computing* 41.6 (2019), B1254–B1283.

[89] M. Thalhammer. “High-order exponential operator splitting methods for time-dependent Schrödinger equations”. *SIAM Journal on Numerical Analysis* 46.4 (2008), pp. 2022–2038.

[90] M. Thalhammer. “Convergence analysis of high-order time-splitting pseudospectral methods for nonlinear Schrödinger equations”. *SIAM Journal on Numerical Analysis* 50.6 (2012), pp. 3231–3258.

[91] M. Thalhammer, M. Caliari, and C. Neuhauser. “High-order time-splitting Hermite and Fourier spectral methods”. *Journal of Computational Physics* 228.3 (2009), pp. 822–832.

[92] G. Wahba. *Spline Models for Observational Data*. Vol. 59. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990, pp. xii+169.

[93] H. Yserentant. *Regularity and Approximability of Electronic Wave Functions*. Vol. 2000. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2010, pp. viii+182.

[94] S. K. Zaremba. "La méthode des "bons treillis" pour le calcul des intégrales multiples". *Applications of number theory to numerical analysis (Proc. Sympos., Univ. Montreal, Montreal, Que., 1971)*. 1972, pp. 39–119.

Curriculum vitae

Education

PhD fellow, Computer Science Department, KU Leuven, Belgium *November 2015 - Present*
Under the supervision of Prof. Dirk Nuyens and Prof. Ronald Cools

Master of Science in a Double Degree Program between: *April 2012 - March 2015*
-Mathematical Statistics in Department of Mathematics, Royal Institute of Technology (KTH), Sweden;
-and Center for Open Systems Management in the School of Science for Open and Environmental Systems, Keio University, Japan.
I received master degrees of science in engineering from both universities on March 2015.

Bachelor of Science and Engineering, Keio University, Japan *April 2008 - March 2012*

Teaching experience

- 2018-2019, second term, Project Mathematical Engineering, supervising project work sessions, preparing the assignments.
- 2017-2018, first term, Scientific Software, supervising exercise sessions, preparing and evaluating the assignments.
- 2016-2017, first term, Scientific Software, supervising exercise sessions, preparing and evaluating the assignments.

Research visit

- November 2018 - December 2018, RICAM, Linz, Austria.

List of publications

Peer reviewed articles

1. Y. Suzuki, G. Suryanarayana, and D. Nuyens. "Strang Splitting in Combination with Rank-1 and Rank-r Lattices for the Time-Dependent Schrödinger Equation". *SIAM Journal on Scientific Computing* 41.6 (2019), B1254–B1283

2. Y. Suzuki and D. Nuyens. "Rank-1 lattices and higher-order exponential splitting for the time-dependent Schrödinger equation". *Proceedings of Monte Carlo and quasi-Monte Carlo methods 2018*. Accepted for publication. Springer, Berlin, 2019/2020.

3. Y. Suzuki and T. Gudmundsson. Markov chain monte carlo for risk measures. In *2014 4th International Conference On Simulation And Modeling Methodologies, Technologies And Applications (SIMULTECH)*, pages 330–338., Aug 2014.

4. Y. Suzuki and J. Imai. Estimating portfolio's risk measures with dimension reduction techniques. *Transactions of the Journal of the Operations Research Society of Japan*, 55:177–193 (in Japanese)., 2012.

Preprints

1. Y. Suzuki and D. Nuyens. "Lattice rules for integration over $\mathbb{R}^d$". *In preparation* (2019)

International conferences

1. Y. Suzuki and D. Nuyens. Lattice rules for integration over $\mathbb{R}^d$. Workshop on High-Dimensional Approximation (HDA2019), Date: 2019/09/9 - 2019/09/13, Location: Zürich, Switzerland.

2. Y. Suzuki and D. Nuyens. Rank-1 lattices and higher-order exponential splitting for the multi-dimensional time-dependent schrödinger equation. 13th International Conference in Monte Carlo & Quasi-Monte Carlo Methods in Scientific Computing (MCQMC2018), Date: 2018/07/01 - 2018/07/06, Location: Rennes, France.

3. Y. Suzuki and D. Nuyens. Quasi-monte carlo for calculating multi-dimensional expectations against gaussian distribution. International Conference on Monte Carlo Methods and Applications (MCM2017), Date: 2017/07/03 - 2017/07/07, Location: Montreal, Canada.

4. Y. Suzuki, G. Suryanarayana and D. Nuyens. Rank-1 lattices and strang splitting for the time-dependent schrödinger equation. Workshop on High-Dimensional Approximation (HDA2017), Date: 2017/02/13 - 2017/02/17, Location: Sydney, Australia.

5. Y. Suzuki and J. Imai. Archimedean copulas for random sums. Quantitative Methods in Finance 2015 (QMF2015), Date: 2015/12/15 - 2015/12/18, Location: Sydney, Australia.

FACULTY OF ENGINEERING SCIENCE
DEPARTMENT OF COMPUTER SCIENCE
NUMA SECTION
Celestijnenlaan 200A box 2402
B-3001 Leuven
yuyas@kth.se