Nonlocal models with truncated interaction kernels

Analysis, finite element methods and shape optimization

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

vorgelegt am Fachbereich IV
der Universität Trier von

Christian Vollmann

Trier, im Mai 2019

Gutachter: Prof. Dr. Volker Schulz
Prof. Dr. Leonhard Frerick
Prof. Dr. Max Gunzburger
Prof. Dr. Martin Siebenborn

Abstract

Nonlocal operators are used in a wide variety of models and applications due to many natural phenomena being driven by nonlocal dynamics. Nonlocal operators are integral operators allowing for interactions between two distinct points in space. The nonlocal models investigated in this thesis involve kernels that are assumed to have a finite range of nonlocal interactions. Kernels of this type are used in nonlocal elasticity and convection-diffusion models as well as finance and image analysis. Also within the mathematical theory they arouse great interest, as they are asymptotically related to fractional and classical differential equations.

The results in this thesis can be grouped according to the following three aspects: modeling and analysis, discretization and optimization.

Mathematical *models* demonstrate their true usefulness when put into numerical practice. For computational purposes, it is important that the support of the kernel is clearly determined. Therefore nonlocal interactions are typically assumed to occur within an Euclidean ball of finite radius. In this thesis we consider more general interaction sets including norm induced balls as special cases and extend established results about well-posedness and asymptotic limits.

The *discretization* of integral equations is a challenging endeavor. Especially kernels which are truncated by Euclidean balls require carefully designed quadrature rules for the implementation of efficient finite element codes. In this thesis we investigate the computational benefits of polyhedral interaction sets as well as geometrically approximated interaction sets. In addition to that we outline the computational advantages of sufficiently structured problem settings.

Shape *optimization* methods have been proven useful for identifying interfaces in models governed by partial differential equations. Here we consider a class of shape optimization problems constrained by nonlocal equations which involve interface-dependent kernels. We derive the shape derivative associated to the nonlocal system model and solve the problem by established numerical techniques.

Zusammenfassung

Nichtlokale Operatoren werden in einer Vielzahl von Modellen und Anwendungen genutzt, da zahlreiche Phänomene durch nichtlokale Dynamiken beeinflusst werden. Nichtlokale Operatoren sind Integraloperatoren, die Interaktionen zwischen zwei disjunkten Punkten im Raum ermöglichen. Die in dieser Arbeit untersuchten nichtlokalen Modelle basieren auf Kernen mit beschränktem Interaktionshorizont. Kerne dieser Art werden beispielsweise in nichtlokalen Elastizitäts- und Diffusionsmodellen sowie in der Finanzmarkttheorie und Bildbearbeitung eingesetzt. Durch ihren direkten Bezug zu fraktionellen Operatoren sowie im weiteren Sinne zu Differentialgleichungen, sind nichtlokale Operatoren auch innermathematisch von großer Bedeutung.

Die Ergebnisse dieser Arbeit können gemäß der folgenden drei Aspekte kategorisiert werden: Modellierung und Analysis, Diskretisierung und Optimierung.

Das Anwendungspotential mathematischer *Modelle* kann erst dann ausgeschöpft werden, wenn diese eine praktikable numerische Umsetzung gewährleisten. Zu diesem Zweck ist es wichtig, dass der Träger des verwendeten Kerns klar definiert ist. In zahlreichen Modellen wird daher die Annahme getroffen, dass nichtlokale Interaktionen innerhalb einer Kugel von endlichem Radius stattfinden. In dieser Arbeit betrachten wir allgemeinere Interaktionsmengen und erweitern bekannte Resultate über Wohlgestelltheit und asymptotische Grenzwerte.

Die *Diskretisierung* von Integralgleichungen stellt den Anwender oftmals vor große Herausforderungen. Insbesondere Kerne mit beschränktem Träger können die Implementierung von beispielsweise Finite-Elemente-Methoden erschweren und erfordern sorgfältig gewählte Quadraturformeln. In dieser Arbeit untersuchen wir die numerischen Vorteile von polyedrischen sowie geometrisch approximierten Interaktionsmengen.

Methoden der *Formoptimierung* werden erfolgreich eingesetzt, um Interface-abhängige Parameter partieller Differentialgleichungen zu identifizieren. In dieser Arbeit wird ein Formoptimierungsproblem betrachtet, das durch eine nichtlokale Gleichung beschränkt ist, bei dem der Kern von einem glatten Interface abhängt. Wir wenden etablierte Techniken der Formoptimierung an, um dieses Problem numerisch zu lösen.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Professor Dr. Volker Schulz for adopting me as a doctoral student into his group, for his guidance and confidence along the way and for the encouragement to develop own research interests. His aptitude and infectious enthusiasm for numerical Mathematics were essential to materialize this thesis. I am equally thankful to my second supervisor Professor Dr. Leonhard Frerick, aka *der freundliche Analytiker*, who has been the lecturer of my very first academic course in Mathematics and since then has always had an open door to share with me his coffee and extraordinary expertise in functional analysis. Furthermore, I am very grateful to Professor Dr. Max Gunzburger for attending my defense and far more for hosting me for half a year in his group at FSU, as well as for the many valuable and challenging ideas, which were indispensable for this thesis. I would also like to thank Professor Dr. Martin Siebenborn for the attendance of being referee and for his mentoring during my PhD including the many coffees over road bikes and the wonderful Eifel.

I am thankful to Dr. Marta D'Elia who guided me through the nonlocal universe and who introduced me to Max. Our numerous discussions with Max were fundamental to the results in Chapter 6 and Chapter 7. Grazie mille!

I owe thanks to the DFG Research Training Group 2126 'Algorithmic Optimization'(ALOP), which employed me as a research assistant at Trier University during my PhD. studies.

Moreover, I thank all my colleagues from the department of Mathematics at Trier University for making everyday work so easygoing. Especially my friend Manuel, who accompanies me from early times on and who has apparently read the whole thesis. And Gena, who harbored me on the tensor manifold E127. A special thank goes to Daniel, David and Stephen for valuable comments on the manuscript.

Finally, I would like to thank my parents Thomas and Gertrud as well as my brother Stefan and my sister-in-law Ina. Vielen Dank für Eure Unterstützung in allen Lebenslagen – besonders in den hungrigen.

Preface

As part of this thesis, a refereed publication was written, which appears as [87] in the bibliography. The content of this work, including figures and results of numerical tests, can be found in Chapter 5.

Contents

Chapter 1

Introduction

Many physical relations and data-based coherences cannot satisfactorily be described by classical differential equations. Often they inherently possess some features, which are not purely local. Thus, mathematical models which are governed by nonlocal operators appear in a large variety of applications. These include among others, anomalous or fractional diffusion [17, 18, 33], peridynamics [39, 78, 98, 43], image processing [47, 56, 66], cardiology [28], machine learning [68], as well as finance and jump processes [55, 12, 11, 83, 42]. Due to their prevalence the analysis of such operators and the development of efficient discretization methods is of great importance. Nonlocal models are distinctively characterized by the fact that they involve integral equations to describe the most dominant coherences between the variables appearing in the model. Consequently, interactions between two or more domains can occur over a nonzero distance. This stands in stark contrast to classical partial differential equations which are governed by local operators.

The central problem in this thesis is the *steady-state nonlocal Dirichlet problem with volume constraints*, given by

$$\begin{cases} -\mathcal{L}u = f & \text{on } \Omega \\ \quad u = g & \text{on } \Omega_I, \end{cases} \tag{1.1}$$

posed on a bounded domain $\Omega \subset \mathbb{R}^d$. The governing operator $\mathcal{L}$ is a *nonlocal convection-diffusion operator* of the form

$$-\mathcal{L}u(\mathbf{x}) = \int_{\mathbb{R}^d} (u(\mathbf{x})\gamma(\mathbf{x},\mathbf{y}) - u(\mathbf{y})\gamma(\mathbf{y},\mathbf{x}))d\mathbf{y}, \tag{1.2}$$

which is determined by a nonnegative *(interaction) kernel* $\gamma\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. Throughout this thesis we consider kernels which can be written as

$$\gamma(\mathbf{x},\mathbf{y}) = \phi(\mathbf{x},\mathbf{y})\chi_{S(\mathbf{x})}(\mathbf{y}), \tag{1.3}$$

for an appropriate positive function $\phi\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, which we refer to as *kernel function*. Here $\chi_{S(\mathbf{x})}$ denotes the indicator function of the *interaction set* $S(\mathbf{x}) \subset \mathbb{R}^d$. The kernel γ, or more precisely the kernel function ϕ, accounts for nonlocal interactions between two

possibly disjoint points in space. By construction (1.3) these interactions are spatially limited to only occur within the interaction set $S(\mathbf{x})$. Although the case $S(\mathbf{x}) = \mathbb{R}^d$ may be of theoretical interest there are no interactions over infinite distances in real-world applications. We therefore mostly consider finitely truncated kernels, i.e., bounded interaction sets. The second equation in (1.1) is called *Dirichlet volume constraint*. It specifies the values of u on the *interaction domain* $\Omega_I \subset \mathbb{R}^d \backslash \Omega$ which consists of all points in the complement of Ω that interact with points in Ω.

In this thesis, we address the following three aspects of problem (1.1).

- **Modeling and analysis.** We consider a variational formulation of (1.1) for different choices of kernel functions ϕ and interaction sets $S(\mathbf{x})$ in (1.3). We address related analytical questions such as well-posedness and asymptotic limits.
- **Discretization.** We strongly focus on numerical computations and therefore extensively discuss the implementation of the finite element method for the nonlocal equation at hand and present approaches to overcome arising challenges. We make a point corroborating all theoretical findings by numerical experiments.
- **Optimization.** We formulate, analyze and numerically solve a shape optimization problem which is constrained by a nonlocal convection-diffusion equation of the form (1.1).

Let us highlight some special cases and applications of problem (1.1). A particular choice of kernel function and interaction set, namely

$$\phi_{d,s}(\mathbf{x}, \mathbf{y}) = \frac{c_{d,s}}{\|\mathbf{y} - \mathbf{x}\|_2^{d+2s}} \quad \text{and} \quad S(\mathbf{x}) = \mathbb{R}^d, \tag{1.4}$$

results in the *(integral) fractional Laplacian* $(-\Delta)^s$ being a special case of the nonlocal operator $-\mathcal{L}$ defined in (1.2) [33]. The fractional Laplacian is perhaps the most prominent representative for nonlocal operators of type (1.2) and frequently found in the literature. For instance, the article [54] gives an overview on its various definitions. In [2] the regularity of variational solutions and finite element approximations of problem (1.1) with the choice (1.4) are discussed; with a supplementing paper [1] on a two-dimensional finite element implementation. Also, in the two recent publications [4] and [5] the authors develop sophisticated finite element codes for this choice of kernel.

The fractional kernel (1.4) is an example for a radially symmetric and singular kernel. However, another class of kernels, which is often seen in literature [36], consists of *integrable* kernels; i.e., kernels for which the integral

$$h(\mathbf{x}) := \int_{\mathbb{R}^d} \gamma(\mathbf{x}, \mathbf{y}) d\mathbf{y} \tag{1.5}$$

is finite. In this case, the nonlocal operator $-\mathcal{L}$ can be written as

$$-\mathcal{L}u(\mathbf{x}) = h(\mathbf{x})u(\mathbf{x}) - \int_{\mathbb{R}^d} u(\mathbf{y})\gamma(\mathbf{y}, \mathbf{x}) d\mathbf{y}. \tag{1.6}$$

Under appropriate conditions on the kernel γ the second convolution-type term in (1.6) can be considered as a compact operator acting on the function u. In this instance, the nonlocal convection-diffusion operator $-\mathcal{L}$ is a compact perturbation of a scalar multiple of the identity operator, which allows us to invoke Fredholm theory for analytical purposes. In [52, 27, 22] the authors derive analytical results for a time-dependent version of (1.1) with integrable kernels of radial type. The results accumulated in these papers are part of the comprehensive monograph [10] on this topic.

The notion of fractional derivatives traces back to the mathematicians Leibniz and Euler and also Fredholm developed his theory more than 100 years ago. More recently, the birth of the *peridynamics* model [78], introduced by Steward Silling in 2000, has led to an increased focus on nonlocal models of the form (1.1); see, e.g., [39, 98, 43] and the references therein. Peridynamics is a nonlocal continuum model for solid mechanics that is free of spatial derivatives. In contrast to classical models it allows for interactions between material points which are separated by a finite distance. A constitutive assumption is that these nonlocal interactions are spatially limited to only occur within an Euclidean ball of finite radius. Since integral equations remain valid in the presence of discontinuities such as cracks, peridynamics has the potential to model fracture and damage without requiring an external criterion for crack initiation and propagation.

Furthermore, nonlocal operators can be used as filters for image denoising [47, 56, 66]. Considering not only local information at a single pixel they enable the reconstruction of important image features by considering intensity patterns from a neighborhood or from all over the image. In fact, classical methods use local (differential) operators and do not necessarily guarantee feature preservation. In this context, kernels which are truncated by rectangular interaction sets, such as balls induced by the supremum norm, are a convenient choice as image data is usually given on regular grids.

When it comes to numerical implementations we are confronted with the "curse of nonlocality"; the price we have to pay for including nonlocality into our model. Regardless of which discretization approach is utilized we have to employ carefully chosen quadrature rules for the occurring integral equations, especially when singular or truncated kernels are involved or discontinuous solutions are to be expected. Nonlocality also leads to densely populated matrices which may necessitate additional memory storage and tailored solvers. For various types of fractional derivatives finite difference schemes are commonly used to discretize corresponding nonlocal operators [89, 90, 57]. For mechanics applications, such as the peridynamics model, we typically have to deal with more complicated domains in two or three dimensional space to model the body of interest. Here, finite difference schemes are ineligible and meshfree methods, in particular particle-type methods [86, 74, 65, 79], are frequently used to discretize problem (1.1).

In this thesis we consider variational methods, and in particular finite element methods. Variational methods are superior to other approximation techniques since they provide a convenient framework for dealing with complicated domains, obtain higher-order convergence rates and facilitate adaptive meshing methods that help to resolve solution misbehaviors such as steep gradients or discontinuities. In addition, a variational framework allows for a rigorous mathematical treatment of operator and solution

properties such as convergence and related stability issues. Motivated by this endeavor, a nonlocal vector calculus has been introduced in [48] and further developed in [37] and [8]. A comprehensive survey is given by [36], which applies the nonlocal vector calculus to the problem at hand (1.1). Therein, the authors derive its variational formulation and well-posedness results for a class of symmetric singular kernels such as (1.4) and a class of symmetric integrable kernels that satisfy (1.5). In [42] similar results are established for nonsymmetric kernels which are related to nonsymmetric jump processes. In both papers, [36] and [42], the kernels are truncated by the Euclidean ball $S(\mathbf{x}) = B_{\delta,2}(\mathbf{x})$, which is why we often refer to this kind of truncation as the *standard* case.

Not only the problem itself but also various optimization problems involving nonlocal models of the form (1.1) are treated in literature. For example matching-type problems are treated in [34, 32, 35] to identify system parameters such as the forcing term or a scalar diffusion parameter. The control variable is typically modeled to be an element of a suitable function space. On the other hand, shape optimization problems are of interest in many applied fields, particularly in the context of partial differential equations. Some examples include aerodynamic shape optimization [69], image segmentation [51] or shape optimization for the identification of interfaces [76, 73, 71, 70]. However, shape optimization techniques applied to nonlocal models can hardly be found in literature. For instance, the articles [16, 81, 64] deal with minimizing (functions of) certain eigenvalues of the fractional Laplacian with respect to the domain Ω. Also, in [29, 15] the energy functional related to fractional equations is minimized. In [19] a functional involving a more general kernel is considered. All of the aforementioned papers are of theoretical nature only. To the best of our knowledge, shape optimization problems involving nonlocal constraint equations of type (1.1) with truncated kernels and numerical methods for solving such problems, cannot yet be found in literature.

Corresponding to the three main aspects addressed in the title of this thesis, we make the following contributions to the field of nonlocal equations.

- **Modeling and analysis.** We extend the aforementioned well-posedness results presented in [36] and [42] to kernels which are truncated by more general interaction sets, thereby allowing for polyhedral interaction sets. This gives natural rise to the question of how variational solutions corresponding to different kernels differ and we establish a framework effecting such a comparison study. This framework is then further applied as follows. On the one hand, we generalize the convergence results given in [33], where it is shown that solutions corresponding to a truncated version of the fractional Laplacian converge as $\delta \to \infty$ to solutions corresponding to the (nontruncated) fractional Laplacian. On the other hand, we apply the comparison result to analyze the effect that geometrically approximated interaction sets have on the global finite element error. We also consider the local limit and extend the results in [85, 38, 10] to general norm induced interaction sets and kernel functions which are radial with respect to any norm.

- **Discretization.** We give a comprehensive description of the implementation of the finite element method for approximating solutions of problem (1.1). We point out related challenges, particularly those arising due to the consideration of truncated kernels of the form (1.3) and reveal the promising computational benefits that polyhedral and approximate interaction sets provide. Furthermore, for hyperrectangular domains and translation invariant kernels we present a structure exploiting finite element code which enables us to solve three-dimensional nonlocal problems.

- **Optimization.** We formulate a shape optimization problem to identify interfaces in nonlocal models. We prove through numerical experiments the applicability of established shape optimization techniques. The crucial ingredient is the derivation of the shape derivative of the nonlocal bilinear form associated to problem (1.1). We derive an explicit expression for this derivative for a specific class of interface-dependent, truncated kernels.

1.1 Structure of the thesis

This thesis is structured in the following way.

Chapter 2 serves as a warm-up chapter that familiarizes the reader with the central problem given in (1.1) and its interpretation as a nonlocal convection-diffusion problem. In Section 2.1 we first formally derive a nonlocal conservation equation from a probabilistic perspective and then we underpin these considerations in Section 2.2 with the concept of a nonlocal vector calculus as introduced in [37]. Means to an end are the notions of a nonlocal flux and a nonlocal divergence operator. Finally, in Section 2.3, based on these discussions we introduce the steady-state nonlocal Dirichlet problem.

Chapter 3 provides analytical results about the well-posedness of problem (1.1) involving possibly nonsymmetric kernel functions and general interaction sets other than the Euclidean ball. For this purpose, we first introduce a variational formulation of (1.1) in Section 3.1 which includes the definition of a nonlocal bilinear form. Then, after setting up some general assumptions on the interaction sets and the kernel in Section 3.2 we provide well-posedness results in Section 3.3 for the two classes of kernels discussed above; namely singular and integrable ones.

In **Chapter** 4 we outline the finite element discretization of problem (1.1) and illuminate the concomitant challenges. Therefore, in Section 4.1 we briefly recall a general finite-dimensional approximation of the weak formulation and then specify to finite element approximations in Section 4.2. The central notion is that of the nonlocal stiffness matrix whose challenging assembly is discussed in Section 4.3. Finally, we introduce the notion of approximate interaction sets in Section 4.4.

Under appropriate assumptions on the kernel and the domain, we show in **Chapter** 5, more precisely in Section 5.1, that the nonlocal stiffness matrix inherits the structure of a multilevel Toeplitz matrix. In Section 5.2 we exploit this structure for solving the resulting linear system and present numerical results in 5.3 including a three-dimensional

study case.

In Section 6.1 of **Chapter** 6 we establish a framework for the comparison of two nonlocal solutions of (1.1), which correspond to two different kernels. A focus is set on the comparison of kernels which solely differ in their interaction sets. In particular we derive precise integral representations to effect the comparison of Euclidean balls and $\|\cdot\|_\infty$-balls for various radii choices. We concretize this special case in Section 6.2 by considering the fractional kernel (1.4). We truncate this kernel by $\|\cdot\|_\infty$-balls and study the convergence of the corresponding solution as the interaction horizon increases. Finally, in Section 6.3 we exploit the comparison results to derive an estimate for the global finite element error when geometrical approximations of the exact interaction sets are used for the assembly of the nonlocal stiffness matrix.

In **Chapter** 7 we treat the local limit as interaction sets shrink to point sets for general norm induced balls. After setting up notation in Section 7.1, we establish convergence results for the nonlocal operator and bilinear form in Section 7.2. Then, in Section 7.3 we estimate the nonlocal energy norm associated with a kernel satisfying our more general setting against the energy norm of the standard setting. This estimate allows us to invoke established results from the standard case to prove the convergence of solutions. Finally, in Section 7.4 we provide the exact scaling constants for several example kernels which we then use in Section 7.5 for numerical tests.

A shape optimization problem to identify interfaces in nonlocal models is the content of **Chapter** 8. After defining the concrete problem setting in Section 8.1, we briefly recall basic concepts of shape optimization in Section 8.2. Central to this chapter is the shape derivative of the nonlocal bilinear form, which is derived in Section 8.3. Finally we present numerical illustrations in Section 8.4.

1.2 Glossary

For the convenience of the reader, we introduce some notation and also list some definitions which are central to this thesis and appear at several places. Thus, the following can be regarded as supplemental glossary.
Throughout, $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ denote points in $\mathbb{R}^d$ and we use bold letters whenever an object is vector-valued. *Two-point functions* map $\mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}^d$ or $\mathbb{R}$ so that $\psi(\mathbf{x},\mathbf{y})$ denotes a scalar two-point function and $\boldsymbol{\psi}(\mathbf{x},\mathbf{y})$ a vector two-point function. *Symmetric* two-point functions satisfy $\psi(\mathbf{x},\mathbf{y}) = \psi(\mathbf{y},\mathbf{x})$ whereas *antisymmetric* two-point functions satisfy $\psi(\mathbf{x},\mathbf{y}) = -\psi(\mathbf{y},\mathbf{x})$. In the remainder we apply the superscript $(\cdot)'$ for indicating a swap of variables, so that for example $\psi'(\mathbf{x},\mathbf{y}) = \psi(\mathbf{y},\mathbf{x})$ or $u'(\mathbf{x}) = u(\mathbf{y})$. The same applies to a bilinear form $A(u,v)$ acting on two functions $u, v\colon \mathbb{R}^d \to \mathbb{R}$, i.e., $A'(u,v) = A(v,u)$ and A is called symmetric if $A = A'$, and antisymmetric if $A = -A'$.
Furthermore, Ω always denotes a bounded domain in $\mathbb{R}^d$, i.e., $\Omega \neq \emptyset$ is bounded, open and connected.

1.2.1 General

Natural numbers and Euclidean space

$\mathbb{N}$	set of natural numbers excluding zero
$\mathbb{R}^d$	Euclidean space of dimension $d \in \mathbb{N}$
$\mathbb{R}^{m\times n}$	set of $m \times n$ matrices in $\mathbb{R}$, where $m, n \in \mathbb{N}$
$\mathbb{R}^{d\times d}_{spd}$	set of symmetric and positive definite $d \times d$ matrices in $\mathbb{R}$
$\|\cdot\|_p$	p-norm in $\mathbb{R}^d$, given by $\|\mathbf{x}\|_p = \sqrt[p]{\sum_{i=1}^d \|x_i\|^p}$ for $p \in [1,\infty)$ and $\|\mathbf{x}\|_\infty = \max_{1\le i\le d} \|x_i\|$; cases of interest are the Euclidean norm ($p = 2$), supremum norm ($p = \infty$) and Manhattan norm ($p = 1$)
$\|x\|$	absolute value of $x \in \mathbb{R}$

Linear algebra

Let $\mathbf{A}$ and $\mathbf{B}$ be two matrices in $\mathbb{R}^{d\times d}$ and let $\mathbf{x}$ and $\mathbf{y}$ denote two vectors in $\mathbb{R}^d$.

$\mathbf{A}^T$	transpose of $\mathbf{A}$
$\det(\mathbf{A})$	determinant of $\mathbf{A}$
$\mathrm{tr}(\mathbf{A})$	trace of $\mathbf{A}$
$\mathbf{A} : \mathbf{B}$	sum of componentwise products, i.e., $\mathbf{A} : \mathbf{B} = \sum_{i,j=1}^d a_{kj} b_{kj}$
$\mathbf{x}^T\mathbf{y}$	scalar product of $\mathbf{x}$ and $\mathbf{y}$
$\mathbf{x}\mathbf{y}^T$	outer product of $\mathbf{x}$ and $\mathbf{y}$
$\mathbf{Id}$	identity matrix in $\mathbb{R}^{d\times d}$

Sets and set operations

Let M and N denote two subsets of $\mathbb{R}^d$.

M^c	complement of M, i.e., $M^c = \mathbb{R}^d \setminus M$
$\|M\|$	d-dimensional volume of M
$\mathrm{int}(M)$	interior of M in $\mathbb{R}^d$
$\overline{M}$	closure of M in $\mathbb{R}^d$
∂M	boundary of M
$\mathrm{diam}(M)$	diameter of M
$M \triangle N$	symmetric difference $M \triangle N = (M \cup N) \setminus (M \cap N)$

$M \dot{\cup} N$	disjoint union of the sets M and N

Differentiation

$d(\cdot)$	derivative with respect to finite-dimensional variables
$d^{\boldsymbol{\alpha}}(\cdot)$	partial derivative $d^{\boldsymbol{\alpha}} = \frac{\partial^{\lvert\boldsymbol{\alpha}\rvert}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}$ where $\boldsymbol{\alpha}$ is a multi-index
$D(\cdot)$	derivative with respect to infinite-dimensional variables
$\mathbf{x}^{\alpha}$	multi-index notation for the monomial $\mathbf{x}^{\alpha} = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$
$\boldsymbol{\alpha}!$	multi-index factorial $\alpha_1! \cdots \alpha_d!$
div	divergence operator
∇	gradient operator or Jacobian
H_u	Hessian of a scalar function u
Δ	Laplace operator
$(-\Delta)^s$	fractional Laplace operator

Banach and Hilbert spaces

Let $\tilde{\Omega} \supset \overline{\Omega}$ be a domain in $\mathbb{R}^d$ containing Ω. Further let $H(\tilde{\Omega})$ denote a vector space of functions $u \colon \tilde{\Omega} \to \mathbb{R}$.

$(\cdot, \cdot)_{H(\tilde{\Omega})}$	inner product on $H(\tilde{\Omega})$
$\lVert \cdot \rVert_{H(\tilde{\Omega})}$	norm on $H(\tilde{\Omega})$
$H^*(\tilde{\Omega})$	dual space of $H(\tilde{\Omega})$
$H_0(\tilde{\Omega})$	set of $H(\tilde{\Omega})$-functions which vanish on the boundary $\partial\tilde{\Omega}$
$H_c(\tilde{\Omega})$	set of $H(\tilde{\Omega})$-functions which vanish on the volume $\tilde{\Omega} \setminus \Omega$

Examples of $H(\tilde{\Omega})$ appearing in this thesis are given by the following spaces.

$L^2(\tilde{\Omega})$	space of square integrable Lebesgue functions
$H^k(\tilde{\Omega})$	Sobolev space $W^{k,2}(\tilde{\Omega})$ of integer order $k \in \mathbb{N}$
$H^s(\tilde{\Omega})$	Sobolev space $W^{s,2}(\tilde{\Omega})$ of fractional (non-integer) order $s \in (0,1)$
$C^k(\tilde{\Omega})$	set of functions with derivative up to order $k \in \mathbb{N} \cup \{0, \infty\}$

Miscellaneous

χ_M	indicator function of a set $M \subset \mathbb{R}^d$
$\text{supp}(u)$	support of a function u
span	linear span
$\mathcal{F}$	Fourier transform
δ_{kj}	Kroenecker delta
$\mathcal{O}(\cdot)$	big $\mathcal{O}$ notation

1.2.2 Specific to nonlocal problems

We now introduce notations and conventions as well as list some definitions of objects which are related to the study of nonlocal models conducted in this thesis.

Nonlocal vector calculus

$\mathcal{D}$	nonlocal divergence
$\mathcal{G}$	nonlocal gradient
$\mathcal{N}$	interaction operator
F^{nl}	nonlocal flux
F^{loc}	local flux

The kernel and interaction sets

Throughout this thesis we consider nonnegative kernels of the form $\gamma(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x}, \mathbf{y})\chi_{S(\mathbf{x})}(\mathbf{y})$ for some appropriate positive kernel function $\phi(\mathbf{x}, \mathbf{y})$ and family of interaction sets $\{S(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$.

γ	kernel accounting for nonlocal interactions in space
γ^s	symmetric part of the kernel γ given by $\gamma^s = \frac{\gamma+\gamma'}{2}$
γ^a	antisymmetric part of the kernel γ given by $\gamma^a = \frac{\gamma-\gamma'}{2}$
γ'	swapped argument, i.e., $\gamma' = \gamma(\mathbf{y}, \mathbf{x})$
γ_*	kernel tagged with $* \in \{a, b\}$ to effect comparison of two nonlocal models
γ_{ab}	difference kernel $\gamma_{ab} = \gamma_a - \gamma_b$
	$\rightarrow$ *similar definitions hold for* ϕ, ϕ^s, ϕ^a, ϕ', ϕ_*, ϕ_{ab}

$\{S(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$	family of interaction sets $S(\mathbf{x}) \subset \mathbb{R}^d$
$S_*(\mathbf{x})$	interaction set tagged with $* \in \{a, b\}$ to effect comparison of two nonlocal models
$B_{\delta,\bullet}(\mathbf{x})$	norm induced ball of radius $\delta > 0$ centered at $\mathbf{x} \in \mathbb{R}^d$, given by $B_{\delta,\bullet}(\mathbf{x}) = \left\{\mathbf{y} \in \mathbb{R}^d \colon \|\mathbf{x} - \mathbf{y}\|_\bullet < \delta\right\}$
$\bullet$	reserved norm indicator used for balls
δ	reserved variable name for interaction radius
Ω_I	nonlocal interaction domain $\Omega_I = \{\mathbf{y} \in \Omega^c \,:\, \mathbf{y} \in S(\mathbf{x}) \text{ for some } \mathbf{x} \in \Omega\}$
Ω_I^*	interaction domain corresponding to γ_* for $* \in \{a, b\}$
Ω_I^{ab}	$\Omega_I^{ab} = \Omega_I^a \cup \Omega_I^b$

Nonlocal convection-diffusion operator

Let $u \colon \Omega \cup \Omega_I \to \mathbb{R}$ be a scalar function.

$\mathcal{L}$	nonlocal operator	$\mathcal{L}u(\mathbf{x}) = \int_{\Omega\cup\Omega_I}(u'\gamma' - u\gamma)d\mathbf{y}$
$\mathcal{L}^d$	diffusion part of $\mathcal{L}$	$\mathcal{L}^d u(\mathbf{x}) = \int_{\Omega\cup\Omega_I}(u' - u)\gamma^s d\mathbf{y}$
$\mathcal{L}^c$	convection part of $\mathcal{L}$	$\mathcal{L}^c u(\mathbf{x}) = \int_{\Omega\cup\Omega_I}(u' + u)\gamma^a d\mathbf{y}$
$\mathcal{L}_\Omega$	part of $\mathcal{L}$ acting on $u\|_\Omega$	$\mathcal{L}_\Omega u(\mathbf{x}) = \int_\Omega(u'\gamma' - u\gamma)d\mathbf{y} + u\int_{\Omega_I}\gamma' d\mathbf{y}$
$\mathcal{L}_{\Omega_I}$	part of $\mathcal{L}$ acting on $u\|_{\Omega_I}$	$\mathcal{L}_{\Omega_I} u(\mathbf{x}) = -\int_{\Omega_I} u'\gamma' d\mathbf{y}$
$\mathcal{L}_\Omega^d$	diffusion part of $\mathcal{L}_\Omega$	$\mathcal{L}_\Omega^d u(\mathbf{x}) = \int_\Omega(u' - u)\gamma^s d\mathbf{y} + u\int_{\Omega_I}\gamma^s d\mathbf{y}$
$\mathcal{L}_\Omega^c$	convection part of $\mathcal{L}_\Omega$	$\mathcal{L}_\Omega^c u(\mathbf{x}) = \int_\Omega(u' + u)\gamma^a d\mathbf{y} + u\int_{\Omega_I}\gamma^a d\mathbf{y}$
$\mathcal{L}^*$	adjoint operator of $\mathcal{L}$	$\mathcal{L}^* u(\mathbf{x}) = \int_{\Omega\cup\Omega_I}(u' - u)\gamma d\mathbf{y}$

Nonlocal Dirichlet problem

Central to this thesis is the nonlocal Dirichlet problem given by $\mathcal{L}u = f$ posed on Ω, where we impose the volume constraint $u = g$ on Ω_I.

u	solution of the nonlocal Dirichlet problem
u_Ω	u restricted to Ω, i.e., $u\|_\Omega$
u_{Ω_I}	u restricted to Ω_I, i.e., $u\|_{\Omega_I}$
u^0	solution of the homogeneous auxiliary problem

u'	swapped argument, i.e., $u'(\mathbf{x}) = u(\mathbf{y})$
u_*	solution corresponding to the kernel γ_* for $* \in \{a, b\}$
f	forcing term
$\tilde{f}$	right-hand side $\tilde{f} = f - \mathcal{L}_{\Omega_I} g$ of the homogeneous auxiliary problem
g	Dirichlet data
g'	swapped argument, i.e., $g'(\mathbf{x}) = g(\mathbf{y})$
g^*	Dirichlet data posed on Ω_I^* for $* \in \{a, b\}$

Nonlocal bilinear and linear form

Let $u, v \colon \Omega \cup \Omega_I \to \mathbb{R}$ denote two scalar functions.

A	nonlocal bilinear form	$A(u,v) = \int_\Omega v \int_\Omega (u\gamma - u'\gamma')\, d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega_I} \gamma \, d\mathbf{y} d\mathbf{x}$
$A_{\Omega\Omega_I}$		$A_{\Omega\Omega_I}(u,v) = (\mathcal{L}_{\Omega_I} u, v)_{L^2(\Omega)}$
A'	swapped arguments	$A'(u,v) = A(v,u)$
A^d	diffusion part of A	$A^d(u,v) = \int_\Omega v \int_\Omega (u - u')\, \gamma^s d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega_I} \gamma^s d\mathbf{y} d\mathbf{x}$
A^c	convection part of A	$A^c(u,v) = \int_\Omega v \int_\Omega (u + u')\, \gamma^a d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega_I} \gamma^a d\mathbf{y} d\mathbf{x}$
$A^{c,s}$	symmetric part of A^c	$A^{c,s}(u,v) = \int_\Omega vu \int_{\Omega\cup\Omega_I} \gamma^a d\mathbf{y} d\mathbf{x}$
$A^{c,a}$	antisymmetric part of A^c	$A^{c,a}(u,v) = \int_\Omega v \int_\Omega u' \gamma^a d\mathbf{y} d\mathbf{x}$.
A^s	symmetric part of A	$A^s = \frac{1}{2}(A + A') = A^d + A^{c,s}$
A^a	antisymmetric part A	$A^a = \frac{1}{2}(A - A') = A^{c,a}$
A^*	adjoint bilinear form	$A^*(u,v) = \int_\Omega v \int_\Omega (u - u')\, \gamma d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega_I} \gamma \, d\mathbf{y} d\mathbf{x}$
$A_{d,s}$	fractional bilinear form	$A_{d,s}(u,v) = \frac{c_{d,s}}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{(u-u')(v-v')}{\|\mathbf{x}-\mathbf{y}\|_2^{d+2s}} d\mathbf{y} d\mathbf{x}$
A^{loc}	local bilinear form	$A^{loc}(u,v) = A_0(u,v) = \int_\Omega \nabla u^T \nabla v d\mathbf{x}$
A_*	effecting comparison	nonlocal bilinear form corresponding to γ_* for $* \in \{a, b\}$
A_{ab}	difference bilinear form	$A_a - A_b$
ℓ	right-hand side of weak formulation	$\ell(v) = \ell_{f,g}(v) = \int_\Omega f v \, d\mathbf{x} - \int_\Omega \mathcal{L}_{\Omega_I} g \; v \, d\mathbf{x}$

Nonlocal vector spaces

$\|\|\|\cdot\|\|\|$	nonlocal energy norm $\|\|\|\cdot\|\|\| = \sqrt{A(\cdot,\cdot)}$
$\|\|\|\cdot\|\|\|_*$	nonlocal energy norm corresponding to the kernel γ_* for $* \in \{a, b\}$
$V(\Omega \cup \Omega_I)$	nonlocal energy space
$\|\cdot\|_{V(\Omega\cup\Omega_I)}$	$\|\cdot\|_{V(\Omega\cup\Omega_I)} = \|\|\|\cdot\|\|\| + \|\cdot\|_{L^2(\Omega\cup\Omega_I)}$
$V_c(\Omega \cup \Omega_I)$	nonlocal constrained energy space
$V_c^*(\Omega)$	dual space of $V_c(\Omega \cup \Omega_I)$
$\tilde{V}(\Omega_I)$	nonlocal volume "trace" space

Discretization via the finite element method

$\mathcal{T}_\Omega^h$	triangulation of Ω into finite elements
$\mathcal{T}_{\Omega_I}^h$	triangulation of Ω_I into finite elements
$\mathcal{T}^h$	triangulation of $\Omega \cup \Omega_I$ into finite elements, given by $\mathcal{T}^h = \mathcal{T}_\Omega^h \cup \mathcal{T}_{\Omega_I}^h$
$\mathcal{E}_j$	the j-th finite element of $\mathcal{T}^h$
J, J_Ω, J_{Ω_I}	total number of finite elements on $\Omega \cup \Omega_I$, Ω and Ω_I, respectively
$\mathbf{x}_k$	the k-th finite element node
K, K_Ω	total number of finite element nodes on $\Omega \cup \Omega_I$ and Ω, respectively
$\{\varphi_k\}_{k=1}^K$	set of basis functions assigned to each node of the finite element mesh
$\mathcal{S}_k$	support of the basis function φ_k
V^h	finite element subspace of $V(\Omega \cup \Omega_I)$, given by $V^h = \operatorname{span}\{\varphi_k \colon 1 \le k \le K\}$
V_c^h	finite element subspace of $V_c(\Omega\cup\Omega_I)$, given by $V_c^h = \operatorname{span}\{\varphi_k \colon 1 \le k \le K_\Omega\}$
h	grid size
u^h	finite element approximation of u
v^h	test function of finite element subspace
$\mathbf{A}^h$	nonlocal stiffness matrix $\mathbf{A}^h = \left(\mathbf{A}_{\Omega\Omega}^h, \mathbf{A}_{\Omega\Omega_I}^h\right) \in \mathbb{R}^{K_\Omega \times K}$
$\mathbf{A}_{\Omega\Omega}^h$	$\mathbf{A}_{\Omega\Omega}^h = (A(\varphi_k, \varphi_j))_{(1\le k,j\le K_\Omega)} \in \mathbb{R}^{K_\Omega \times K_\Omega}$
$\mathbf{A}_{\Omega\Omega_I}^h$	$\mathbf{A}_{\Omega\Omega_I}^h = (A_{\Omega\Omega_I}(\varphi_k, \varphi_j))_{(1\le k\le K_\Omega,\ K_\Omega+1\le j\le K)} \in \mathbb{R}^{K_\Omega \times (K-K_\Omega)}$,
$\mathbf{u}_\Omega^h$	discrete solution vector

$\mathbf{f}^h$	discrete forcing term vector
$\mathbf{g}^h$	discrete Dirichlet data vector
$\mathbf{x}_a^{bary}$	barycenter of the element $\mathcal{E}_a$

Numerical quadrature

During the assembly process of the nonlocal stiffness matrix, the main task is to numerically compute double integrals of the form $\int_{\mathcal{E}_a}\int_{\mathcal{E}_b\cap S(\mathbf{x})}\psi(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}$.

$S(\mathcal{E}_a)$	set of all element indices $1 \leq b \leq J$ so that integrals over $\mathcal{E}_a \times \mathcal{E}_b$ have nonzero contributions to the nonlocal stiffness matrix
$\mathcal{E}_I$	interaction domain of the element $\mathcal{E}$
$\mathring{S}(\mathcal{E})$	subset of $S(\mathcal{E})$ consisting of elements which neither necessitate special treatment of the outer integral nor a subdivision task
$\partial S(\mathcal{E})$	subset of $S(\mathcal{E})$ consisting of elements which necessitate special treatment of the outer integral and subdivision task for the inner integral
$\tilde{S}(\mathcal{E})$	subset of $S(\mathcal{E})$ consisting of elements which necessitate subdivision task for the inner integral
$\mathcal{E}_a^0$	$\mathcal{E}_a^0 = \{\mathbf{x} \in \mathcal{E}_a \colon \mathcal{E}_b \cap S(\mathbf{x}) = \emptyset\}$ for an element $\mathcal{E}_b$
$\mathcal{E}_a^1$	$\mathcal{E}_a^1 = \mathcal{E}_a \backslash \mathcal{E}_a^0$ for an element $\mathcal{E}_b$
Q_{out}	quadrature order of a quadrature rule on the outer element $\mathcal{E}_a$
$Q_{out}^{special}$	quadrature order of a specialized quadrature rule on the outer element $\mathcal{E}_a$
$\mathbf{x}_q^a$	the q-th quadrature point of a rule on $\mathcal{E}_a$
w_q^a	the q-th quadrature weight of a rule on $\mathcal{E}_a$
Q_{in}	quadrature order of a quadrature rule on polyhedral subelements of the inner element $\mathcal{E}_b \cap S(\mathbf{x}_q^a)$
$Q_{in}^{nonpoly}$	quadrature order of a quadrature rule on nonpolyhedral subelements of the inner element $\mathcal{E}_b \cap S(\mathbf{x}_q^a)$
J_S	total number of subelements of $\mathcal{E}_b \cap S(\mathbf{x}_q^a)$
J_S^{poly}	number of polyhedral subelements of $\mathcal{E}_b \cap S(\mathbf{x}_q^a)$
$\mathbf{y}_p^b$	the p-th quadrature point of a rule on a subelement of the inner element $\mathcal{E}_b \cap S(\mathbf{x}_q^a)$
w_p^b	the p-th quadrature weight of a rule on a subelement of the inner element $\mathcal{E}_b \cap S(\mathbf{x}_q^a)$

A^h	approximate bilinear form resulting from numerical quadrature
ℓ^h	approximate linear form resulting from numerical quadrature
$S^h(\mathbf{x})$	approximate interaction set
$B^h_{\delta,\bullet}(\mathbf{x})$	approximate norm induced ball

Nonlocal shape optimization

Γ	interface modeled as an element of a suitable shape space
Ω_i	subdomain of Ω_i resulting from the partition induced by the interface
γ_{ij}	partial kernel accounting for nonlocal interactions between Ω_i and Ω_j
γ_Γ	interface-dependent kernel composed of partial kernels γ_{ij}
$\mathcal{L}_\Gamma$, A_Γ, ℓ_Γ	respective nonlocal objects corresponding to γ_Γ
f_Γ	interface-dependent forcing term
$\bar{u}$	given measurements on Ω
$\tilde{A}_\Gamma$	perturbed nonlocal bilinear form $\tilde{A}_\Gamma = A_\Gamma + c_{per}A^{loc}$
c_{per}	perturbation parameter $c_{per} > 0$
J	real-valued shape functional
j	L^2 tracking-type functional
j_{reg}	perimeter regularization
F_t	family of vector fields on Ω for $t > 0$ with $F_0 = id$
$dF_t(\mathbf{x})$	Jacobian of F_t at $\mathbf{x}$
Γ_t	transformed shape effected by F_t
$DJ(\Gamma)[\mathbf{V}]$	shape derivative of J at Γ in direction of $\mathbf{V}$
$\frac{d}{dt}\Big\vert_{t=0^+}$	derivative with respect to $t > 0$ evaluated at $t = 0$
L	Lagrangian
$\dot{u}$ or $D_m(u)$	material derivative of u
ℓ^*	right-hand side of the adjoint equation
ψ_{ij}	integrand of the bilinear form A_Γ
ψ^*_{ij}	integrand of the adjoint bilinear form A^*_Γ

Chapter 2

Introduction to nonlocal convection-diffusion equations

The central problem considered in this thesis is given by

$$\begin{cases} -\mathcal{L}u = f & \text{on } \Omega \\ u = g & \text{on } \Omega_I, \end{cases} \tag{2.1}$$

where $\Omega \subset \mathbb{R}^d$ denotes a bounded domain,

$$-\mathcal{L}u(\mathbf{x}) = \int_{\Omega \cup \Omega_I} (u(\mathbf{x})\gamma(\mathbf{x},\mathbf{y}) - u(\mathbf{y})\gamma(\mathbf{y},\mathbf{x}))d\mathbf{y}$$

denotes a nonlocal convection-diffusion operator and $\gamma\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ denotes a nonnegative function, which we refer to as *kernel*. The set Ω_I is called the *interaction domain* corresponding to the domain Ω and the kernel γ. It consists of those points in the complement domain $\mathbb{R}^d \backslash \Omega$ that interact with points in Ω. The second equation in (2.1) is called *Dirichlet volume constraint* since it specifies the values of u on the domain Ω_I, which has nonzero volume in $\mathbb{R}^d$. This stands in stark contrast to Dirichlet constraints for partial differential equations which are imposed on hypersurfaces. Thus, we naturally refer to (2.1) as *steady-state nonlocal Dirichlet problem with volume constraints*.

This chapter is based on [37, 36], [10, Chapter 1-2], and it is organized as follows. In Section 2.1 we first formally derive a nonlocal conservation equation from a probabilistic perspective to familiarize with problem (2.1). Afterwards in Section 2.2 we underpin these considerations with the concept of a nonlocal vector calculus as introduced in [37]. Means to an end are the notions of a nonlocal flux and a nonlocal divergence operator. Finally, in Section 2.3, based on these discussions we define the steady-state nonlocal convection-diffusion equation which represents the fundamental problem of this thesis.

2.1 A nonlocal conservation equation

Let $u(\mathbf{x}, t)$ denote a density of particles at a point $\mathbf{x} \in \mathbb{R}^d$ in space and at time $t \in [0, \infty)$. We may also think of $u(\mathbf{x}, t)$ as being a heat density or a density accounting for

individuals. Furthermore, let $\gamma(\mathbf{x},\mathbf{y})$ be a nonnegative two-point function accounting for the probability, or say tendency, that a particle moves from location $\mathbf{x}$ to location $\mathbf{y}$. We do not require this kernel to be symmetric and allow for jumps meaning that particles can move with "infinite" speed. We also denote by $f(\mathbf{x},t)$ an external source density taking effect at location $\mathbf{x}\in\mathbb{R}^d$ and time $t\in[0,\infty)$. Then, over a time horizon $[t_1,t_2)$, where $t_2>t_1\geq 0$, we can state the following nonlocal conservation equation

$$\begin{aligned} u(\mathbf{x},t_2) =& u(\mathbf{x},t_1) + \int_{t_1}^{t_2}\int_{\mathbb{R}^d} u(\mathbf{y},s)\gamma(\mathbf{y},\mathbf{x})d\mathbf{y}ds \\ &- \int_{t_1}^{t_2}\int_{\mathbb{R}^d} u(\mathbf{x},s)\gamma(\mathbf{x},\mathbf{y})d\mathbf{y}ds + \int_{t_1}^{t_2} f(\mathbf{x},s)ds, \end{aligned} \tag{2.2}$$

where $\int_{\mathbb{R}^d} u(\mathbf{y},s)\gamma(\mathbf{y},\mathbf{x})d\mathbf{y}$ accounts for the rate at which particles arrive at location $\mathbf{x}$ from all other places and $-\int_{\mathbb{R}^d} u(\mathbf{x},s)\gamma(\mathbf{x},\mathbf{y})d\mathbf{y}$ is the rate at which particles leave location $\mathbf{x}$. The last term in (2.2), i.e., $\int_{t_1}^{t_2} f(\mathbf{x},s)ds$, accounts for the amount of particles produced or absorbed at location $\mathbf{x}$ over the time horizon $[t_1,t_2)$. With other words, the difference between the initial density $u(\mathbf{x},0)$ and the evolved density $u(\mathbf{x},t)$ after time $t>0$ (i.e., $t_1=0,\ t_2=t$) at a fixed location $\mathbf{x}$ is given by the difference of mass which arrived and left over the prescribed time horizon (i.e., the *flux density at location* $\mathbf{x}$) plus the mass which is produced by the external source f. See Figure 2.1 for an illustration of a nonlocal convection-diffusion process.
Some simple manipulations of equation (2.2) lead to an equivalent relation

$$\begin{aligned} &\frac{u(\mathbf{x},t_2)-u(\mathbf{x},t_1)}{t_2-t_1} = \\ &\qquad \frac{1}{t_2-t_1}\left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d}\left(u(\mathbf{y},s)\gamma(\mathbf{y},\mathbf{x})-u(\mathbf{x},s)\gamma(\mathbf{x},\mathbf{y})\right)d\mathbf{y}ds + \int_{t_1}^{t_2} f(\mathbf{x},s)ds\right). \end{aligned}$$

By formally taking the limit $t_2\to t_1$ we eventually arrive at the nonlocal evolution equation

$$u_t(\mathbf{x},t) = \int_{\mathbb{R}^d}\left(u(\mathbf{y},t)\gamma(\mathbf{y},\mathbf{x})-u(\mathbf{x},t)\gamma(\mathbf{x},\mathbf{y})\right)d\mathbf{y} + f(\mathbf{x},t), \tag{2.3}$$

where $u_t := \frac{\partial}{\partial t}u$ denotes the partial time derivative. Let us define the nonlocal operator

$$\mathcal{L}u(\mathbf{x},t) := \int_{\mathbb{R}^d}\left(u(\mathbf{y},t)\gamma(\mathbf{y},\mathbf{x})-u(\mathbf{x},t)\gamma(\mathbf{x},\mathbf{y})\right)d\mathbf{y}, \tag{2.4}$$

then equation (2.3) can be written more concisely as

$$u_t = \mathcal{L}u + f.$$

The operator $\mathcal{L}$ and with it the evolution equation (2.3) is deemed to be "nonlocal", since the evaluation of $\mathcal{L}u(\mathbf{x},t)$ at location $\mathbf{x}$ involves not only information of the density u at $\mathbf{x}$, but of all values in a neighborhood of fixed size.

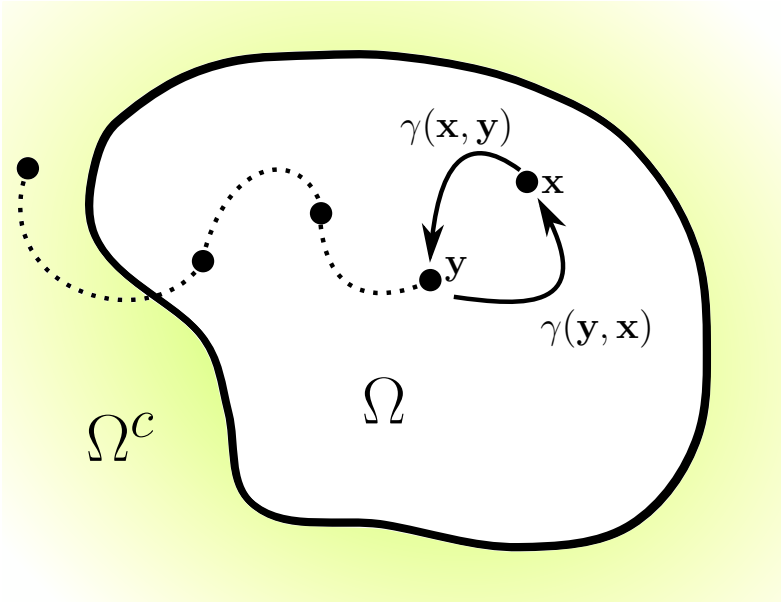


Figure 2.1: Illustration of a nonlocal convection-diffusion process.

By considering equation (2.2) on a domain $\Omega \subset \mathbb{R}^d$, not necessarily bounded, we can describe the evolution of the total mass by integrating the density over Ω. After some straightforward manipulations we arrive at

$$\int_\Omega (u(\mathbf{x},t) - u(\mathbf{x},0))d\mathbf{x} = \int_0^t \left(\int_\Omega \int_{\mathbb{R}^d} (u(\mathbf{y},s)\gamma(\mathbf{y},\mathbf{x}) - u(\mathbf{x},s)\gamma(\mathbf{x},\mathbf{y}))d\mathbf{y}d\mathbf{x} \right) ds + \int_0^t \int_\Omega f(\mathbf{x},s)d\mathbf{x}ds.$$

Since the integrand $(\mathbf{x},\mathbf{y}) \mapsto (u(\mathbf{y},s)\gamma(\mathbf{y},\mathbf{x}) - u(\mathbf{x},s)\gamma(\mathbf{x},\mathbf{y}))$ is an antisymmetric function, we find the identity

$$\int_\Omega \int_\Omega (u(\mathbf{y},s)\gamma(\mathbf{y},\mathbf{x}) - u(\mathbf{x},s)\gamma(\mathbf{x},\mathbf{y}))d\mathbf{y}d\mathbf{x} = 0$$

leading to

$$\int_\Omega (u(\mathbf{x},t) - u(\mathbf{x},0))d\mathbf{x} = \int_0^t \left(\int_\Omega \int_{\Omega^c} (u(\mathbf{y},s)\gamma(\mathbf{y},\mathbf{x}) - u(\mathbf{x},s)\gamma(\mathbf{x},\mathbf{y}))d\mathbf{y}d\mathbf{x} \right) ds + \int_0^t \int_\Omega f(\mathbf{x},s)d\mathbf{x}ds. \tag{2.5}$$

The expression

$$\int_\Omega \int_{\Omega^c} (u(\mathbf{y},s)\gamma(\mathbf{y},\mathbf{x}) - u(\mathbf{x},s)\gamma(\mathbf{x},\mathbf{y}))d\mathbf{y}d\mathbf{x} \tag{2.6}$$

can be considered the *nonlocal flux out of* Ω *into* Ω^c *at time* s. Assuming the absence of external forces, i.e., $f = 0$, and a zero flux out of Ω we find that

$$\int_\Omega (u(\mathbf{x}, t) - u(\mathbf{x}, 0))d\mathbf{x} = 0, \tag{2.7}$$

meaning that the total mass is preserved as time evolves in the prescribed situation. We return to this notion in Section 2.2 within the framework of a nonlocal vector calculus and relate this to partial differential equations in order to define a reasonable *nonlocal* divergence operator.
Without assuming any symmetry properties for the kernel γ such nonlocal evolution equations can be used to describe convective-diffusive processes [10, 42]. In fact, if γ is not symmetric then there are pairs of locations $(\mathbf{x}, \mathbf{y})$ for which particles tend to move more likely from one to the other place than going the opposite direction, thereby inducing convective effects.

2.2 A nonlocal vector calculus

We introduce a nonlocal vector calculus and largely follow the exposition in [37]. The key notion here is that of a nonlocal divergence operator $\mathcal{D}$, which also leads to a nonlocal gradient operator $\mathcal{G}$. In combination with appropriate diffusion and convection coefficients, we can then use these operators as building blocks to design general nonlocal operators which mimic classical differential operators. The ensuing structural insight into the governing operator also enables us to set up a variational framework by introducing the concept of weak solutions and thus positioning the finite element method as an accessible discretization approach.
The basis is given by the (classical) Gauss theorem

$$\int_\Omega \operatorname{div} \mathbf{q} \; d\mathbf{x} = \int_{\partial\Omega} \mathbf{q}^T \mathbf{n} \; ds =: F^{loc}(\Omega, \Omega^c; \mathbf{q}), \tag{2.8}$$

which states an important relation between the classical divergence operator $\operatorname{div} \mathbf{q}$ and the (local) flux $F^{loc}(\Omega, \Omega^c; \mathbf{q}) = \int_{\partial\Omega} \mathbf{q}^T \mathbf{n} ds$, where $\mathbf{q}\colon \mathbb{R}^d \to \mathbb{R}^d$ denotes a sufficiently smooth vector field and $\mathbf{n}\colon \partial\Omega \to \mathbb{R}^d$ the outer normal vector field of Ω. Assume we have $F^{nl}(\Omega, \Omega^c; \boldsymbol{\nu})$ accounting for a nonlocal flux (see, e.g., (2.6)), then we can define the nonlocal divergence as the *linear* operator $\mathcal{D}$, which satisfies a nonlocal version of the Gauss theorem, namely

$$\int_\Omega \mathcal{D}\boldsymbol{\nu} d\mathbf{x} = F^{nl}(\Omega, \Omega^c; \boldsymbol{\nu}), \tag{2.9}$$

where $\boldsymbol{\nu}\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$. Thus, the important underpinning for the development of a nonlocal vector calculus is the notion of an appropriate nonlocal flux. That is why we start from abstract balance laws and then review several notions related to classical differential equations so as to provide a comparison context for the subsequent development of the nonlocal convection-diffusion model.

2.2.1 An abstract balance law

Let $F(\Omega_1, \Omega_2; u)$ account for the flux of a density function $u\colon \Omega \times [0,\infty) \to \mathbb{R}$ out of Ω_1 into Ω_2, where $\Omega_1, \Omega_2 \subset \Omega \subset \mathbb{R}^d$ are two disjoint domains. Let $f\colon \Omega \to \mathbb{R}$ account for an external source density, then based on (2.5) and (2.6) we formulate an abstract balance law as follows

$$\frac{\partial}{\partial t}\int_{\widetilde{\Omega}} u(\mathbf{x},t)d\mathbf{x} = \int_{\widetilde{\Omega}} f(\mathbf{x},t)d\mathbf{x} - F(\widetilde{\Omega}, \widetilde{\Omega}^c; u) \quad \text{for all } \ \widetilde{\Omega} \subset \Omega, t > 0. \tag{2.10}$$

In words, equation (2.10) states that the temporal rate of change of the amount of u, i.e., $\frac{\partial}{\partial t}\int_{\widetilde{\Omega}} u(\mathbf{x},t)d\mathbf{x}$, is given by the amount of u created within $\widetilde{\Omega}$ by the external source f minus the flux of u out of $\widetilde{\Omega}$. If $f = 0$ and $F(\widetilde{\Omega}, \widetilde{\Omega}^c; u) = 0$, then

$$\int_{\widetilde{\Omega}} u(\mathbf{x},0)d\mathbf{x} = \int_{\widetilde{\Omega}} u(\mathbf{x},t)d\mathbf{x} \quad \text{for all } \ t > 0,$$

which means that when there are no sources of u in $\widetilde{\Omega}$ and there is no flux of u out of $\widetilde{\Omega}$ then the total mass of u is preserved; compare to (2.7).

2.2.2 The local case

We start with the notion of a local flux; see also (2.8). For this purpose, let $\mathbf{q}\colon \mathbb{R}^d \to \mathbb{R}^d$ denote a sufficiently smooth vector field, $\partial\Omega_{12} := \overline{\Omega}_1 \cap \overline{\Omega}_2$ the common boundary of two disjoint domains $\Omega_1, \Omega_2 \subset \Omega$ and let $\mathbf{n}_i\colon \partial\Omega_i \to \mathbb{R}^d$ denote the outer normal vector field of Ω_i. Then

$$F^{loc}(\Omega_1, \Omega_2; \mathbf{q}) := \int_{\partial\Omega_{12}} \mathbf{q}^T\mathbf{n}_1 ds \tag{2.11}$$

is called the *local flux out of* Ω_1 *into* Ω_2 and $\mathbf{x} \mapsto \mathbf{q}(\mathbf{x})^T\mathbf{n}_1(\mathbf{x}) = -\mathbf{q}(\mathbf{x})^T\mathbf{n}_2(\mathbf{x})$ is referred to as *flux density along* $\partial\Omega_{12}$. The flux accounts for interactions between the two domains Ω_1 and Ω_2 which only occur across their common boundary. In Figure 2.2 this situation is pictured. In case that these two regions are separated, so that $\partial\Omega_{12} = \emptyset$, the flux is zero which is why it is deemed to be *local*. The vector field $\mathbf{q} = \mathbf{q}(u)$ is related to some density function u through a physical law, such as Fourier's heat law. Since $\mathbf{n}_1 = -\mathbf{n}_2$ and $\partial\Omega_{12} = \partial\Omega_{21}$, the local flux satisfies the *action-reaction principle*

$$F^{loc}(\Omega_1, \Omega_2; \mathbf{q}) = \int_{\partial\Omega_{12}} \mathbf{q}^T\mathbf{n}_1 ds = -\int_{\partial\Omega_{21}} \mathbf{q}^T\mathbf{n}_2 ds = -F^{loc}(\Omega_2, \Omega_1; \mathbf{q}). \tag{2.12}$$

With other words, the flux out of Ω_1 into Ω_2 is the negative of the flux out of Ω_2 into Ω_1.

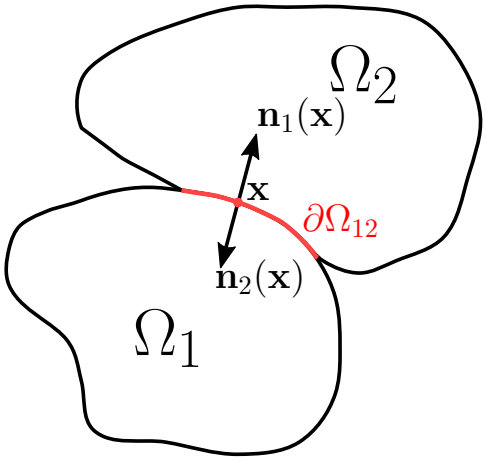


Figure 2.2: Illustration of a local flux. The two domains Ω_1 and Ω_2 only interact across their common boundary $\partial\Omega_{12}$. For $\mathbf{x} \in \partial\Omega_{12}$ the relation $\mathbf{n}_1(\mathbf{x}) = -\mathbf{n}_2(\mathbf{x})$ holds for the respective outer normal vector fields.

Now, by inserting the local flux (2.11) with $\mathbf{q} = \mathbf{q}(u)$ into the balance law (2.10) and applying the classical Gauss theorem (2.8) we find

$$\frac{\partial}{\partial t}\int_{\widetilde{\Omega}} u(\mathbf{x},t)d\mathbf{x} = \int_{\widetilde{\Omega}} f(\mathbf{x},t)d\mathbf{x} - \int_{\widetilde{\Omega}} \operatorname{div}\mathbf{q}(u)d\mathbf{x} \quad \text{for all} \quad \widetilde{\Omega} \subset \Omega, t > 0.$$

Since $\widetilde{\Omega} \subset \Omega$ is arbitrary, we obtain the *local evolution equation*

$$u_t + \operatorname{div}\mathbf{q}(u) = f \quad \text{in } \Omega.$$

For example, Fourier's heat law postulates

$$\mathbf{q}(u) = -\mathbf{D}\nabla u, \tag{2.13}$$

where the diffusivity $\mathbf{D}\colon \Omega \to \mathbb{R}^{d\times d}_{spd}$ is given by a symmetric and positive definite matrix. In this case, we obtain the heat equation

$$u_t - \operatorname{div}\mathbf{D}\nabla u = f \quad \text{in } \Omega.$$

Additional convective effects can be included into the model by considering

$$u_t - \operatorname{div}\mathbf{D}\nabla u + \operatorname{div}(\mathbf{v}u) = f \quad \text{in } \Omega, \tag{2.14}$$

where $\mathbf{v}\colon \Omega \to \mathbb{R}^d$ is referred to as *velocity vector field*.

2.2.3 The nonlocal case

For the local case, we have discussed local fluxes, local balance laws and local convection-diffusion. We next go through the same list for the nonlocal case, mimicking at every step

what we did for the local case. We also consider additional notions such as a nonlocal divergence operator and its adjoint operator as well as a nonlocal vector calculus.
Let us take up the considerations presented in the introductory Section 2.1, in particular the therein derived flux term (2.6). Recalling our intuition that particles may jump we have that an appropriate nonlocal flux $F^{nl}(\Omega_1, \Omega_2; u)$ must relate all points in Ω_1 to all points in Ω_2. Therefore, let $\psi = \psi(u)\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be a two-point function, then we denote by

$$F^{nl}(\Omega_1, \Omega_2; u) = \int_{\Omega_1} \int_{\Omega_2} \psi(u)(\mathbf{x}, \mathbf{y}) d\mathbf{y} d\mathbf{x} \tag{2.15}$$

the *nonlocal flux of u out of Ω_1 into Ω_2* and $\mathbf{x} \mapsto \int_{\Omega_2} \psi(\mathbf{x}, \mathbf{y}) d\mathbf{y}$ denotes the *flux density from $\mathbf{x} \in \Omega_1$ into Ω_2*. See Figure 2.3 for an illustration of the nonlocal flux. We note that the flux given in (2.6) is a special case of (2.15) with

$$\psi(u) := u'\gamma' - u\gamma.$$

The flux is nonlocal because it may be nonzero even if the closures of Ω_1 and Ω_2 have an empty intersection. This is in stark contrast to classical local interactions for which the flux between Ω_1 and Ω_2 vanishes if their common boundary is empty.

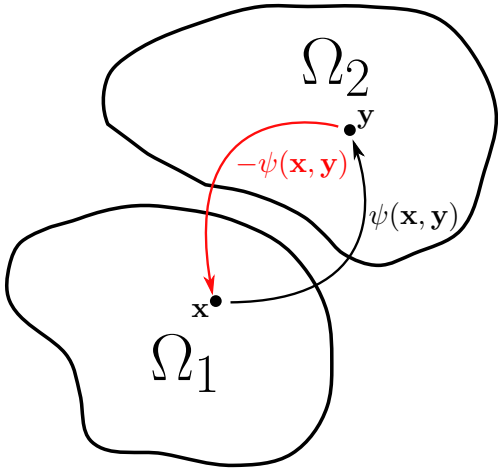


Figure 2.3: Illustration of a nonlocal flux. All points of the two domains Ω_1 and Ω_2 are mutually related through the antisymmetric function ψ. Therefore, the nonlocal flux $F^{nl}(\Omega_1, \Omega_2; u)$ (2.15) may be nonzero despite the fact that $\Omega_1 \cap \Omega_2 = \emptyset$.

Imitating (2.12) we now require that the nonlocal flux fulfills a *nonlocal action-reaction principle*

$$F^{nl}(\Omega_1, \Omega_2; u) = -F^{nl}(\Omega_2, \Omega_1; u). \tag{2.16}$$

It is straightforward to show, that the following are equivalent:

i) The *nonlocal action-reaction principle* (2.16) holds for all $\Omega_1, \Omega_2 \subset \Omega$.

ii) The two-point function ψ is *antisymmetric* (almost sure) on $\Omega \times \Omega$.

iii) There are *no self-interactions*, i.e., $\int_{\widetilde{\Omega}} \int_{\widetilde{\Omega}} \psi(\mathbf{x}, \mathbf{y}) d\mathbf{y} d\mathbf{x} = 0$ for all $\widetilde{\Omega} \subset \Omega$.

A nonlocal divergence and gradient operator

Inserting the definition of the nonlocal flux (2.15) into the nonlocal version of the Gauss theorem (2.9) gives

$$\int_{\widetilde{\Omega}} \mathcal{D}\boldsymbol{\nu} d\mathbf{x} = F^{nl}(\widetilde{\Omega}, \widetilde{\Omega}^c; \boldsymbol{\nu}) = \int_{\widetilde{\Omega}} \int_{\widetilde{\Omega}^c} \psi(\boldsymbol{\nu})(\mathbf{x}, \mathbf{y}) d\mathbf{y} d\mathbf{x} \quad \text{for all } \widetilde{\Omega} \subset \Omega.$$

Here, $\boldsymbol{\nu} = \boldsymbol{\nu}(u) \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ is a vector valued two-point function and mimics the role of $\mathbf{q} = \mathbf{q}(u)$. Similarly to (2.13) the relation $\boldsymbol{\nu}(u)$ is explained by postulating a physical law. Since we require a nonlocal action-reaction principle to hold so that there are no self-interactions, we find

$$\int_{\widetilde{\Omega}} \mathcal{D}\boldsymbol{\nu} d\mathbf{x} = \int_{\widetilde{\Omega}} \int_{\mathbb{R}^d} \psi(\boldsymbol{\nu})(\mathbf{x}, \mathbf{y}) d\mathbf{y} d\mathbf{x} \quad \text{for all } \widetilde{\Omega} \subset \Omega.$$

Since $\widetilde{\Omega} \subset \Omega$ is arbitrary we therefore obtain

$$\mathcal{D}\boldsymbol{\nu}(\mathbf{x}) = \int_{\mathbb{R}^d} \psi(\boldsymbol{\nu})(\mathbf{x}, \mathbf{y}) d\mathbf{y} \quad \text{for all } \mathbf{x} \in \Omega.$$

Let us recapitulate our steps until now towards an appropriate definition of a nonlocal divergence operator. After introducing the notion of a nonlocal flux, we have first required that the nonlocal flux satisfies an action-reaction principle provoking that the integrand ψ is antisymmetric. Then, we have further required a nonlocal Gauss theorem to hold, meaning that the integral over the nonlocal divergence $\mathcal{D}\boldsymbol{\nu}$ of a vector $\boldsymbol{\nu}$ over any domain $\widetilde{\Omega} \subset \Omega$ shall equal the nonlocal flux of $\boldsymbol{\nu}$ out of that domain. We complete this list by thirdly postulating that the nonlocal divergence operator $\mathcal{D}$ is linear in $\boldsymbol{\nu}$ as is the case for the classical divergence operator. A definition satisfying all three requirements is given by

$$\psi(\boldsymbol{\nu}) := (\boldsymbol{\nu} + \boldsymbol{\nu}')^T \boldsymbol{\alpha} \tag{2.17}$$

for some *antisymmetric* vector-valued two-point function $\boldsymbol{\alpha} \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, resulting in

$$\mathcal{D}\boldsymbol{\nu}(\mathbf{x}) = \mathcal{D}_\alpha \boldsymbol{\nu}(\mathbf{x}) = \int_{\mathbb{R}^d} (\boldsymbol{\nu}(\mathbf{x}, \mathbf{y}) + \boldsymbol{\nu}(\mathbf{y}, \mathbf{x}))^T \boldsymbol{\alpha}(\mathbf{x}, \mathbf{y}) d\mathbf{y} \quad \text{for all } \mathbf{x} \in \Omega. \tag{2.18}$$

In [8], the authors derive a more general definition. They show that the three assumptions just mentioned on the nonlocal divergence operator allow the application of the Schwartz kernel theorem. Our definition (2.18) can be considered a special case, but is sufficiently general for our needs.
Finally we define a nonlocal gradient operator. Due to the classical duality relation

$\nabla = -\operatorname{div}^*$ with respect to the L^2 inner product, it stands to reason to define the corresponding nonlocal gradient operator as $\mathcal{G} := -\mathcal{D}^*$. Thus $\mathcal{G}$ is uniquely determined through

$$(u, \mathcal{D}\boldsymbol{\nu})_{L^2(\mathbb{R}^d)} = (\mathcal{G}u, \boldsymbol{\nu})_{L^2(\mathbb{R}^d \times \mathbb{R}^d)},$$

for all $u\colon \mathbb{R}^d \to \mathbb{R}$ and $\boldsymbol{\nu}\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^k$. We find

$$\mathcal{G}(u)(\mathbf{x}, \mathbf{y}) = -(u(\mathbf{x}) - u(\mathbf{y}))\boldsymbol{\alpha}(\mathbf{x}, \mathbf{y}) \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^d. \tag{2.19}$$

Similar to their local counterparts, $\mathcal{D}$ is scalar-valued and $\mathcal{G}$ is vector-valued. Also, constant functions are contained in the kernel of the nonlocal gradient, since obviously $\mathcal{G}(u) = 0$ for $u = const$.

Nonlocal interaction domains and integral theorems

In the classical integral theorems we find boundary integrals accounting for interactions. Since interactions may occur at distance in the nonlocal case, i.e., points in Ω may now interact with points in $\overline{\Omega}^c$, we find double integrals in the nonlocal integral theorems accounting for these interactions.
Let $\Omega \subset \mathbb{R}^d$ be an open set. Then for a function $\boldsymbol{\alpha}\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ we define the corresponding *interaction domain* Ω_I by

$$\Omega_I := \{\mathbf{y} \in \Omega^c\colon \boldsymbol{\alpha}(\mathbf{x}, \mathbf{y}) \neq 0 \text{ for some } \mathbf{x} \in \Omega\},$$

so that Ω_I consists of those points outside of Ω that interact with points in Ω. We find that Ω_I mimics the role of the boundary $\partial\Omega$ in the following integral theorems.
We have defined the nonlocal divergence operator $\mathcal{D}$ in such a way that a nonlocal analogue of the classical Gauss theorem holds, specifically,

$$\int_\Omega \mathcal{D}\boldsymbol{\nu} d\mathbf{x} = F^{nl}(\Omega, \Omega^c; \boldsymbol{\nu}),$$

where F^{nl} is given in (2.15) with ψ as in (2.17). By denoting the flux density from $\mathbf{x} \in \Omega$ into Ω_I by

$$\mathcal{N}(\boldsymbol{\nu})(\mathbf{x}) := -\int_{\Omega \cup \Omega_I} (\boldsymbol{\nu} + \boldsymbol{\nu}')^T \boldsymbol{\alpha}\, d\mathbf{y} \quad \text{for all} \quad \mathbf{x} \in \Omega_I,$$

we can write this more concisely as

$$\int_\Omega \mathcal{D}\boldsymbol{\nu} d\mathbf{x} = \int_{\Omega_I} \mathcal{N}\boldsymbol{\nu}(\mathbf{x}) d\mathbf{x}.$$

We refer to this equation, which trivially holds by definition of $\mathcal{N}$ and $\mathcal{D}$, as *nonlocal Gauss theorem*. By applying this theorem to $(\mathbf{x}, \mathbf{y}) \mapsto v(\mathbf{x})\boldsymbol{\nu}(\mathbf{x}, \mathbf{y})$ we obtain the *nonlocal integration by parts formula*

$$\int_\Omega v\mathcal{D}(\boldsymbol{\nu})\, d\mathbf{x} = -\int_{\Omega \cup \Omega_I} \int_{\Omega \cup \Omega_I} \mathcal{G}(v)^T \boldsymbol{\nu}\, d\mathbf{y} d\mathbf{x} + \int_{\Omega_I} v\mathcal{N}(\boldsymbol{\nu})\, d\mathbf{x},$$

and with $\boldsymbol{\nu} := (\boldsymbol{\Theta}\mathcal{G}(u))$ as special case the *nonlocal Green's first identity*

$$\int_\Omega v\mathcal{D}(\boldsymbol{\Theta}\mathcal{G}(u))\,d\mathbf{x} = -\int_{\Omega\cup\Omega_I}\int_{\Omega\cup\Omega_I}\mathcal{G}(v)^T\boldsymbol{\Theta}\mathcal{G}(u)\,d\mathbf{y}d\mathbf{x} + \int_{\Omega_I} v\mathcal{N}(\boldsymbol{\Theta}\mathcal{G}(u))\,d\mathbf{x},$$

where $u, v\colon \mathbb{R}^d \to \mathbb{R}$ are scalar functions, $\boldsymbol{\nu}\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ a two-point vector function and $\boldsymbol{\Theta}\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ a two-point matrix function.

2.3 Nonlocal convection-diffusion equations

With the notion of a nonlocal divergence at hand we are in a position to formulate a nonlocal evolution equation in analogy to (2.14). Nonlocal balance laws now read as

$$\frac{\partial}{\partial t}\int_{\widetilde{\Omega}} u(\mathbf{x},t)\,dx = \int_{\widetilde{\Omega}} f\,d\mathbf{x} - \int_{\widetilde{\Omega}}\mathcal{D}(\boldsymbol{\nu})\,d\mathbf{x} \quad \text{for all} \quad \widetilde{\Omega}\subseteq\Omega.$$

Because $\widetilde{\Omega}\subseteq\Omega$ is arbitrary, we arrive at the *nonlocal evolution equation*

$$u_t + \mathcal{D}(\boldsymbol{\nu}) = f \qquad \text{for all} \quad \mathbf{x}\in\Omega,\ t>0.$$

A nonlocal convection-diffusion flux arises from the relation

$$\boldsymbol{\nu} = \boldsymbol{\nu}(u) = -\boldsymbol{\Theta}\mathcal{G}(u) + \boldsymbol{\mu}u,$$

where $\boldsymbol{\Theta}\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times d}_{spd}$ denotes a two-point matrix function, which we refer to as *nonlocal diffusion coefficient* and $\boldsymbol{\mu}\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ a two-point velocity vector field. This flux leads to the *nonlocal convection-diffusion equation*

$$u_t - \mathcal{D}(\boldsymbol{\Theta}\mathcal{G}(u)) + \mathcal{D}(\boldsymbol{\mu}u) = f \quad \text{for all} \quad \mathbf{x}\in\Omega,\ t>0. \tag{2.20}$$

Without loss of generality we can assume that $\boldsymbol{\Theta}$ and $\boldsymbol{\mu}$ are symmetric in $(\mathbf{x},\mathbf{y})$ [42, Section 1.1]. By inserting the definitions of $\mathcal{D}$ (2.18) and $\mathcal{G}$ (2.19) we find

$$\begin{aligned}
\mathcal{L}u(\mathbf{x}) &= \mathcal{D}(\boldsymbol{\Theta}\mathcal{G}u) - \mathcal{D}(\boldsymbol{\mu}u)\\
&=: \mathcal{L}^d u(\mathbf{x}) - \mathcal{L}^c u(\mathbf{x})\\
&= 2\int_{\mathbb{R}^d}(u(\mathbf{y})-u(\mathbf{x}))\boldsymbol{\alpha}^T(\boldsymbol{\Theta}\boldsymbol{\alpha})d\mathbf{y} - \int_{\mathbb{R}^d}(u(\mathbf{y})+u(\mathbf{x}))\boldsymbol{\mu}^T\boldsymbol{\alpha}d\mathbf{y}\\
&= \int_{\Omega\cup\Omega_I}\left(u(\mathbf{y})\gamma(\mathbf{y},\mathbf{x}) - u(\mathbf{x})\gamma(\mathbf{x},\mathbf{y})\right)d\mathbf{x}
\end{aligned} \tag{2.21}$$

with

$$\gamma = 2\boldsymbol{\alpha}^T(\boldsymbol{\Theta}\boldsymbol{\alpha}) + \boldsymbol{\mu}^T\boldsymbol{\alpha}\colon \mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}. \tag{2.22}$$

The representation of $\mathcal{L}$ in (2.21) coincides with the one given in (2.4) derived from the nonlocal conservation equation in the introductory Section 2.1. The kernel γ, given in

(2.22), consists of a symmetric component $2\boldsymbol{\alpha}^T(\boldsymbol{\Theta}\boldsymbol{\alpha})$, accounting for diffusive effects, and an antisymmetric component $\boldsymbol{\mu}^T\boldsymbol{\alpha}$ inducing convective behavior; explaining the superscripts "d" and "c" in (2.21).
On the other hand, starting from a model as in Section 2.1 with a given kernel γ (not necessarily a probability density function), we can decompose this kernel into its symmetric and antisymmetric parts $\gamma = \gamma^s + \gamma^a$, where $\gamma^s := \frac{\gamma+\gamma'}{2}$ and $\gamma^a := \frac{\gamma-\gamma'}{2}$. Then by defining

$$\boldsymbol{\alpha}(\mathbf{x},\mathbf{y}) := \frac{(\mathbf{x}-\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|_2}, \qquad \boldsymbol{\Theta}_\gamma(\mathbf{x},\mathbf{y}) := 2\gamma^s(\mathbf{x},\mathbf{y})\,\mathbf{Id}, \qquad \boldsymbol{\mu}_\gamma(\mathbf{x},\mathbf{y}) := \gamma^a(\mathbf{x},\mathbf{y})\boldsymbol{\alpha}(\mathbf{x},\mathbf{y}), \tag{2.23}$$

we find $\gamma = \gamma^s + \gamma^a = 2\boldsymbol{\alpha}^T(\boldsymbol{\Theta}\boldsymbol{\alpha}) + \boldsymbol{\mu}^T\boldsymbol{\alpha}$. Thus by fixing the gradient operator $\mathcal{D} = \mathcal{D}_{\boldsymbol{\alpha}}$, for any kernel γ, we can define appropriate convection-diffusion parameters $\boldsymbol{\mu}_\gamma$ and $\boldsymbol{\Theta}_\gamma$, such that we obtain the interpretation $\mathcal{L}_\gamma u = \mathcal{D}(\boldsymbol{\Theta}_\gamma \mathcal{G} u) - \mathcal{D}(\boldsymbol{\mu}_\gamma u)$. However, note that the mapping

$$(\boldsymbol{\alpha}, \boldsymbol{\Theta}, \boldsymbol{\mu}) \mapsto 2\boldsymbol{\alpha}^T(\boldsymbol{\Theta}\boldsymbol{\alpha}) + \boldsymbol{\mu}^T\boldsymbol{\alpha}$$

is surjective, but not injective. Consequently, different divergence operators $\mathcal{D}_{\boldsymbol{\alpha}}$ and convection parameters as well as (possibly anisotropic) diffusion coefficients may lead to the same kernel function γ.

2.3.1 The nonlocal Dirichlet problem

In this thesis we set a focus on the spatial discretization of nonlocal equations. Hence, throughout the remainder of this work *we from now on consider the stationary version* of (2.20) given by

$$-\mathcal{L}u(\mathbf{x}) = \int_{\mathbb{R}^d} (u(\mathbf{x})\gamma(\mathbf{x},\mathbf{y}) - u(\mathbf{y})\gamma(\mathbf{y},\mathbf{x}))d\mathbf{y} = f(\mathbf{x}) \quad \text{for all} \quad \mathbf{x} \in \Omega.$$

In general $-\mathcal{L}$ is not injective. If we consider for example divergence free convection parameters, i.e., $\mathcal{D}\boldsymbol{\mu} = 0$, then by the first identity in (2.21) and the linearity of $\mathcal{D}$ all constant functions are contained in the kernel of $\mathcal{L}$. Contrasting with partial differential equations it stands to reason to define nonlocal counterparts of boundary constraints. The role of the boundary is taken by the interaction domain Ω_I in the nonlocal setting and we define the *nonlocal Dirichlet problem with volume constraints* by

$$\begin{cases} -\mathcal{L}u(\mathbf{x}) = \displaystyle\int_{\mathbb{R}^d} (u(\mathbf{x})\gamma(\mathbf{x},\mathbf{y}) - u(\mathbf{y})\gamma(\mathbf{y},\mathbf{x}))d\mathbf{y} = f & \text{on } \Omega \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad u = g & \text{on } \Omega_I. \end{cases}$$

We point out that due to the kernel representation $\gamma = 2\boldsymbol{\alpha}^T(\boldsymbol{\Theta}\boldsymbol{\alpha}) + \boldsymbol{\mu}^T\boldsymbol{\alpha}$ we can rewrite the interaction domain as

$$\Omega_I = \{\mathbf{y} \in \Omega^c \colon \gamma(\mathbf{x},\mathbf{y}) \neq 0 \text{ for some } \mathbf{x} \in \Omega\}.$$

These kind of constraints are in line with our intuition that particles are allowed to jump and nonlocal interactions therefore also occur over distance. In the homogeneous case, i.e., $g = 0$, the interaction domain Ω_I resembles a hostile environment, meaning that particles which jump outside Ω are absorbed immediately; see also [10, Chapter 2].

Remark 2.3.1. *Throughout the remainder of this thesis we only operate on the level of the kernel, i.e., further specializations of the nonlocal operator such as appropriate conditions for well-posedness are realized through requirements imposed on the kernel. However, due to relation* (2.23) *we can always interpret the symmetric part of the kernel as nonlocal diffusion and the antisymmetric part as nonlocal convection.*

Chapter 3

Variational formulation and extended well-posedness results

We introduce a weak formulation of the problem (2.1) and derive results about the existence and uniqueness of such a weak solution. In [36, 42] and [10, Chapter 2 and 4] such problems have been analyzed with regard to well-posedness statements for several choices of kernels. We focus on two example classes of truncated kernels as they are introduced in [36].
The first one consists of integrable, not necessarily symmetric kernels. A kernel γ is called integrable, if the integral $\int \gamma(\mathbf{x}, \mathbf{y}) d\mathbf{y}$ is finite for all $\mathbf{x} \in \Omega$. In [36] well-posedness is shown for the symmetric case. We extend these results to the nonsymmetric case by adding additional conditions on the antisymmetric (convective) part of the kernel. Similar conditions have been derived in [42] but on the level of the operators appearing within the nonlocal vector calculus. The second class consists of singular and symmetric kernels, which contains as special case the fractional kernel.
In the above mentioned literature, nonlocal interactions are limited to occur up to a finite distance, such that the kernels of interest are truncated by the Euclidean ball. This is motivated by the fact, that there are no interactions over infinite distance in real world applications. The case of infinite interactions, i.e., unbounded interaction sets, is not included in the following well-posedness theory but can be found, e.g., in [10, Chapter 2 and 4] or [2].
In this chapter we extend the well-posedness results to general *interaction sets*. We impose appropriate conditions on these sets to rely on existing results holding for the standard case, i.e., truncation by Euclidean balls. A clearly determined support of the kernel is important for numerical computations. Instead of restricting ourselves to nonlocal interactions occurring within Euclidean balls, we can now consider interaction phenomena described through, e.g., general (semi-)norm induced balls, given by

$$B_{\delta,\bullet}(\mathbf{x}) := \left\{ \mathbf{y} \in \mathbb{R}^d \colon \|\mathbf{y} - \mathbf{x}\|_\bullet := \|\mathbf{A}^{\frac{1}{2}}(\mathbf{y} - \mathbf{x})\|_p < \delta \right\},$$

where $\mathbf{A} \in \mathbb{R}^{d \times d}_{spd}$ denotes a symmetric and positive definite matrix, $0 < p \leq +\infty$ and $\delta > 0$. Thus including d-dimensional possibly transformed convex norm balls ($p \geq 1$)

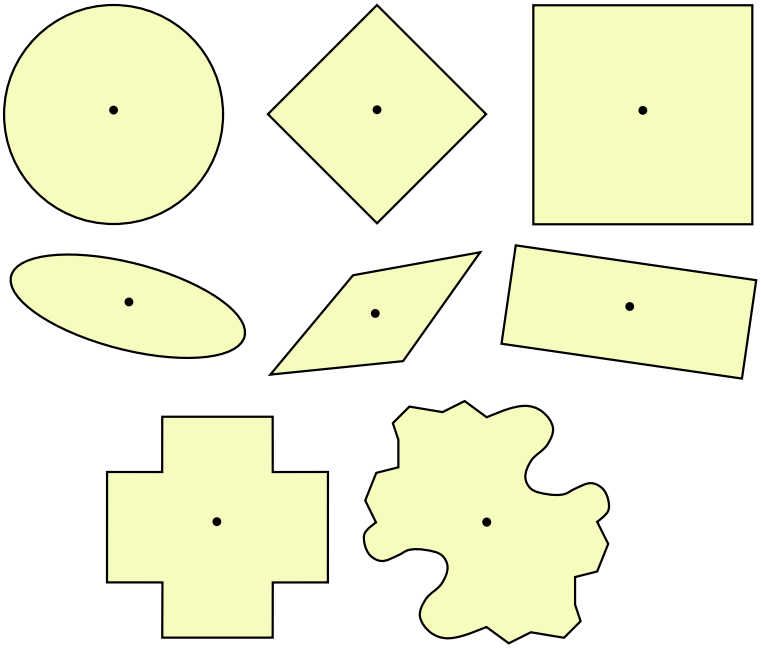


Figure 3.1: Different examples of interaction sets. For instance, in the top row we find from left to right the $\|\cdot\|_2$-, $\|\cdot\|_1$-, and $\|\cdot\|_\infty$-ball, respectively. In the second row these norms are additionally transformed by a suitable symmetric and positive definite matrix. In the bottom row there are two examples of interaction sets which are not induced by a norm.

and nonconvex seminorm balls ($p < 1$) as wells as possibly rotated hyperrectangles ($p = +\infty$). Special cases of interest are the balls $B_{\delta,2}, B_{\delta,\infty}$ and $B_{\delta,1}$, induced by the Euclidean, supremum and Manhattan norm, respectively. Also, we can consider any translation invariant family $S(\mathbf{x}) := \mathbf{x} + S$, where $S \subset \mathbb{R}^d$ is a bounded set satisfying $S = -S$. Further examples are provided in Section 3.2. In Figure 3.1 we find several considerable interaction sets.

This chapter is organized as follows. We first introduce the weak formulation in Section 3.1 including the definition of a nonlocal bilinear form. Then, after setting up some general assumptions on the interaction sets and the kernel in Section 3.2 we provide well-posedness results in Section 3.3 for the two classes of kernels mentioned above. For these results we apply the Lax-Milgram theorem.

3.1 Weak formulation

For the purpose of defining an appropriate weak formulation of problem (2.1), let us separate the variables of a potential solution $u\colon \Omega \cup \Omega_I \to \mathbb{R}$ by

$$u = (u_\Omega, u_{\Omega_I}),$$

where $u_\Omega := u|_\Omega$ and $u_{\Omega_I} := u|_{\Omega_I}$. According to this decomposition we can also split the nonlocal convection-diffusion operator $\mathcal{L}$ into two components, one acting on u_Ω and one acting on u_{Ω_I}. More precisely, let $\mathbf{x} \in \Omega$, then

$$\begin{aligned} -\mathcal{L}u(\mathbf{x}) &= \int_{\Omega\cup\Omega_I} (u\gamma - u'\gamma')d\mathbf{y} \\ &= \underbrace{\int_\Omega (u_\Omega\gamma - u'_\Omega\gamma')d\mathbf{y} + u_\Omega \int_{\Omega_I} \gamma d\mathbf{y}}_{=:-\mathcal{L}_\Omega u(\mathbf{x})} \underbrace{- \int_{\Omega_I} u'_{\Omega_I}\gamma' d\mathbf{y}}_{=:\mathcal{L}_{\Omega_I} u(\mathbf{x})}. \end{aligned} \tag{3.1}$$

Since we require $u_{\Omega_I} = g$, integrals over Ω_I only involve the given Dirichlet data and the first equation in (2.1) given by $-\mathcal{L}u = f$ is equivalent to

$$-\mathcal{L}_\Omega u_\Omega = f - \mathcal{L}_{\Omega_I} g =: \tilde{f}. \tag{3.2}$$

Now we require this equivalently reformulated equation (3.2) to hold in a weak sense by testing with functions $v\colon \Omega \cup \Omega_I \to \mathbb{R}$ that satisfy the constraint $v = 0$ on Ω_I and we obtain

$$\int_\Omega (-\mathcal{L}_\Omega u_\Omega)\, v d\mathbf{x} = \int_\Omega (f - \mathcal{L}_{\Omega_I} g)\; v\, d\mathbf{x}. \tag{3.3}$$

We define the linear functional

$$\ell(v) := \ell_{f,g}(v) := \int_\Omega f v\, d\mathbf{x} - \int_\Omega \mathcal{L}_{\Omega_I} g\; v\, d\mathbf{x} \tag{3.4}$$

and the *nonlocal bilinear form*

$$A(u,v) := \int_\Omega (-\mathcal{L}_\Omega u)\, v d\mathbf{x} = \int_\Omega v \int_\Omega \left(u\gamma - u'\gamma'\right) d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega_I} \gamma\; d\mathbf{y} d\mathbf{x}. \tag{3.5}$$

Then equation (3.3) reads as

$$A(u,v) = \ell(v).$$

We note that $A(u,v)$ only involves function evaluations of u and v on Ω. We now define the corresponding "energy" seminorm

$$|||u||| := \sqrt{A(u,u)}. \tag{3.6}$$

Further, we define the *nonlocal energy space* and *nonlocal volume-constrained energy space* by

$$\begin{cases} V(\Omega\cup\Omega_I) := \{u \in L^2(\Omega\cup\Omega_I) \;:\; \|u\|_{V(\Omega\cup\Omega_I)} := |||u||| + \|u\|_{L^2(\Omega\cup\Omega_I)} < \infty\} \\ V_c(\Omega\cup\Omega_I) := \{u \in V(\Omega\cup\Omega_I) \;:\; u = 0 \text{ on } \Omega_I\}, \end{cases}$$

respectively. In the subsequent program we consider kernels for which $|||\cdot|||$ defines a norm on $V_c(\Omega\cup\Omega_I)$ implying in particular that $A(u,u)$ is nonnegative for $u \in V_c(\Omega\cup\Omega_I)$.

We denote by $V_c^*(\Omega)$ the dual space of $V_c(\Omega \cup \Omega_I)$ with respect to the standard $L^2(\Omega)$ duality pairing and also define the volume "trace" space

$$\widetilde{V}(\Omega_I) := \{v|_{\Omega_I} \,:\, v \in V(\Omega \cup \Omega_I)\}.$$

We can finally define a weak formulation of (2.1) as

$$\begin{aligned} &\textit{given } f \in V_c'(\Omega) \textit{ and } g \in \widetilde{V}(\Omega_I), \textit{ find } u^0 \in V_c(\Omega \cup \Omega_I) \textit{ such that} \\ &\qquad A(u^0, v) = \ell(v) \quad \textit{for all} \quad v \in V_c(\Omega \cup \Omega_I). \\ &\textit{Then } u := (u^0_\Omega, g) \in V(\Omega \cup \Omega_I) \textit{ is called the weak solution of } (2.1). \end{aligned} \tag{3.7}$$

Remark 3.1.1. *For $u \in V_c(\Omega \cup \Omega_I)$, so that $u = 0$ on Ω_I, we obtain $-\mathcal{L}u = -\mathcal{L}_\Omega u$. Thus u^0 in (3.7) can be interpreted as the weak solution of the homogeneous auxiliary problem*

$$\begin{cases} -\mathcal{L}u = \tilde{f} & \text{on } \Omega \\ \quad u = 0 & \text{on } \Omega_I. \end{cases}$$

Assumed (3.7) admits a unique weak solution $u^0 = (u^0_\Omega, 0) \in V_c(\Omega \cup \Omega_I)$, then due to (3.2) the function $u := (u^0_\Omega, g) \in V(\Omega \cup \Omega_I)$ represents the unique weak solution of the inhomogeneous problem (2.1).

We want to point out that the nonlocal bilinear form does not have to be symmetric, since we do not require the kernel to be symmetric. However, let us in greater detail illuminate some inherent structures. For this purpose, we introduce its symmetric and antisymmetric part, namely

$$A^s := \frac{1}{2}(A + A') \qquad \text{and} \qquad A^a := \frac{1}{2}(A - A').$$

Then we find $A^a(u,u) = 0$ which implies $A(u,u) = A^s(u,u)$, so that

$$|||u||| = \sqrt{A(u,u)} = \sqrt{A^s(u,u)} \quad \text{for all} \quad u \in V(\Omega \cup \Omega_I).$$

Thus $A^s \colon V(\Omega \cup \Omega_I) \times V(\Omega \cup \Omega_I) \to \mathbb{R}$ defines the symmetric bilinear form associated to the energy norm. Let us now derive precise representations of A^s and A^a. To this end we consider the composition of the kernel into its symmetric and antisymmetric part, i.e., $\gamma = \gamma^s + \gamma^a$. Inserted into (3.5) yields

$$\begin{aligned} A(u,v) &= \int_\Omega v \int_\Omega (u\gamma - u'\gamma')\, d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega_I} \gamma d\mathbf{y} d\mathbf{x} \\ &= \underbrace{\int_\Omega v \int_\Omega (u - u')\, \gamma^s d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega_I} \gamma^s d\mathbf{y} d\mathbf{x}}_{=:A^d(u,v)} \\ &\qquad + \underbrace{\int_\Omega v \int_\Omega (u + u')\, \gamma^a d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega_I} \gamma^a d\mathbf{y} d\mathbf{x}}_{=:A^c(u,v)}. \end{aligned} \tag{3.8}$$

As pointed out in Section 2.3 and in particular (2.21), the bilinear form $A^d(u,v) = (-\mathcal{L}^d_\Omega u, v)_{L^2(\Omega)}$ accounts for **d**iffusion, whereas $A^c(u,v) = (-\mathcal{L}^c_\Omega u, v)_{L^2(\Omega)}$ takes into account **c**onvective effects. If the kernel is symmetric, i.e., $\gamma = \gamma^s$ and $\gamma^a = 0$, we obtain $A = A^d$. By the symmetry of γ^s and Fubini's theorem we further find that the first term in $A^d(u,v)$ equals

$$\int_\Omega v \int_\Omega (u - u')\, \gamma^s d\mathbf{y} d\mathbf{x} = -\int_\Omega \int_\Omega v'\, (u - u')\, \gamma^s d\mathbf{y} d\mathbf{x}$$

and therefore

$$A^d(u,v) = \frac{1}{2} \int_\Omega \int_\Omega (v - v')\, (u - u')\, \gamma^s d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega_I} \gamma^s d\mathbf{y} d\mathbf{x}, \tag{3.9}$$

implying that $A^d \colon V(\Omega \cup \Omega_I) \times V(\Omega \cup \Omega_I) \to \mathbb{R}$ is symmetric. For $u, v \in V_c(\Omega \cup \Omega_I)$ so that $u, v = 0$ on Ω_I we can add some appropriate zero terms to arrive at the representation

$$A^d(u,v) = \frac{1}{2} \int_{\Omega \cup \Omega_I} \int_{\Omega \cup \Omega_I} (v - v')\, (u - u')\, \gamma^s d\mathbf{y} d\mathbf{x}, \tag{3.10}$$

which is ubiquitous in the related literature; see, e.g., [36, 38, 85]. While A^d is symmetric we find that A^c is neither symmetric nor antisymmetric. By splitting the first term in A^c we find

$$A^c(u,v) = \int_\Omega v \int_\Omega u' \gamma^a d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega \cup \Omega_I} \gamma^a d\mathbf{y} d\mathbf{x}.$$

By the antisymmetry of γ^a and Fubini's theorem we find for the first term

$$\int_\Omega v \int_\Omega u' \gamma^a d\mathbf{y} d\mathbf{x} = \int_\Omega \int_\Omega v' u (\gamma^a)' d\mathbf{y} d\mathbf{x} = -\int_\Omega u \int_\Omega v' \gamma^a d\mathbf{y} d\mathbf{x}$$

implying its antisymmetry in u and v. Since the second term of A^c is obviously symmetric in u and v we find that the symmetric part $A^{c,s}$ of A^c is given by

$$A^{c,s}(u,v) = \int_\Omega vu \int_{\Omega \cup \Omega_I} \gamma^a d\mathbf{y} d\mathbf{x}$$

and the antisymmetric part $A^{c,a}$ of A^c by

$$A^{c,a}(u,v) = \int_\Omega v \int_\Omega u' \gamma^a d\mathbf{y} d\mathbf{x}.$$

Finally we are in a position to derive precise representations for the symmetric and antisymmetric part of the original bilinear form A. Due to $A = A^d + A^c$ (see (3.8)) we find

$$\begin{aligned} A^s(u,v) &= A^d(u,v) + A^{c,s}(u,v) \\ &= \int_\Omega v \int_\Omega (u - u')\, \gamma^s d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega_I} \gamma^s d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega \cup \Omega_I} \gamma^a d\mathbf{y} d\mathbf{x} \end{aligned}$$

and

$$A^a(u,v) = A^{c,a}(u,v) = \int_\Omega v \int_\Omega u' \gamma^a d\mathbf{y} d\mathbf{x}.$$

By invoking (3.9) we find the following representation of the nonlocal energy seminorm

$$|||u|||^2 = A^s(u,u) = \int_\Omega u^2 \int_{\Omega\cup\Omega_I} \gamma^a d\mathbf{y} d\mathbf{x} + \frac{1}{2}\int_\Omega \int_\Omega (u-u')^2 \gamma^s d\mathbf{y} d\mathbf{x} + \int_\Omega u^2 \int_{\Omega_I} \gamma^s d\mathbf{y} d\mathbf{x}.$$

We see later, that the symmetric part γ^s, i.e., the part which accounts for diffusion, has to dominate the antisymmetric part γ^a, i.e., the convection component, to achieve that $|||\cdot|||$ is a well-defined norm.

Remark 3.1.2. *We have applied two different splitting approaches to the nonlocal bilinear form A. Namely, for any $u, v \in V_c(\Omega \cup \Omega_I)$ we have*

$$A(u,v) = A^d(u,v) + A^c(u,v) = \underbrace{A^d(u,v) + A^{c,s}(u,v)}_{A^s(u,v)} + \underbrace{A^{c,a}(u,v)}_{=A^a(u,v)}.$$

We note that the symmetric kernel γ^s contributes with A^d solely to the symmetric part of A, whereas γ^a contributes to both, A^s and A^a, since A^c falls into $A^c = A^{c,s} + A^{c,a}$. The decomposition resulting from splitting the kernel into its symmetric and antisymmetric part leading to $A = A^d + A^c$ with associated operators

$$\mathcal{L}^d_\Omega u(\mathbf{x}) = \int_\Omega (u-u')\gamma^s d\mathbf{y} + u\int_{\Omega_I} \gamma^s d\mathbf{y}$$

and

$$\mathcal{L}^c_\Omega u(\mathbf{x}) = \int_\Omega (u+u')\gamma^a d\mathbf{y} + u\int_{\Omega_I} \gamma^a d\mathbf{y}$$

is interesting from the physical perspective. It allows us to interpret the problem as a convection-diffusion model; see also (2.23).
For the analysis however, it is interesting to consider the splitting on the operator level, i.e., the splitting of the bilinear form into its symmetric and antisymmetric parts leading to $A = A^s + A^a$ with associated operators

$$\mathcal{L}^s_\Omega u(\mathbf{x}) = \mathcal{L}^d_\Omega u(\mathbf{x}) + u(\mathbf{x})\int_{\Omega_I} \gamma^a(\mathbf{x},\mathbf{y}) d\mathbf{y} \quad \textit{and} \quad \mathcal{L}^a_\Omega u(\mathbf{x}) = \int_\Omega u(\mathbf{y})\gamma^a(\mathbf{x},\mathbf{y}) d\mathbf{y}.$$

However, we have seen that splitting the kernel does not lead to the operator splitting, i.e., the two approaches do not match. This is due the structure of the nonlocal convection-diffusion operator, which becomes clearer for integrable kernels. In that case the operator inherits the structure $\mathcal{L}_\gamma = h_\gamma id + \mathcal{K}_\gamma$. We note that the first summand $h_\gamma id$ is symmetric for any type of kernel, (anti-)symmetric or not, explaining the difference $(\mathcal{L}^s_\Omega - \mathcal{L}^d_\Omega)u(\mathbf{x}) = u(\mathbf{x})\int_{\Omega_I}\gamma^a(\mathbf{x},\mathbf{y})d\mathbf{y}$. Finally we note that if $\gamma = \gamma^s$, then $A(u,v) = A^d(u,v) = A^s(u,v)$.

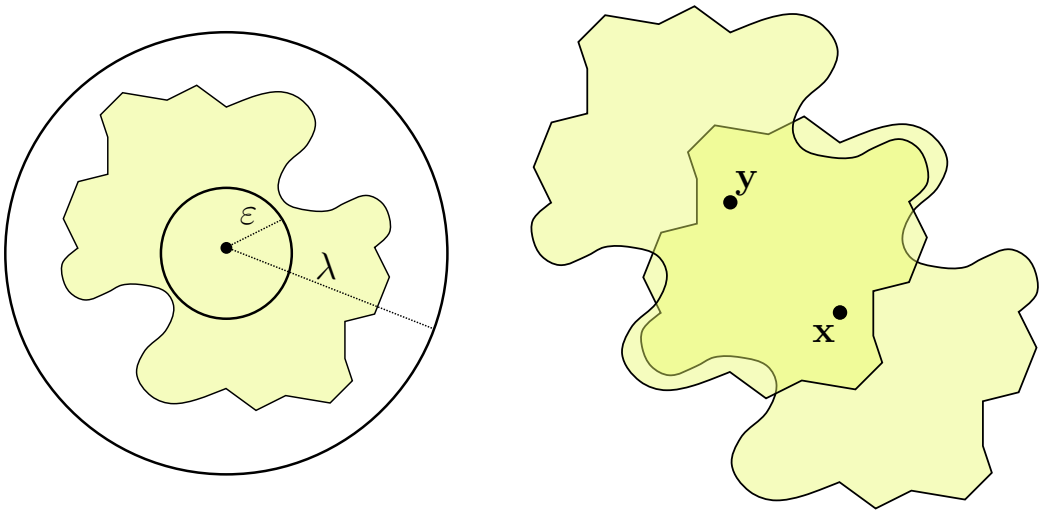


Figure 3.2: Illustration of the requirements (S1) and (S2) on the family of interaction sets. The yellow colored domain indicates an interaction set, which satisfies the sandwich property (S1) (left) and also the symmetry property (S2) (right).

3.2 Kernel conditions and interaction sets

Let $\{S(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ with $S(\mathbf{x}) \subset \mathbb{R}^d$ denote a family of sets that satisfy the following two properties.

(S1) *Sandwich property:* There exist two radii $\varepsilon, \lambda > 0$ such that

$$B_{\varepsilon,2}(\mathbf{x}) \subset S(\mathbf{x}) \subset B_{\lambda,2}(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^d.$$

(S2) *Symmetry:* For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, it holds that

$$\mathbf{y} \in S(\mathbf{x}) \quad \text{if and only if} \quad \mathbf{x} \in S(\mathbf{y}),$$

which implies that the indicator function $(\mathbf{x}, \mathbf{y}) \mapsto \chi_{S(\mathbf{x})}(\mathbf{y})$ is symmetric in $(\mathbf{x}, \mathbf{y})$.

Figure 3.2 depicts a nonstandard interaction set which satisfies these requirements. Now we impose two general conditions on the kernel γ.

(K1) First, we assume that there exists a positive constant $\gamma_0 > 0$ such that the symmetric part of the kernel satisfies

$$\gamma^s(\mathbf{x}, \mathbf{y}) \geq \gamma_0 > 0 \qquad \text{for all } \mathbf{x} \in \Omega \cup \Omega_I,\ \mathbf{y} \in B_{\varepsilon,2}(\mathbf{x}),$$

with $\varepsilon > 0$ from (S1).

(K2) Second, we assume that there exists a *positive* function $\phi\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ which we refer to as the *kernel function*, such that we can express the kernel as the product

$$\gamma(\mathbf{x},\mathbf{y}) = \phi(\mathbf{x},\mathbf{y})\chi_{S(\mathbf{x})}(\mathbf{y}),$$

where

$$\chi_{S(\mathbf{x})}(\mathbf{y}) := \begin{cases} 1 & : \mathbf{y} \in S(\mathbf{x}) \\ 0 & : else \end{cases}$$

denotes the indicator function.

Due to the symmetry condition (S2) we arrive at the representations

$$\gamma^s(\mathbf{x},\mathbf{y}) = \phi^s(\mathbf{x},\mathbf{y})\chi_{S(\mathbf{x})}(\mathbf{y}) \quad \text{and} \quad \gamma^a(\mathbf{x},\mathbf{y}) = \phi^a(\mathbf{x},\mathbf{y})\chi_{S(\mathbf{x})}(\mathbf{y}).$$

Requirement (K1) implies that A^d defines an inner product on $V_c(\Omega \cup \Omega_I) \times V_c(\Omega \cup \Omega_I)$. In the remaining sections we discuss kernels for which the constrained energy spaces are Banach.

Truncation (K2) of the kernel serves to restrict nonlocal interactions to occur only within the set $S(\mathbf{x})$ so that we refer to $S(\mathbf{x})$ as the *interaction set* for the point $\mathbf{x}$. An equivalent definition of the interaction domain Ω_I is given by

$$\Omega_I = \{\mathbf{y} \in \Omega^c \,:\, \mathbf{y} \in S(\mathbf{x}) \text{ for some } \mathbf{x} \in \Omega\}. \tag{3.11}$$

In Figure 3.3 the interaction domain Ω_I is depicted for two different choices of interaction sets. The boundedness of Ω and assumption (S1) on the interaction sets imply that $\Omega \cup \Omega_I$ is bounded. Examples of interaction sets are given by

$$S(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^d \,:\, \eta(\mathbf{x},\mathbf{y}) < 0\}$$

for a symmetric two-point function $\eta(\mathbf{x},\mathbf{y})$ for which there exist radii $\varepsilon_\eta, \lambda_\eta > 0$ with

$$\varepsilon_\eta \le \|\mathbf{x}-\mathbf{y}\|_2 - \eta(\mathbf{x},\mathbf{y}) \le \lambda_\eta \quad \text{for all } \mathbf{y} \in S(\mathbf{x}),$$

such that (S1) is fulfilled. If in addition $\eta(\mathbf{x},\mathbf{y})$ is translation invariant, i.e., $\eta(\mathbf{x},\mathbf{y}) = \eta(\mathbf{x}+\mathbf{b},\mathbf{y}+\mathbf{b})$ for all $\mathbf{b} \in \mathbb{R}^d$, then

$$S(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^d \,:\, \eta(\mathbf{y}-\mathbf{x},\mathbf{0}) < 0\} = \{(\mathbf{x}+\mathbf{z}) \in \mathbb{R}^d \,:\, \eta(\mathbf{z},\mathbf{0}) < 0\} = \mathbf{x} + S(\mathbf{0}),$$

where $S(\mathbf{0}) = -S(\mathbf{0})$. Note that $S(\mathbf{x}) := \mathbf{x} + S$, with $S \subset \mathbb{R}^d$ being any bounded set satisfying $B_{\varepsilon,2}(\mathbf{0}) \subset S = -S$, fulfills (S1) and (S2).
The particular interaction set $S(\mathbf{x})$ that is in ubiquitous use in nonlocal modeling corresponds to the symmetric and translation invariant two-point function $\eta(\mathbf{x},\mathbf{y}) := \|\mathbf{x}-\mathbf{y}\|_2 - \delta$ for $\delta > 0$ independent of $\mathbf{x}$, in which case $S(\mathbf{x})$ is simply the Euclidean ball $B_{\delta,2}(\mathbf{x})$ of radius δ centered at $\mathbf{x}$. However, within our more general framework, we consider general interaction sets $S(\mathbf{x})$ that satisfy (S1) and (S2). For instance, other special cases of interest are defined by

$$\eta(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_\bullet - \delta$$

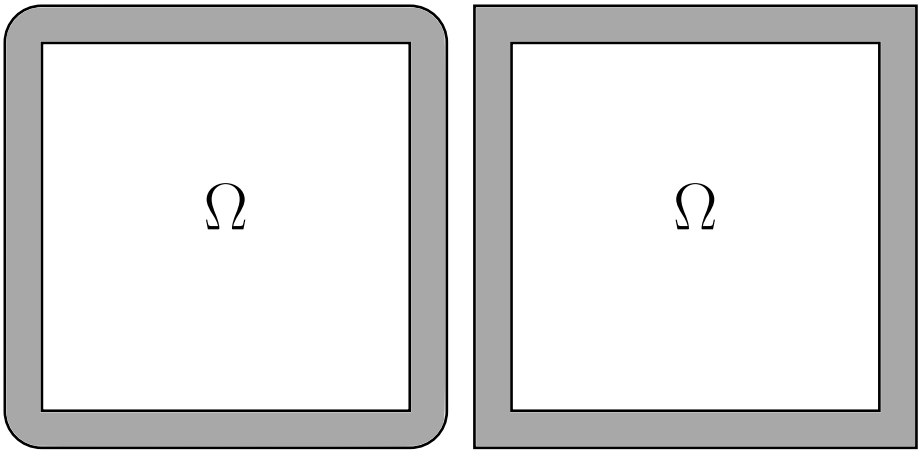


Figure 3.3: A rectangular domain Ω enclosed by its gray colored interaction domain Ω_I. On the left-hand side the standard case of Euclidean balls is depicted, whereas on the right-hand side balls induced by the supremum norm are used as interaction sets.

for an arbitrary norm $\|\cdot\|_\bullet$ in $\mathbb{R}^d$. We note that the sandwich property (S1) follows from the norm equivalences in $\mathbb{R}^d$. This leads to interaction sets $S(\mathbf{x})$ that are balls

$$B_{\delta,\bullet}(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^d : \|\mathbf{x}-\mathbf{y}\|_\bullet < \delta\} = \mathbf{x} + \{\mathbf{z} \in \mathbb{R}^d : \|\mathbf{z}\|_\bullet < \delta\}.$$

In these cases we refer to δ as the *interaction horizon* or simply as the *horizon*. In addition to the Euclidean ball $B_{\delta,2}(\mathbf{x})$, we can consider, e.g., the balls $B_{\delta,1}(\mathbf{x})$ and $B_{\delta,\infty}(\mathbf{x})$ induced by the Manhattan and supremum norm, respectively.

3.3 Well-posedness of the weak formulation

In this section, we introduce two concrete classes of kernels $\gamma\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ of the form (K2) both of which are well studied for Euclidean interaction balls and show the well-posedness of the weak formulation (3.7). We note that the "smoothing" property of the inverse nonlocal convection-diffusion operator depends on the regularity imposed on the kernel γ, or more precisely, on the kernel function ϕ.

We invoke the Lax-Milgram theorem (see, e.g., [44, Lemma 2.2]) to state the well-posedness of the weak formulation (3.7) and therefore need to show that A is a coercive and continuous bilinear form on some carefully chosen Hilbert space. For showing the coercivity we can exploit $A^a(u,u) = 0$, so that $A(u,u) = A^s(u,u)$.

3.3.1 Integrable kernels

In this section we consider a class of integrable kernels, not necessarily symmetric, for which the weak convection-diffusion problem (3.7) is well-posed. For this purpose we

show that the nonlocal bilinear form A is coercive and continuous on the constrained space

$$L_c^2(\Omega \cup \Omega_I) := \left\{ u \in L^2(\Omega \cup \Omega_I) \ : \ u = 0 \text{ on } \Omega_I \right\}.$$

Due to the composition $A = A^d + A^c$ given in (3.8) it stands to reason to impose separate conditions on the symmetric and antisymmetric parts of the kernel γ. Also, in doing so, we can rely on established results for the special case of a symmetric kernel $\gamma = \gamma^s$. The following requirements determining this class of integrable kernels are understood to be in addition to those included in (K1) and (K2).
For the diffusion component we assume that there exist two positive constants $\underline{\gamma}^s$, $\overline{\gamma}^s > 0$ such that

$$\begin{cases} \underline{\gamma}^s = \inf\limits_{\mathbf{x}\in\Omega} \displaystyle\int_{\Omega\cup\Omega_I} \gamma^s(\mathbf{x},\mathbf{y}) d\mathbf{y}, \\ \qquad \sup\limits_{\mathbf{x}\in\Omega} \displaystyle\int_{\Omega\cup\Omega_I} (\gamma^s)^2(\mathbf{x},\mathbf{y}) d\mathbf{y} = (\overline{\gamma}^s)^2. \end{cases} \tag{3.12}$$

The second requirement and the boundedness of $\Omega \cup \Omega_I$ imply that γ^s is a Hilbert-Schmidt kernel, since $\gamma^s \in L^2(\Omega \cup \Omega_I \times \Omega \cup \Omega_I)$. Also, by applying the Cauchy-Schwarz inequality, we find the estimate

$$\sup_{\mathbf{x}\in\Omega} \int_{\Omega\cup\Omega_I} |\gamma^s(\mathbf{x},\mathbf{y})| \, d\mathbf{y} \le \overline{\gamma}^s \sqrt{|\Omega \cup \Omega_I|}, \tag{3.13}$$

where $|\Omega \cup \Omega_I|$ denotes the d-dimensional volume of $\Omega \cup \Omega_I$. For the convection component we assume that there exist two constants $\underline{\gamma}^a \in \mathbb{R}$ and $\overline{\gamma}^a > 0$ such that

$$\begin{cases} \underline{\gamma}^a = \inf\limits_{\mathbf{x}\in\Omega} \displaystyle\int_{\Omega\cup\Omega_I} \gamma^a(\mathbf{x},\mathbf{y}) d\mathbf{y} \quad \text{and} \quad \underline{\gamma}^s + \underline{\gamma}^a > 0, \\ \qquad \sup\limits_{\mathbf{x}\in\Omega} \displaystyle\int_{\Omega\cup\Omega_I} |\gamma^a(\mathbf{x},\mathbf{y})| \, d\mathbf{y} = \overline{\gamma}^a. \end{cases} \tag{3.14}$$

The family of interaction sets $\{S(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ can be any family satisfying (S1) and (S2). We find that the lower bounds in (3.12) and (3.14) guarantee the coercivity of the nonlocal bilinear form whereas the upper bounds imply the continuity. Furthermore, as we see below, the requirement $\underline{\gamma}^s + \underline{\gamma}^a > 0$ in (3.14) assures that the diffusive part "dominates" the convection-diffusion equation and the coercivity is not deteriorated by the convection part. Examples for symmetric kernels are provided by the kernels used in [6, 7, 10] and trivially include the constant kernel. Considerable nonsymmetric kernels are introduced in the following example.

Example 3.3.1. *An example class of nonsymmetric kernels fulfilling the requirements* (3.12) *and* (3.14) *can be constructed as follows. Let*

$$\gamma \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}, \quad \gamma(\mathbf{x},\mathbf{y}) := \big(\gamma^s(\mathbf{x},\mathbf{y}) + \gamma^a(\mathbf{x}-\mathbf{y})\big)\chi_{S(\mathbf{x})}(\mathbf{y})$$

for any family of interaction sets $\{S(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ *satisfying* (S1) *and* (S2)*, for a symmetric* γ^s *satisfying* (K1) *and* (3.12) *and for an antisymmetric* γ^a *which is translation invariant, i.e.,* $\gamma^a(-\mathbf{z}) = -\gamma^a(\mathbf{z})$ *for* $\mathbf{z} = \mathbf{x} - \mathbf{y}$*. The latter implies that* $\underline{\gamma}^a = 0$ *in* (3.14)*. An example is given by* $\gamma^a(\mathbf{z}) := z_i$*, for* $1 \le i \le d$*.*

Recall that because of $A(u,u) = A^s(u,u)$ it suffices to show the coercivity of the symmetrized bilinear form

$$A^s \colon L_c^2(\Omega \cup \Omega_I) \times L_c^2(\Omega \cup \Omega_I) \to \mathbb{R}.$$

Furthermore, since we have separate conditions on each, the diffusion and convection part, respectively, we treat the symmetric parts of the corresponding bilinear forms, i.e., the summands of $A^s = A^d + A^{c,s}$, separately in the following. For the diffusion part A^d we find the following Lemma which is an immediate consequence of Lemma 4.7 from [36].

Lemma 3.3.2. *For any $u, v \in L_c^2(\Omega \cup \Omega_I)$ we find*

$$A^d(u,u) \geq \underline{\gamma}^s \|u\|^2_{L^2(\Omega)} \tag{3.15}$$

with $\underline{\gamma}^s > 0$ from (3.12). In particular, A^d is coercive on $L_c^2(\Omega \cup \Omega_I)$.

Proof. Let $u, v \in L_c^2(\Omega \cup \Omega_I)$. Recall that $A^d(u,v) = (-\mathcal{L}^d_\Omega u, v)$, with $\mathcal{L}^d$ from (2.21). Also, since $u, v = 0$ on Ω_I we can rely on the representation of A^d given in (3.10). Since the family of interaction sets is assumed to be uniformly bounded (S1) there is a radius $\lambda > 0$ such that $S(\mathbf{x}) \subset B_{\lambda,2}(\mathbf{x})$ for all $\mathbf{x} \in \Omega \cup \Omega_I$. Then let us define the kernel

$$\tilde{\gamma}(\mathbf{x},\mathbf{y}) := \gamma^s(\mathbf{x},\mathbf{y}) \chi_{B_{\lambda,2}(\mathbf{x})}(\mathbf{y}) = \phi^s(\mathbf{x},\mathbf{y}) \chi_{S(\mathbf{x})}(\mathbf{y}) \chi_{B_{\lambda,2}(\mathbf{x})}(\mathbf{y})$$

and the operator

$$\tilde{\mathcal{L}} u(\mathbf{x}) := \int_{\Omega \cup \Omega_I} (u(\mathbf{y}) - u(\mathbf{x})) \tilde{\gamma}(\mathbf{x},\mathbf{y}) d\mathbf{y}$$

as well as the associated bilinear form

$$\tilde{A}(u,v) := (-\tilde{\mathcal{L}} u, v)_{L^2(\Omega)}.$$

Clearly, we have $\tilde{\gamma} = \gamma^s$ on $\Omega \cup \Omega_I \times \Omega \cup \Omega_I$ and therefore also $\tilde{\mathcal{L}} u = \mathcal{L}^d u$ on Ω and $\tilde{A} = A^d$ on $L_c^2(\Omega \cup \Omega_I) \times L_c^2(\Omega \cup \Omega_I)$. Then, due to the general assumptions on the interaction sets (i.e., (S1) and (S2)) and on the kernel (i.e, (K1) and (K2)) as well as the requirements on γ^s (i.e., (3.12)), the auxiliary kernel $\tilde{\gamma}$ and its associated operators $\tilde{\mathcal{L}}$ and $\tilde{A}$ fulfill the requirements of Lemma 4.7 in [36], which states the desired result (3.15).

□

Inequality (3.15) is sometimes referred to as *nonlocal Poincaré inequality* in the literature [36]. We finally obtain that the coercivity of A^d is not deteriorated by the symmetric part $A^{c,s}$ of A^c, such that A determines a coercive bilinear form on $L_c^2(\Omega \cup \Omega_I)$.

Lemma 3.3.3 (***Coercivity***)**.** *For any $u \in L_c^2(\Omega \cup \Omega_I)$ we find*

$$A(u,u) \geq \underline{\gamma} \|u\|^2_{L^2(\Omega)}$$

where $\underline{\gamma} := \underline{\gamma}^s + \underline{\gamma}^a > 0$. In particular, A is coercive on $L_c^2(\Omega \cup \Omega_I)$.

Proof. By requirement (3.14) on γ^a we obtain the estimate

$$A^{c,s}(u,v) = \int_\Omega u^2 \int_{\Omega\cup\Omega_I} \gamma^a d\mathbf{y}d\mathbf{x} \geq \underline{\gamma}^a \|u\|^2_{L^2(\Omega)}.$$

Combining this estimate with Lemma 3.3.2 we find that

$$\begin{aligned} A(u,u) = A^s(u,u) &= (A^d + A^{c,s})(u,u) \\ &\geq \underline{\gamma}_s \|u\|^2_{L^2(\Omega)} + \underline{\gamma}_a \|u\|^2_{L^2(\Omega)} = \underline{\gamma}\|u\|^2_{L^2(\Omega)}, \end{aligned}$$

where $\underline{\gamma} := \underline{\gamma}^s + \underline{\gamma}^a > 0$ by assumption (3.14). □

The latter result shows that $|||\cdot||| = \sqrt{A(\cdot,\cdot)}$ defines a norm and $A^s(\cdot,\cdot)$ an inner product on the constrained space $L^2_c(\Omega\cup\Omega_I)$. The next lemma provides the continuity of the nonlocal bilinear form under slightly more general conditions on the kernel than required in this section. In Chapter 6 we return to this more general result.

Lemma 3.3.4. *Let $\varrho\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfy*

$$\overline{\varrho} := \max\left\{\sup_{\mathbf{x}\in\Omega}\int_{\Omega\cup\Omega_I} |\varrho| d\mathbf{y},\ \sup_{\mathbf{x}\in\Omega}\int_\Omega |\varrho'| d\mathbf{y}\right\} < +\infty, \tag{3.16}$$

and consider the bilinear form

$$\begin{aligned} B : L^2_c(\Omega\cup\Omega_I)\times L^2_c(\Omega\cup\Omega_I) &\to \mathbb{R}, \\ B(u,v) &:= \int_\Omega v \int_\Omega (u\varrho - u'\varrho') d\mathbf{y}d\mathbf{x} + \int_\Omega vu \int_{\Omega_I} \varrho d\mathbf{y}d\mathbf{x}, \end{aligned}$$

where $\Omega \subset \mathbb{R}^d$. Then for any $u, v \in L^2_c(\Omega\cup\Omega_I)$ we obtain the bound

$$|B(u,v)| \leq 2\overline{\varrho}\|u\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)},$$

implying that B is continuous on $L^2_c(\Omega\cup\Omega_I)$.

Proof. We first note that due to the integrability of the kernel (3.16) we can write

$$B(u,v) = \int_\Omega vu \int_{\Omega\cup\Omega_I} \varrho d\mathbf{y}d\mathbf{x} - \int_\Omega v \int_\Omega u'\varrho' d\mathbf{y}d\mathbf{x}.$$

We find the bound on B by applying the Cauchy-Schwarz inequality. Let us first define

$$\tilde{\varrho} := \sup_{\mathbf{x}\in\Omega}\int_{\Omega\cup\Omega_I} |\varrho| d\mathbf{y}, \qquad \tilde{\varrho}' := \sup_{\mathbf{x}\in\Omega}\int_\Omega |\varrho'| d\mathbf{y}.$$

Then for any $u, v \in L^2_c(\Omega\cup\Omega_I)$ we have

$$|B(u,v)| \leq \int_\Omega |vu| \left(\int_{\Omega\cup\Omega_I} |\varrho| d\mathbf{y}\right) d\mathbf{x} + \int_\Omega\int_\Omega |vu'||\varrho'| d\mathbf{y}d\mathbf{x}$$

$$
\begin{aligned}
&\leq \tilde{\varrho}\|v\|_{L^2(\Omega)}\|u\|_{L^2(\Omega)} + \sqrt{\int_\Omega v^2 \int_\Omega |\varrho'| d\mathbf{y} d\mathbf{x}} \sqrt{\int_\Omega \int_\Omega (u')^2 |\varrho'| d\mathbf{y} d\mathbf{x}} \\
&\leq \tilde{\varrho}\|v\|_{L^2(\Omega)}\|u\|_{L^2(\Omega)} + \sqrt{\tilde{\varrho}'}\|v\|_{L^2(\Omega)} \sqrt{\tilde{\varrho}}\|u\|_{L^2(\Omega)} \\
&\leq 2\overline{\varrho}\|u\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}.
\end{aligned}
$$

The second inequality follows from Cauchy-Schwarz inequality, where for the second summand applied to the integrand $|v|\sqrt{|\varrho'|} \cdot |u'|\sqrt{|\varrho'|}$. The third inequality follows from Fubini's theorem applied to the second factor in the second summand. □

We can immediately deduce the continuity of the nonlocal bilinear form in our setting.

Corollary 3.3.5 (***Continuity***)**.** *For any $u, v \in L_c^2(\Omega \cup \Omega_I)$ we obtain*

$$
|A(u,v)| \leq 2\overline{\gamma}\|u\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)},
$$

where $\overline{\gamma} := \overline{\gamma}^s \sqrt{|\Omega \cup \Omega_I|} + \overline{\gamma}^a$.

Proof. Due to the upper bounds given in (3.13) and (3.14) on the symmetric and anti-symmetric part of the kernel, we find the estimate

$$
\begin{aligned}
\sup_{\mathbf{x}\in\Omega} \int_{\Omega\cup\Omega_I} |\gamma| d\mathbf{y} &\leq \sup_{\mathbf{x}\in\Omega} \int_{\Omega\cup\Omega_I} |\gamma^s| d\mathbf{y} + \sup_{\mathbf{x}\in\Omega} \int_{\Omega\cup\Omega_I} |\gamma^a| d\mathbf{y} \\
&\leq \overline{\gamma}^s \sqrt{|\Omega \cup \Omega_I|} + \overline{\gamma}^a.
\end{aligned}
$$

Since $|\gamma^s|$ and $|\gamma^a|$ are symmetric, the same bound $\overline{\gamma} := \overline{\gamma}^s \sqrt{|\Omega \cup \Omega_I|} + \overline{\gamma}^a$ holds for $\sup_{\mathbf{x}\in\Omega} \int_{\Omega\cup\Omega_I} |\gamma'| d\mathbf{y}$. Thus, the result follows by Lemma 3.3.4. □

The well-posedness of the weak formulation (3.7) in the space $L_c^2(\Omega \cup \Omega_I)$ then is a consequence of the Lax-Milgram theorem and the results established above. We obtain the a priori estimate

$$
\|u\|_{L^2(\Omega)} \leq C\|f - \mathcal{L}_{\Omega_I} g\|_{L^2(\Omega)} \leq \tilde{C}(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega_I)}), \quad \text{for some } \tilde{C} > 0,
$$

so that there is no gain of regularity for the solution of the nonlocal volume-constrained problem. Furthermore, the latter results also lead to the following equivalence of spaces.

Corollary 3.3.6 (***Equivalence of spaces***)**.** *Combining Lemma 3.3.2 and Corollary 3.3.5 we find the norm equivalence*

$$
\sqrt{\underline{\gamma}}\ \|u\|_{L^2(\Omega)} \leq |||u||| \leq \sqrt{2\overline{\gamma}}\ \|u\|_{L^2(\Omega)}
$$

for all $u \in L_c^2(\Omega \cup \Omega_I)$, implying that the normed vector spaces $(V_c(\Omega \cup \Omega_I), |||\cdot|||)$ and $\left(L_c^2(\Omega \cup \Omega_I), \|\cdot\|_{L^2(\Omega)}\right)$ are equivalent. Thus the nonlocal constrained energy space $(V_c(\Omega \cup \Omega_I), |||\cdot|||)$ is a Banach space. Also, if $\gamma = \gamma^s$, such that $A = A^s = A^d$ on $V_c(\Omega \cup \Omega_I) \times V_c(\Omega \cup \Omega_I)$ we find that the well-posedness of (3.7) *also holds on $V_c(\Omega \cup \Omega_I)$ by the Riesz representation theorem. In addition to that, the equivalence between the constrained spaces also imply the equivalence between the unconstrained spaces*

$$
\left(V(\Omega \cup \Omega_I)\,,\ |||\cdot||| + \|\cdot\|_{L^2(\Omega\cup\Omega_I)}\right) \sim \left(L^2(\Omega \cup \Omega_I)\,,\ \|\cdot\|_{L^2(\Omega\cup\Omega_I)}\right).
$$

3.3.2 Singular symmetric kernels

In this subsection we consider a class of singular symmetric kernels for which the nonlocal diffusion problem is relatable to space-fractional diffusion equations. We treat such kernels in Chapter 5 and Section 6.2. Since we restrict ourselves to symmetric kernels $\gamma = \gamma^s$, thereby provoking the absence of convective effects, we consider a pure diffusion model. In this case the nonlocal operator takes the form

$$\mathcal{L}u(\mathbf{x}) = \mathcal{L}^d u(\mathbf{x}) = \int_{\mathbb{R}^d} (u(\mathbf{y}) - u(\mathbf{x}))\, \gamma(\mathbf{x}, \mathbf{y}) d\mathbf{y}$$

and the nonlocal bilinear form reduces to the symmetric diffusion part

$$\begin{aligned} A(u, v) = A^d(u, v) &= \frac{1}{2} \int_\Omega \int_\Omega (v - v')\, (u - u')\, \gamma^s d\mathbf{y} d\mathbf{x} + \int_\Omega vu \int_{\Omega_I} \gamma^s d\mathbf{y} d\mathbf{x} \\ &= \frac{1}{2} \int_{\Omega \cup \Omega_I} \int_{\Omega \cup \Omega_I} (u - u')(v - v') \gamma^s d\mathbf{y} d\mathbf{x}, \end{aligned}$$

for all $u, v \in V_c(\Omega \cup \Omega_I)$. Let $s \in (0, 1)$, then in addition to the requirements included in (K1) and (K2) we assume the existence of two constants $0 < \underline{\gamma} \leq \overline{\gamma} < \infty$ such that

$$\underline{\gamma} \leq \gamma(\mathbf{x}, \mathbf{y}) \|\mathbf{y} - \mathbf{x}\|_2^{d+2s} \leq \overline{\gamma} \quad \text{for all } \mathbf{x} \in \Omega,\ \mathbf{y} \in S(\mathbf{x}), \tag{3.17}$$

where $\{S(\mathbf{x})\}_{\mathbf{x} \in \mathbb{R}^d}$ could be any family of interaction sets satisfying (S1) and (S2). Requirement (3.17) serves to limit the strength of the singularity of $\gamma(\mathbf{x}, \mathbf{y})$ at $\mathbf{y} = \mathbf{x}$ as well as its decay rate for large values of $\|\mathbf{y} - \mathbf{x}\|_2$. Note that we allow for general interaction sets $S(\mathbf{x})$ but use the norm $\|\cdot\|_2$ for the two purposes just mentioned. Due to the equivalence of norms in $\mathbb{R}^d$, without loss of generality, we can replace $\|\cdot\|_2$ in (3.17) by other norms such as the $\|\cdot\|_1$- or $\|\cdot\|_\infty$-norm. An example of these kernels is given by

$$\gamma(\mathbf{x}, \mathbf{y}) := \frac{\sigma(\mathbf{x}, \mathbf{y})}{\|\mathbf{y} - \mathbf{x}\|_2^{d+2s}} \chi_{S(\mathbf{x})}(\mathbf{y}) \tag{3.18}$$

for a symmetric function $\sigma \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ that is bounded from below and above by some positive constants. Note that this kernel could also be chosen to be radial with respect to any discrete norm. When σ is constant, we sometimes refer to this kernel as the *fractional kernel* because, if $\delta = \infty$ and σ attains an appropriate scaling constant, (3.18) becomes the kernel for the fractional Laplace operator [54].

We recall the definition of the fractional Sobolev space

$$H^s(\Omega \cup \Omega_I) := \left\{ u \in L^2(\Omega \cup \Omega_I) :\ \|u\|_{H^s(\Omega \cup \Omega_I)} := \|u\|_{L^2(\Omega \cup \Omega_I)} + |u|_{H^s(\Omega \cup \Omega_I)} < \infty \right\}$$

for $s \in (0, 1)$, where we define the seminorm

$$|u|^2_{H^s(\Omega \cup \Omega_I)} := \int_{\Omega \cup \Omega_I} \int_{\Omega \cup \Omega_I} \frac{(u(\mathbf{x}) - u(\mathbf{y}))^2}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}} d\mathbf{y} d\mathbf{x}.$$

The volume constrained fractional Sobolev space is then defined as

$$H_c^s(\Omega \cup \Omega_I) := \{u \in H^s(\Omega \cup \Omega_I) \ : \ u = 0 \text{ on } \Omega_I\}$$

for which $|u|_{H^s(\Omega\cup\Omega_I)}$ is a norm [59].
We show that $A = A^s$ is a coercive bilinear form on the constrained fractional Sobolev space $\left(H_c^s(\Omega \cup \Omega_I), |\cdot|_{H^s(\Omega\cup\Omega_I)}\right)$. For this purpose we first need to provide a Poincaré type inequality for the energy norm associated to this class of kernels. With the same considerations as in the proof of Lemma 3.3.2, the proof of the following Lemma is a consequence of Lemma 4.3 established in [36] and is therefore omitted here.

Lemma 3.3.7 (***Poincaré inequality***)**.** *For any $u \in H_c^s(\Omega \cup \Omega_I)$ we find*

$$\|u\|_{L^2(\Omega)} \leq C_P |||u||| \tag{3.19}$$

for some constant $C_P > 0$.

For completeness we provide the following observation which we need in Section 6.2.

Remark 3.3.8 (***Uniform Poincaré inequality***)**.** *For the special case of norm induced interaction sets $S(\mathbf{x}) = B_{\delta,\bullet}(\mathbf{x})$ the Poincaré constant C_P in (3.19) depends on the interaction horizon δ. However, by carefully analyzing the proof of Lemma 4.3 from [36], which is established for the standard case $S(\mathbf{x}) = B_{\delta,2}(\mathbf{x})$, we find that we can extend the therein established Poincaré inequality to a uniform one. More precisely, in this paper the authors prove the statement by contradicting the existence of a sequence $(u_k)_k \in V_c^{\mathbb{N}}$ with $\|u_k\|^2_{L^2(\Omega\cup\Omega_I)} = 1$, for which $|||u_k|||^2 < \frac{1}{k}$. Their first step is to show that this sequence is bounded in $H^s(\Omega \cup \Omega_I)$. More precisely, they derive*

$$\|u_k\|_{H^s(\Omega\cup\Omega_I)} \leq 1 + 4|\Omega \cup \Omega_I|\delta^{-(d+2s)}. \tag{3.20}$$

Once this step is established, the desired result then follows from statements independent of δ. Since $\delta^{-(d+2s)} \leq 1$ for all $\delta \in [1, +\infty)$ we can use $1 + 4|\Omega \cup \Omega_I|$ as a uniform bound in (3.20). Consequently we obtain a uniform Poincaré constant such that the inequality holds for all $\delta \in [1, +\infty)$. Using the auxiliary kernel $\tilde{\gamma}$ from the proof of Lemma 3.3.2, we can establish a uniform Poincaré inequality for the case of general norm induced interaction sets $S(\mathbf{x}) = B_{\delta,\bullet}(\mathbf{x})$ which holds for all $\delta \in [C_\bullet, +\infty)$, where $C_\bullet > 0$ is a constant depending on the norm equivalence $\|\cdot\|_2 \sim \|\cdot\|_\bullet$.

We are now in the position to prove the coercivity of A on $\left(H_c^s(\Omega \cup \Omega_I), |\cdot|_{H^s(\Omega\cup\Omega_I)}\right)$.

Lemma 3.3.9 (***Coercivity***)**.** *For any $u \in H_c^s(\Omega \cup \Omega_I)$ we find*

$$A(u, u) \geq C_{coer} |u|^2_{H^s(\Omega\cup\Omega_I)}$$

for some constant $C_{coer} > 0$. In particular, A is coercive on $H_c^s(\Omega \cup \Omega_I)$.

Proof. Let $u \in H_c^s(\Omega \cup \Omega_I)$. Due to assumption (S1), for any $\mathbf{x} \in \Omega \cup \Omega_I$ we find

$$\sup_{\mathbf{y} \in (\Omega \cup \Omega_I) \setminus S(\mathbf{x})} \frac{1}{\|\mathbf{y}-\mathbf{x}\|_2^{d+2s}} \leq \sup_{\mathbf{y} \in (\Omega \cup \Omega_I) \setminus B_{\varepsilon,2}(\mathbf{x})} \frac{1}{\|\mathbf{y}-\mathbf{x}\|_2^{d+2s}} \leq \frac{1}{\varepsilon^{d+2s}}.$$

Consequently

$$\overline{\overline{\gamma}} := \sup_{\mathbf{x} \in \Omega \cup \Omega_I} \sup_{\mathbf{y} \in \Omega \cup \Omega_I \setminus S(\mathbf{x})} \frac{1}{\|\mathbf{y}-\mathbf{x}\|_2^{d+2s}} \leq \frac{1}{\varepsilon^{d+2s}} < +\infty.$$

Thus we find

$$\begin{aligned}
|u|_{H^s(\Omega\cup\Omega_I)}^2 &= \int_{\Omega\cup\Omega_I} \int_{(\Omega\cup\Omega_I)\cap S(\mathbf{x})} \frac{(u(\mathbf{x})-u(\mathbf{y}))^2}{\|\mathbf{y}-\mathbf{x}\|_2^{d+2s}} d\mathbf{y}d\mathbf{x} \\
&\quad + \int_{\Omega\cup\Omega_I} \int_{(\Omega\cup\Omega_I)\setminus S(\mathbf{x})} \frac{(u(\mathbf{x})-u(\mathbf{y}))^2}{\|\mathbf{y}-\mathbf{x}\|_2^{d+2s}} d\mathbf{y}d\mathbf{x} \\
&\leq \underline{\gamma}^{-1} A(u,u) + \overline{\overline{\gamma}} \int_{\Omega\cup\Omega_I} \int_{(\Omega\cup\Omega_I)\setminus S(\mathbf{x})} (u(\mathbf{x})-u(\mathbf{y}))^2 d\mathbf{y}d\mathbf{x} \\
&\leq \underline{\gamma}^{-1} A(u,u) + \overline{\overline{\gamma}} \int_{\Omega\cup\Omega_I} \int_{\Omega\cup\Omega_I} (u(\mathbf{x})-u(\mathbf{y}))^2 d\mathbf{y}d\mathbf{x} \\
&\leq \underline{\gamma}^{-1} A(u,u) + 2\overline{\overline{\gamma}} \int_{\Omega\cup\Omega_I} \int_{\Omega\cup\Omega_I} \left(u(\mathbf{x})^2 + u(\mathbf{y})^2\right) d\mathbf{y}d\mathbf{x} \\
&= \underline{\gamma}^{-1} A(u,u) + 4\overline{\overline{\gamma}} |\Omega \cup \Omega_I| \|u\|_{L^2(\Omega\cup\Omega_I)}^2.
\end{aligned}$$

The first inequality follows from the assumptions (3.17) on the kernel and (K2). The second inequality follows because the integrand is nonnegative and the third inequality follows from the estimate $(a-b)^2 \leq 2(a^2+b^2)$ for $a, b \in \mathbb{R}$. Now invoking the nonlocal Poincaré inequality (3.19) we obtain

$$\begin{aligned}
|u|_{H^s(\Omega\cup\Omega_I)}^2 &\leq \underline{\gamma}^{-1} A(u,u) + 4\overline{\overline{\gamma}} |\Omega \cup \Omega_I| C_P^2 A(u,u) \\
&\leq \max\{\underline{\gamma}_1^{-1}, 4\overline{\overline{\gamma}} C_P^2 |\Omega \cup \Omega_I|\} A(u,u).
\end{aligned}$$

Thus, $C_{coer} := \max\{\underline{\gamma}_1^{-1}, 4\overline{\overline{\gamma}} C_P^2 |\Omega \cup \Omega_I|\}$ provides the necessary constant. □

The next lemma establishes the continuity of the nonlocal bilinear form A on the constrained fractional Sobolev space $H_c^s(\Omega \cup \Omega_I)$.

Lemma 3.3.10 (***Continuity***)**.** *For any $u, v \in H_c^s(\Omega \cup \Omega_I)$ we find*

$$|A(u,v)| \leq \overline{\gamma} |u|_{H^s(\Omega\cup\Omega_I)} |v|_{H^s(\Omega\cup\Omega_I)} \tag{3.21}$$

with $\overline{\gamma} > 0$ from (3.17)*. In particular, A is continuous on $H_c^s(\Omega \cup \Omega_I)$.*

Proof. Due to (K1) in combination with (S1) the bilinear form $A = A^d$ defines a definite inner product on $H_c^s(\Omega \cup \Omega_I)$. Thus by Cauchy-Schwarz inequality we obtain

$$|A(u,v)|^2 \leq A(u,u) A(v,v).$$

Now let $u \in H_c^s(\Omega \cup \Omega_I)$, then it follows from assumption (3.17) that

$$\begin{aligned} A(u,u) &= \int_{\Omega\cup\Omega_I} \int_{(\Omega\cup\Omega_I)\cap S(\mathbf{x})} (u(\mathbf{x}) - u(\mathbf{y}))^2 \phi(\mathbf{x},\mathbf{y}) d\mathbf{y} d\mathbf{x} \\ &\leq \overline{\gamma} \int_{\Omega\cup\Omega_I} \int_{\Omega\cup\Omega_I} \frac{(u(\mathbf{x}) - u(\mathbf{y}))^2}{\|\mathbf{y}-\mathbf{x}\|_2^{d+2s}} d\mathbf{y} d\mathbf{x} = \overline{\gamma} |u|^2_{H^s(\Omega\cup\Omega_I)}. \end{aligned}$$

The stated inequality (3.21) is an immediate consequence. □

The well-posedness of the weak formulation (3.7) in the space $H_c^s(\Omega \cup \Omega_I)$ is a consequence of the Lax-Milgram theorem and the results established above. In case of homogeneous Dirichlet constraints $g = 0$ we further obtain the a priori estimate

$$\|u\|_{H^s(\Omega)} \leq C \|f\|_{H^{-s}(\Omega)}, \quad 0 < s < 1.$$

In contrast, for second-order elliptic partial differential equations, we have

$$\|u\|_{H^1(\Omega)} \leq C \|f\|_{H^{-1}(\Omega)}$$

so that the nonlocal volume-constrained problem with this type of singular kernels results in a lower gain in regularity. More precisely, for second-order elliptic partial differential equations, there is gain of regularity of 2 whereas for the nonlocal volume-constrained problem, there is gain of regularity of $2s$ with $s \in (0,1)$.

As for the integrable kernel class, the latter results also lead to the following equivalence of spaces.

Corollary 3.3.11 (***Equivalence of spaces***)**.** *Combining Lemmata 3.3.9 and 3.3.10 we find the norm equivalence*

$$\sqrt{C_{coer}}\ |u|_{H^s(\Omega\cup\Omega_I)} \leq |||u||| \leq \sqrt{\overline{\gamma}}\ |u|_{H^s(\Omega\cup\Omega_I)}$$

for all $u \in H_c^s(\Omega \cup \Omega_I)$, implying that the normed vector spaces $(V_c(\Omega \cup \Omega_I), |||\cdot|||)$ and $\left(H_c^s(\Omega \cup \Omega_I), |\cdot|_{H^s(\Omega\cup\Omega_I)}\right)$ are equivalent. Thus the nonlocal constrained energy space $(V_c(\Omega \cup \Omega_I), |||\cdot|||)$ is a Banach space and the well-posedness of (3.7) *also holds on $V_c(\Omega \cup \Omega_I)$ by the Riesz representation theorem. Moreover, the equivalence between the constrained spaces also imply the equivalence between the unconstrained spaces*

$$\left(V(\Omega\cup\Omega_I)\,,\, |||\cdot||| + \|\cdot\|_{L^2(\Omega\cup\Omega_I)}\right) \sim \left(H^s(\Omega\cup\Omega_I)\,,\, |\cdot|_{H^s(\Omega\cup\Omega_I)} + \|\cdot\|_{L^2(\Omega\cup\Omega_I)}\right).$$

3.4 Concluding remarks

We have shown the well-posedness of the weak formulation (3.7) by applying the Lax-Milgram theorem for which the coercivity of the symmetrized bilinear form A^s is crucial. The latter implies that the regarded problems exhibit properties comparable to that of elliptic partial differential operators. For the class of possibly nonsymmetric integrable

kernels treated in Subsection 3.3.1 we thereby only allow for a "mild" convection component which does not deteriorate the ellipticity property. However, for this class of nonsymmetric integrable kernels there may arise situations for which the problem is well-posed but the associated bilinear form is *not* coercive; see Chapter 8 for a respective counterexample and a related discussion.

Chapter 4

Discretization using the finite element method

Given the variational formulation (3.7) of the nonlocal volume-constrained problem (2.1), it is natural to consider its finite-dimensional approximation within that variational framework. Specifically, we focus on finite element approximations and discuss in detail the assembly of the *nonlocal* stiffness matrix which requires numerical quadrature for the double integral appearing in the nonlocal bilinear form (3.5). We expose the concomitant difficulties and challenges arising along the assembly process which contrasts strongly with the local setting and thereby reveal the price we have to pay for including nonlocality into our model.
In addition to that, we examine two different types of interaction sets in the context of finite element implementations, which both have not attained detailed discussion in literature yet. On the one hand, in Chapter 3 we have established a well-posedness theory which allows us to model nonlocal interactions to occur within polyhedral interaction sets, such as balls induced by the supremum norm. Here, we discuss the computational advantage of such interaction sets over Euclidean balls which are ubiquitous in the literature. On the other hand, in terms of efficiency we advertise geometric approximations to the exact interaction set. A further analysis of the latter is then exhibited in Section 6.3.
In this thesis the solving procedure of the resulting discrete system and related issues are not discussed at this general level. For state of the art finite element discretizations, the assembly process represents the restrictive factor when aiming to solve nonlocal equations on very fine meshes. Thus, the size of the discrete system is manageably small in our experiments. Exceptions for the latter statement arise in special settings; see, e.g., Chapter 5.

This chapter is organized as follows. In Section 4.1 we recall a finite-dimensional approximation of the weak formulation (3.7) which applies in general to variational problems. We then specify to finite element approximations in Section 4.2. The central notion is that of the nonlocal stiffness matrix whose challenging assembly is discussed in the subsequent Section 4.3. Finally we introduce the notion of *approximate* interaction

sets in Section 4.4.

4.1 Finite-dimensional approximation

Let $\{V^h\}_{h>0} \subset V(\Omega \cup \Omega_I)$ denote a sequence of finite-dimensional subspaces and assume that $\{V^h\}_{h>0}$ is dense in $V(\Omega \cup \Omega_I)$ as $h \to 0$, i.e., for any $u \in V(\Omega \cup \Omega_I)$, there exists a sequence $\{u^h \in V^h\}_{h>0}$ such that

$$|||u - u^h||| \to 0 \quad \text{as} \quad h \to 0.$$

The family of constrained finite-dimensional subspaces is then given by

$$V_c^h := \left\{ u^h \in V^h \colon u^h = 0 \ \text{ on } \ \Omega_I \right\} \subset V^h,$$

which is dense in $V_c(\Omega \cup \Omega_I)$ as $h \to 0$. Let $u \in V(\Omega \cup \Omega_I)$ denote the solution of the variational problem (3.7) then we seek the Ritz-Galerkin approximation $u^h \in V^h$ of u determined by posing the variational problem (3.7) on the finite-dimensional subspace, i.e.,

$$\begin{aligned} &\textit{given } f \in V_c'(\Omega) \textit{ and } g \in \widetilde{V}(\Omega_I), \textit{ find } u^h \in V^h \textit{ such that } u^h|_{\Omega_I} = g \textit{ and} \\ &\qquad A(u^h|_\Omega, v^h) = \ell(v^h) \quad \textit{for all } v^h \in V_c^h. \end{aligned} \tag{4.1}$$

For the two classes of kernels discussed in Subsection 3.3.1 and Subsection 3.3.2, respectively, problem (4.1) satisfies the *best approximation property*

$$|||u - u^h||| \leq \inf_{v^h \in V_c^h} |||u - v^h||| \to 0 \ \text{ as } \ h \to 0. \tag{4.2}$$

In fact, since u and u^h satisfy (3.7) and (4.1), the approximation u^h is the $|||\cdot|||$-orthogonal projection of u onto V^h (the so-called *Galerkin projection property*). Then result (4.2) follows from the assumed density of the spaces $V^h \subset V(\Omega \cup \Omega_I)$ as $h \to 0$. Also, due to the norm equivalences and Poincaré inequalities established for the two classes of kernels we find

- for the class of integrable kernels (see Subsection 3.3.1)

$$||u - u^h||_{L^2(\Omega \cup \Omega_I)} \to 0 \ \text{ as } \ h \to 0,$$

- and for the class of singular and symmetric kernels (see Subsection 3.3.2)

$$||u - u^h||_{H^s(\Omega \cup \Omega_I)} \to 0 \ \text{ as } \ h \to 0,$$

  and

$$||u - u^h||_{L^2(\Omega \cup \Omega_I)} \to 0 \ \text{ as } \ h \to 0.$$

Remark 4.1.1 (***Convergence rates***)**.** *Precise convergence rates with respect to the interpolation parameter $h > 0$ typically depend on the regularity of the (analytic) weak solution and the approximation quality of the implemented finite-dimensional subspaces (ansatz spaces). In this work such regularity and convergence issues are not investigated at an analytical level.*

4.2 Finite element approximation

In this thesis we consider finite element approximations in order to define appropriate finite-dimensional energy spaces V^h, where $h > 0$ is a parameter determining the grid size of the underlying finite element mesh. Finite element methods for nonlocal volume constrained problems such as (2.1) have been studied, e.g., in [96, 5, 2].

For simplicity we assume that Ω is a polyhedral domain and that the interaction sets $\{S(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ are simply connected. The first requirement implies that we can exactly triangulate Ω into finite (polyhedral) elements $\mathcal{T}_\Omega^h = \{\mathcal{E}_j\}_{j=1}^{J_\Omega}$ with

$$\overline{\Omega} = \bigcup_{j=1}^{J_\Omega} \overline{\mathcal{E}}_j.$$

We assume that the corners of the boundary $\partial\Omega$ are vertices of the finite element mesh. In the standard case of Euclidean interaction sets $S(\mathbf{x}) = B_{\delta,2}(\mathbf{x})$ the interaction domain Ω_I is not polyhedral. However we assume that $\mathcal{T}_{\Omega_I}^h = \{\mathcal{E}_j\}_{j=J_\Omega+1}^{J}$ triangulates a polyhedral approximation Ω_I^{approx} of Ω_I, such that

$$\overline{\Omega}_I^{approx} = \bigcup_{j=J_\Omega+1}^{J} \overline{\mathcal{E}}_j.$$

We achieve such an approximation by replacing curved corners of Ω_I by polyhedral corners, such that $\Omega_I \subset \Omega_I^{approx}$. Since we use the correct interaction set in the assembly process, the Dirichlet data g is not evaluated on $\Omega_I^{approx}\backslash\Omega_I$ so that we do not need to extend g to Ω_I^{approx}. For simplicity we from now on refer to Ω_I^{approx} as Ω_I. See Figure 4.1 for an illustration of Ω_I^{approx}. Furthermore we assume that the two triangulations $\mathcal{T}_\Omega^h$ and $\mathcal{T}_{\Omega_I}^h$ coincide at their common boundary $\overline{\Omega} \cap \overline{\Omega}_I$ ensuring that $\mathcal{T}^h := \mathcal{T}_\Omega^h \cup \mathcal{T}_{\Omega_I}^h$ defines a triangulation of $\Omega \cup \Omega_I$. By "coincide" we mean that the vertices of the finite element mesh $\mathcal{T}_\Omega^h$ and the vertices of $\mathcal{T}_{\Omega_I}^h$ are identical on $\overline{\Omega} \cap \overline{\Omega}_I$. See Figure 4.2 for a feasible configuration of such a matching grid.
Let $\{\mathbf{x}_k\}_{k=1}^K$ denote the mesh vertices of the triangulation $\mathcal{T}^h$. Then this kind of separate but matching triangulation of Ω and Ω_I naturally leads to the decomposition

$$\{\mathbf{x}_k\}_{k=1}^K = \{\mathbf{x}_k\}_{k=1}^{K_\Omega} \dot\cup \{\mathbf{x}_j\}_{k=K_\Omega+1}^K \tag{4.3}$$

of the mesh vertices. Here, $\{\mathbf{x}_k\}_{k=1}^{K_\Omega}$ contains all *interior* vertices of Ω whereas $\{\mathbf{x}_k\}_{k=K_\Omega+1}^K$ contains all vertices in $\Omega^c \cap \overline{\Omega}_I$ including the joint vertices on $\overline{\Omega} \cap \overline{\Omega}_I$.

Now let us assign to each vertex of the triangulation $\mathcal{T}^h$ a continuous piecewise polynomial function $\varphi_k \colon \Omega \cup \Omega_I \to \mathbb{R}$ leading to a family of basis function $\{\varphi_k\}_{k=1}^K$. We require $\varphi_k(\mathbf{x}_j) = \delta_{kj}$, where δ_{kj} denotes the Kroenecker delta function, and define the finite-dimensional finite element spaces

$$V^h := \operatorname{span}\{\varphi_k \colon 1 \le k \le K\} \quad \text{and} \quad V_c^h := \operatorname{span}\{\varphi_k \colon 1 \le k \le K_\Omega\}. \tag{4.4}$$

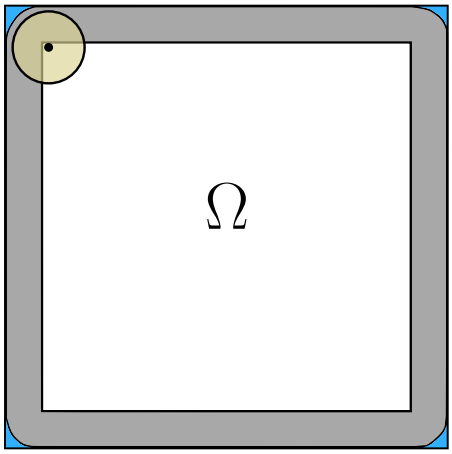


Figure 4.1: Illustration of Ω_I^{approx}. The gray colored area represents the true interaction domain Ω_I of a square domain Ω. A rectangular approximation $\Omega_I^{approx} \supset \Omega_I$ is achieved by replacing the curved corners by polygonal ones. The additional area $\Omega_I^{approx} \backslash \Omega_I$ is colored in blue. However, during the assembly process we use the original interaction set (here the yellow disc as an example), which does not intersect the blue area.

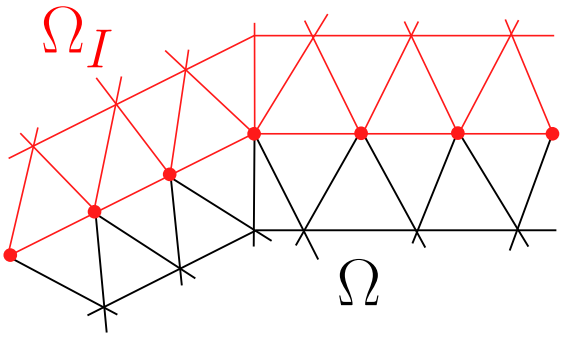


Figure 4.2: Illustration of a matching grid. The black mesh indicates the triangulation $\mathcal{T}_\Omega^h$ of the domain Ω and the red mesh indicates the triangulation $\mathcal{T}_{\Omega_I}^h$ of the interaction domain. At their common boundary $\overline{\Omega} \cap \overline{\Omega}_I$ they share the same vertices (here red dots) so that $\mathcal{T}^h = \mathcal{T}_\Omega^h \cup \mathcal{T}_{\Omega_I}^h$ defines a proper triangulation of $\Omega \cup \Omega_I$. Also note that we assign the nodes on $\overline{\Omega} \cap \overline{\Omega}_I$ to the triangulation of Ω_I as indicated by the red color.

Remark 4.2.1 (***Finite elements used in this thesis***)**.** *In all numerical experiments throughout this thesis we use either P_1 elements on triangular meshes or Q_1 elements on rectangular meshes (see, e.g., [44, pp. 22–23] for their definitions). Nonetheless the following considerations hold in general.*

The interpolation $u^h \in V^h$ of the weak solution $u \in V(\Omega \cup \Omega_I)$ effected by the finite-dimensional subspace V^h is given by

$$u^h(\mathbf{x}) = \sum_{k=1}^{K} u_k^h \varphi_k(\mathbf{x}) \quad (\mathbf{x} \in \Omega \cup \Omega_I),$$

where $u_k^h := u(\mathbf{x}_k)$. Since $u = g$ on Ω_I we can employ the decomposition of the vertices and obtain

$$u^h(\mathbf{x}) = \underbrace{\sum_{k=1}^{K_\Omega} u_k^h \varphi_k(\mathbf{x})}_{=:u_\Omega^h(\mathbf{x})} + \underbrace{\sum_{k=K_\Omega+1}^{K} g_k^h \varphi_k(\mathbf{x})}_{=:g^h(\mathbf{x})} \quad (\mathbf{x} \in \Omega \cup \Omega_I), \tag{4.5}$$

where $g_k^h := g(\mathbf{x}_k)$. Clearly u_Ω^h, i.e., the discrete finite element solution restricted to the interior of the domain Ω, defines the degree of freedom and is an element of V_c^h. According to the discrete variational formulation (4.1) the approximation u_Ω^h is determined by

$$A(u_\Omega^h, v^h) = \ell(v^h) = (f, v^h)_{L^2(\Omega)} - (\mathcal{L}_{\Omega_I} g^h, v^h)_{L^2(\Omega)} \quad \text{for all } v^h \in V_c^h, \tag{4.6}$$

where $\mathcal{L}_{\Omega_I}$ is defined in (3.1). Let us define

$$A_{\Omega\Omega_I}(u, v) := (\mathcal{L}_{\Omega_I} u, v)_{L^2(\Omega)}.$$

Then due to the definitions of V_c^h given in (4.4) as well as u_Ω^h and g^h given in (4.5) the system (4.6) is equivalent to

$$\sum_{k=1}^{K_\Omega} u_k^h A(\varphi_k, \varphi_j) = (f, \varphi_j)_{L^2(\Omega)} - \sum_{k=K_\Omega+1}^{K} g_k^h A_{\Omega\Omega_I}(\varphi_k, \varphi_j) \quad \text{for all } 1 \le j \le K_\Omega. \tag{4.7}$$

Let us define

$$\begin{aligned}
\mathbf{A}_{\Omega\Omega}^h &:= (A(\varphi_k, \varphi_j))_{(1 \le k,j \le K_\Omega)} \in \mathbb{R}^{K_\Omega \times K_\Omega}, \\
\mathbf{A}_{\Omega\Omega_I}^h &:= (A_{\Omega\Omega_I}(\varphi_k, \varphi_j))_{(1 \le k \le K_\Omega,\ K_\Omega+1 \le j \le K)} \in \mathbb{R}^{K_\Omega \times (K - K_\Omega)}, \\
\mathbf{u}_\Omega^h &:= \left(u_k^h\right)_{1 \le k \le K_\Omega} \in \mathbb{R}^{K_\Omega}, \\
\mathbf{f}^h &:= \left((f, \varphi_j)_{L^2(\Omega)}\right)_{1 \le j \le K_\Omega} \in \mathbb{R}^{K_\Omega}, \\
\mathbf{g}^h &:= \left(g_k^h\right)_{K_\Omega+1 \le k \le K} \in \mathbb{R}^{(K - K_\Omega)}.
\end{aligned} \tag{4.8}$$

Then we can rewrite (4.7) more concisely into the discrete system

$$\mathbf{A}^h_{\Omega\Omega}\mathbf{u}^h_\Omega = \mathbf{f}^h - \mathbf{A}^h_{\Omega\Omega_I}\mathbf{g}^h \tag{4.9}$$

with unknown $\mathbf{u}^h_\Omega \in \mathbb{R}^{K_\Omega}$, or even more condensed as

$$\mathbf{A}^h \begin{pmatrix} \mathbf{u}^h_\Omega \\ \mathbf{g}^h \end{pmatrix} = \mathbf{f}^h,$$

where $\mathbf{A}^h := \left(\mathbf{A}^h_{\Omega\Omega}, \mathbf{A}^h_{\Omega\Omega_I}\right) \in \mathbb{R}^{K_\Omega \times K}$. The entries a_{kj} of $\mathbf{A}^h$ are given by

$$\begin{aligned} a_{kj} &= A(\varphi_k, \varphi_j) + A_{\Omega\Omega_I}(\varphi_k, \varphi_j) = (-\mathcal{L}\varphi_k, \varphi_j)_{L^2(\Omega)} \\ &= \int_\Omega \varphi_j \int_{\Omega\cup\Omega_I} (\varphi_k\gamma - \varphi_k{}'\gamma')d\mathbf{y}d\mathbf{x} \end{aligned} \tag{4.10}$$

for $1 \le k \le K$, i.e., $\varphi_k \in V^h$, and $1 \le j \le K_\Omega$, i.e., $\varphi_j \in V^h_c$. We refer to $\mathbf{A}^h$ as *nonlocal stiffness matrix*. Since we consider kernels that satisfy (K2) such that $\gamma = \phi\chi_{S(\cdot)}$ we finally arrive at

$$a_{kj} = \int_\Omega \varphi_j \int_{S(\mathbf{x})} (\varphi_k\phi - \varphi_k{}'\phi')d\mathbf{y}d\mathbf{x}. \tag{4.11}$$

In the next two subsections we discuss this term for the two types of kernels introduced in the previous chapter. Before doing so, we end this part with some remarks.

Remark 4.2.2 (***Role of the Dirichlet data***)**.** *Since the Dirichlet data g is a given system parameter, we may think of using other approaches to compute the term $\mathcal{L}_{\Omega_I}g$ in* (4.6) *(instead of interpolating g by utilizing the finite element mesh). In this case we would only need to assemble the part $\mathbf{A}^h_{\Omega\Omega}$ of the nonlocal stiffness matrix as well as the adjusted right-hand side $\tilde{f} = f - \mathcal{L}_{\Omega_I}g$. However, note that the entries $a_{kj} = A(\varphi_k, \varphi_j)$ of $\mathbf{A}^h_{\Omega\Omega}$ given in* (4.10) *still involve integrals over Ω_I which is why we prefer to use a finite element mesh for the whole domain $\Omega \cup \Omega_I$. In the homogeneous case, i.e., $g = 0$, system* (4.9) *reduces to*

$$\mathbf{A}^h_{\Omega\Omega}\mathbf{u}^h_\Omega = \mathbf{f}^h$$

so that we only need to assemble $\mathbf{A}^h_{\Omega\Omega}$ in this case.

Remark 4.2.3 (***Unbounded interaction sets***)**.** *For the derivation of the nonlocal stiffness matrix there is no need for interaction sets to be bounded so that the finite element discretization of a nonlocal model associated to a kernel with infinite interactions, i.e., $S(\mathbf{x}) = \mathbb{R}^d$, leads to the same representations of the matrix entries. However, in this case $\Omega_I = \Omega^c$ which necessitates a strategy to cope with infinite domains within the finite element implementation. Also, a well-posedness theory different to the one developed in Chapter 3 is needed.*

Remark 4.2.4 (***Continuous interpolation***)**.** *We restrict ourselves to continuous basis functions, so that the interpolants g^h and f^h of g and f, respectively, are necessarily continuous. However, the well-posedness theory from Chapter 3 merely requires square integrability, so that g and f may have discontinuities. The same holds for the solution u and its interpolant u^h in case of integrable kernels and for kernels of fractional type where $s \leq 0.5$ [36].*

4.2.1 Integrable kernels

For the class of integrable kernels discussed in Subsection 3.3.1 we can split the inner integral in (4.11) so that

$$a_{kj} = \int_\Omega \varphi_j \varphi_k \int_{S(\mathbf{x})} \phi \, d\mathbf{y} d\mathbf{x} - \int_\Omega \varphi_j \int_{S(\mathbf{x})} \varphi_k{}' \phi' \, d\mathbf{y} d\mathbf{x}. \tag{4.12}$$

We typically consider basis functions which have compact support consisting of a connected patch of elements (see also Remark 4.2.1). In order to compute the occurring double integrals in (4.12) it is therefore desirable to determine the support of the integrand such that zero evaluations are avoided. Also, this enables us to assign integrals over finite elements to the support of the respective basis functions and thereby assign the computed integral value to the correct entry in the nonlocal stiffness matrix; further details are provided in Section 4.3. By setting $\mathcal{S}_k := \mathrm{supp}(\varphi_k)$, we find

$$a_{kj} = \int_{\mathcal{S}_k \cap \mathcal{S}_j} \varphi_j \varphi_k \int_{S(\mathbf{x})} \phi \, d\mathbf{y} d\mathbf{x} - \int_{\mathcal{S}_j} \varphi_j \int_{\mathcal{S}_k \cap S(\mathbf{x})} \varphi_k{}' \phi' \, d\mathbf{y} d\mathbf{x}. \tag{4.13}$$

Remark 4.2.5. *We use representation* (4.13) *for implementing the nonlocal stiffness matrix associated to integrable kernels.*

4.2.2 Singular kernels

For the class of singular and symmetric kernels discussed in Subsection 3.3.2 the nonlocal model (2.1) reduces to a pure diffusion model. Thus the nonlocal bilinear form reduces to $A = A^d$, where A^d is the diffusion part defined in (3.8). We can therefore consider

$$\begin{aligned} a_{kj} &= A^d(\varphi_k, \varphi_j) + A_{\Omega\Omega_I}(\varphi_k, \varphi_j) \\ &= \int_\Omega \varphi_j \int_{\Omega \cap S(\mathbf{x})} (\varphi_k - \varphi_k{}') \, \phi d\mathbf{y} d\mathbf{x} + \int_\Omega \varphi_j \varphi_k \int_{\Omega_I \cap S(\mathbf{x})} \phi d\mathbf{y} d\mathbf{x} \\ &\quad - \int_\Omega \varphi_j \int_{\Omega_I \cap S(\mathbf{x})} \varphi_k{}' \phi d\mathbf{y} d\mathbf{x}, \end{aligned} \tag{4.14}$$

for $1 \leq k \leq K$, i.e., $\varphi_k \in V^h$, and $1 \leq j \leq K_\Omega$, i.e., $\varphi_j \in V_c^h$. Note that the last integral in (4.14) vanishes in the homogeneous setting, i.e., if $\varphi_k \in V_c^h$. Since the kernel function

may have singularities at $(\mathbf{x},\mathbf{x})$ (see, e.g., (3.18)) we cannot split the first inner integral in (4.14). Instead, in order to increase the mollifying effect of the term $(\varphi_k - \varphi_k{}')$ in the integrand, in numerical simulations it is reasonable to consider the equivalent representation of this term given in (3.9) leading to

$$\begin{aligned} a_{kj} = \frac{1}{2}\int_\Omega \int_{\Omega\cap S(\mathbf{x})} (\varphi_j - \varphi_j{}')\,(\varphi_k - \varphi_k{}')\,\phi d\mathbf{y}d\mathbf{x} + \int_\Omega \varphi_j\varphi_k \int_{\Omega_I\cap S(\mathbf{x})} \phi d\mathbf{y}d\mathbf{x} \\ - \int_\Omega \varphi_j \int_{\Omega_I\cap S(\mathbf{x})} \varphi_k{}'\phi d\mathbf{y}d\mathbf{x}. \end{aligned} \tag{4.15}$$

For specifying the domain of integration in terms of the supports of the basis functions, we need to have a closer look at the integrand $(\varphi_j - \varphi_j{}')\,(\varphi_k - \varphi_k{}')\,\phi$ from the first summand in (4.15). Since ϕ is assumed to be positive (see (K2)), we find

$$\begin{aligned} &(\varphi_k(\mathbf{y}) - \varphi_k(\mathbf{x}))(\varphi_j(\mathbf{y}) - \varphi_j(\mathbf{x})) = 0 \\ \Leftrightarrow\ &(\mathbf{x},\mathbf{y}) \in (\mathcal{S}_k^c \times \mathcal{S}_k^c) \cup \left(\mathcal{S}_j^c \times \mathcal{S}_j^c\right) \cup \{(\mathbf{x},\mathbf{x})\colon \mathbf{x}\in\Omega\cup\Omega_I\}\,. \end{aligned} \tag{4.16}$$

Because $\{(\mathbf{x},\mathbf{x})\colon \mathbf{x}\in\Omega\cup\Omega_I\}$ has zero $2d$-dimensional volume we can neglect it in the integral. Further we find

$$\begin{aligned} &\left((\mathcal{S}_k^c \times \mathcal{S}_k^c) \cup \left(\mathcal{S}_j^c \times \mathcal{S}_j^c\right)\right)^c = (\mathcal{S}_k^c \times \mathcal{S}_k^c)^c \cap \left(\mathcal{S}_j^c \times \mathcal{S}_j^c\right)^c \\ =&\big((\mathcal{S}_k \times \mathcal{S}_k) \cup (\mathcal{S}_k^c \times \mathcal{S}_k) \cup (\mathcal{S}_k \times \mathcal{S}_k^c)\big) \cap \left((\mathcal{S}_j \times \mathcal{S}_j) \cup (\mathcal{S}_j^c \times \mathcal{S}_j) \cup (\mathcal{S}_j \times \mathcal{S}_j^c)\right) \\ =&(I\times I) \cup (D_k \times I) \cup (I \times D_k) \cup (D_j \times I) \cup (I \times D_j) \\ &\cup (D_j \times D_k) \cup (D_k \times \mathcal{S}_j) \cup (C \times I) \cup (I \times C), \end{aligned}$$

where we set

$$I := \mathcal{S}_k \cap \mathcal{S}_j, \quad D_k := \mathcal{S}_k \cap \mathcal{S}_j^c, \quad D_j := \mathcal{S}_k^c \cap \mathcal{S}_j \quad \text{and} \quad C := (\mathcal{S}_k \cup \mathcal{S}_j)^c.$$

By exploiting the symmetry of the integrand $(\varphi_j - \varphi_j{}')\,(\varphi_k - \varphi_k{}')\,\phi$ in $(\mathbf{x},\mathbf{y})$ and noting that $\mathcal{S}_j \subset \Omega$ for $\varphi_j \in V_c^h$, we find

$$\begin{aligned} a_{kj} = \frac{1}{2}\Bigg(&\int_{\mathcal{S}_k\cap\mathcal{S}_j} \int_{(\mathcal{S}_k\cap\mathcal{S}_j)\cap S(\mathbf{x})} (\varphi_k - \varphi_k{}')(\varphi_j - \varphi_j{}')\phi(\mathbf{x},\mathbf{y}) d\mathbf{y}d\mathbf{x} \\ &- 2\int_{\mathcal{S}_k\cap\mathcal{S}_j} \varphi_j \int_{(\mathcal{S}_k\setminus\mathcal{S}_j)\cap\Omega\cap S(\mathbf{x})} \varphi_k{}'\phi d\mathbf{y}d\mathbf{x} \\ &- 2\int_{\mathcal{S}_j\setminus\mathcal{S}_k} \varphi_j \int_{\mathcal{S}_k\cap\Omega\cap S(\mathbf{x})} \varphi_k{}'\phi d\mathbf{y}d\mathbf{x} \\ &+ 2\int_{\mathcal{S}_k\cap\mathcal{S}_j} \varphi_j\varphi_k \int_{(\mathcal{S}_k\cap\mathcal{S}_j)^c\cap\Omega\cap S(\mathbf{x})} \phi d\mathbf{y}d\mathbf{x}\Bigg) \\ &+ \int_{\mathcal{S}_k\cap\mathcal{S}_j} \varphi_j\varphi_k \int_{\Omega_I\cap S(\mathbf{x})} \phi d\mathbf{y}d\mathbf{x} - \int_{\mathcal{S}_j} \varphi_j \int_{\mathcal{S}_k\cap\Omega_I\cap S(\mathbf{x})} \varphi_k{}'\phi d\mathbf{y}d\mathbf{x}. \end{aligned} \tag{4.17}$$

Furthermore, since $\mathcal{S}_k \cap \mathcal{S}_j \subset \Omega$ we can merge the fourth and fifth integrals in (4.17) leading to

$$\begin{aligned} & \int_{\mathcal{S}_k\cap\mathcal{S}_j} \varphi_j\varphi_k \int_{(\mathcal{S}_k\cap\mathcal{S}_j)^c\cap\Omega\cap S(\mathbf{x})} \phi d\mathbf{y}d\mathbf{x} + \int_{\mathcal{S}_k\cap\mathcal{S}_j} \varphi_j\varphi_k \int_{\Omega_I\cap S(\mathbf{x})} \phi d\mathbf{y}d\mathbf{x} \\ = & \int_{\mathcal{S}_k\cap\mathcal{S}_j} \varphi_j\varphi_k \int_{(\mathcal{S}_k\cap\mathcal{S}_j)^c\cap S(\mathbf{x})} \phi d\mathbf{y}d\mathbf{x}. \end{aligned}$$

Also, by decomposing the last integral in (4.17) into

$$\begin{aligned} & \int_{\mathcal{S}_j} \varphi_j \int_{\mathcal{S}_k\cap\Omega_I\cap S(\mathbf{x})} {\varphi_k}'\phi d\mathbf{y}d\mathbf{x} \\ = & \int_{\mathcal{S}_k\cap\mathcal{S}_j} \varphi_j \int_{\mathcal{S}_k\cap\Omega_I\cap S(\mathbf{x})} {\varphi_k}'\phi d\mathbf{y}d\mathbf{x} + \int_{\mathcal{S}_j\setminus\mathcal{S}_k} \varphi_j \int_{\mathcal{S}_k\cap\Omega_I\cap S(\mathbf{x})} {\varphi_k}'\phi d\mathbf{y}d\mathbf{x}, \end{aligned}$$

we can sum the second, third and last integral in (4.17) as follows

$$\begin{aligned} & \int_{\mathcal{S}_k\cap\mathcal{S}_j} \varphi_j \int_{(\mathcal{S}_k\setminus\mathcal{S}_j)\cap\Omega\cap S(\mathbf{x})} {\varphi_k}'\phi d\mathbf{y}d\mathbf{x} + \int_{\mathcal{S}_j\setminus\mathcal{S}_k} \varphi_j \int_{\mathcal{S}_k\cap\Omega\cap S(\mathbf{x})} {\varphi_k}'\phi d\mathbf{y}d\mathbf{x} \\ & + \int_{\mathcal{S}_j} \varphi_j \int_{\mathcal{S}_k\cap\Omega_I\cap S(\mathbf{x})} {\varphi_k}'\phi d\mathbf{y}d\mathbf{x} \\ = & \int_{\mathcal{S}_k\cap\mathcal{S}_j} \varphi_j \int_{(\mathcal{S}_k\setminus\mathcal{S}_j)\cap S(\mathbf{x})} {\varphi_k}'\phi d\mathbf{y}d\mathbf{x} + \int_{\mathcal{S}_j\setminus\mathcal{S}_k} \varphi_j \int_{\mathcal{S}_k\cap S(\mathbf{x})} {\varphi_k}'\phi d\mathbf{y}d\mathbf{x}. \end{aligned}$$

These considerations finally lead to

$$\begin{aligned} a_{kj} = & \frac{1}{2} \int_{\mathcal{S}_k\cap\mathcal{S}_j} \int_{(\mathcal{S}_k\cap\mathcal{S}_j)\cap S(\mathbf{x})} (\varphi_k - {\varphi_k}')(\varphi_j - {\varphi_j}')\phi(\mathbf{x},\mathbf{y})\ d\mathbf{y}d\mathbf{x} \\ & + \int_{\mathcal{S}_k\cap\mathcal{S}_j} \varphi_j\varphi_k \int_{(\mathcal{S}_k\cap\mathcal{S}_j)^c\cap S(\mathbf{x})} \phi\ d\mathbf{y}d\mathbf{x} \\ & - \int_{\mathcal{S}_k\cap\mathcal{S}_j} \varphi_j \int_{(\mathcal{S}_k\setminus\mathcal{S}_j)\cap S(\mathbf{x})} {\varphi_k}'\phi\ d\mathbf{y}d\mathbf{x} \\ & - \int_{(\mathcal{S}_j\setminus\mathcal{S}_k)} \varphi_j \int_{\mathcal{S}_k\cap S(\mathbf{x})} {\varphi_k}'\phi\ d\mathbf{y}d\mathbf{x}. \end{aligned} \tag{4.18}$$

See Figure 4.3 for an exemplary configuration of basis function supports and the resulting integration domains appearing in (4.13) and (4.18).

Remark 4.2.6. *We use representation* (4.18) *for the assembly of the nonlocal stiffness matrix whenever symmetric singular kernels are involved. We note that for symmetric integrable kernels the representations* (4.13) *and* (4.18) *are equivalent.*

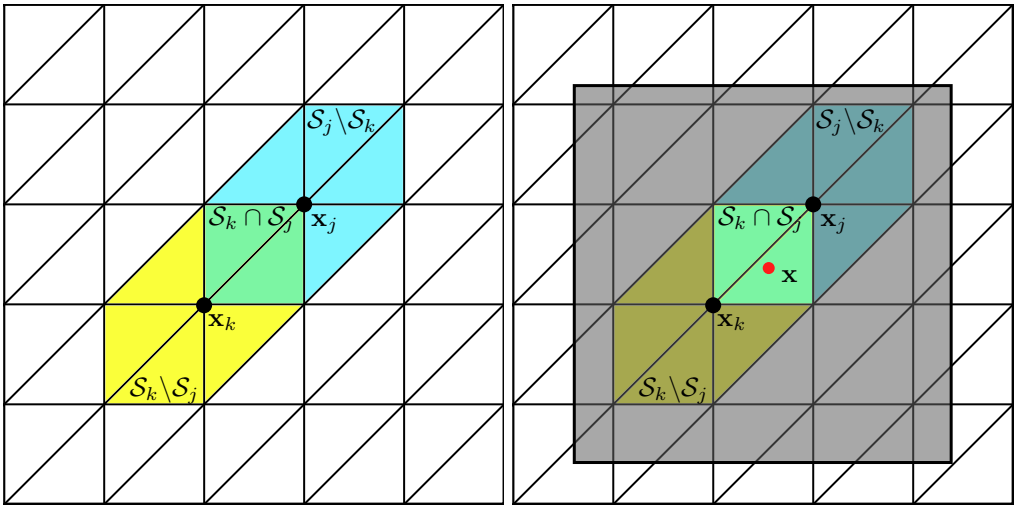


Figure 4.3: Overlapping supports of two basis functions on a triangular mesh and resulting integration domains appearing in (4.13) and (4.18). In both images we find the yellow colored support $\mathcal{S}_k$ of the basis function φ_k centered at $\mathbf{x}_k$ and likewise cyan colored for index j. Their intersection $\mathcal{S}_k \cap \mathcal{S}_j$ is colored green. Furthermore, on the right-hand side the integration domain of the second integral in (4.18) is exemplarily depicted. More precisely, the outer integral integrates over points $\mathbf{x}$ in the green colored intersection $\mathcal{S}_k \cap \mathcal{S}_j$, whereas the inner integral integrates over points $\mathbf{y}$ in the shady domain $(\mathcal{S}_k \cap \mathcal{S}_j)^c \cap S(\mathbf{x})$ (here $S(\mathbf{x}) = B_{\delta,\infty}(\mathbf{x})$).

4.3 Assembly and related challenges

In this section we discuss in detail the assembly of the nonlocal stiffness matrix $\mathbf{A}^h$ and reveal challenges germane to this endeavor. The assembly of the forcing term vector $\mathbf{f}^h$ is standard and therefore not discussed here.
In terms of the triangulation $\mathcal{T}^h$ of $\Omega \cup \Omega_I$ the stiffness matrix entry as given in (4.11) can also be written as

$$a_{kj} = \sum_{a=1}^{J_\Omega} \sum_{b=1}^{J} \int_{\mathcal{E}_a} \varphi_j \int_{\mathcal{E}_b \cap S(\mathbf{x})} (\varphi_k \phi - \varphi_k{}' \phi') d\mathbf{y} d\mathbf{x}.$$

Let us define for an element $\mathcal{E} \in \mathcal{T}^h$ and a family of interaction sets $\{S(\mathbf{x})\}_{\mathbf{x} \in \mathbb{R}^d}$ the set

$$S(\mathcal{E}) = \left\{ b \in \{1, \dots, J\} : |\mathcal{E}_I \cap \overline{\mathcal{E}_b}| \neq 0 \right\}, \tag{4.19}$$

where $|\cdot|$ denotes the d-dimensional volume measure and $\mathcal{E}_I$ denotes the interaction domain of the element $\mathcal{E}$, given by

$$\mathcal{E}_I = \{ \mathbf{y} \in \mathcal{E}^c : \ \mathbf{y} \in S(\mathbf{x}) \text{ for some } \mathbf{x} \in \mathcal{E} \}.$$

Then we can also write

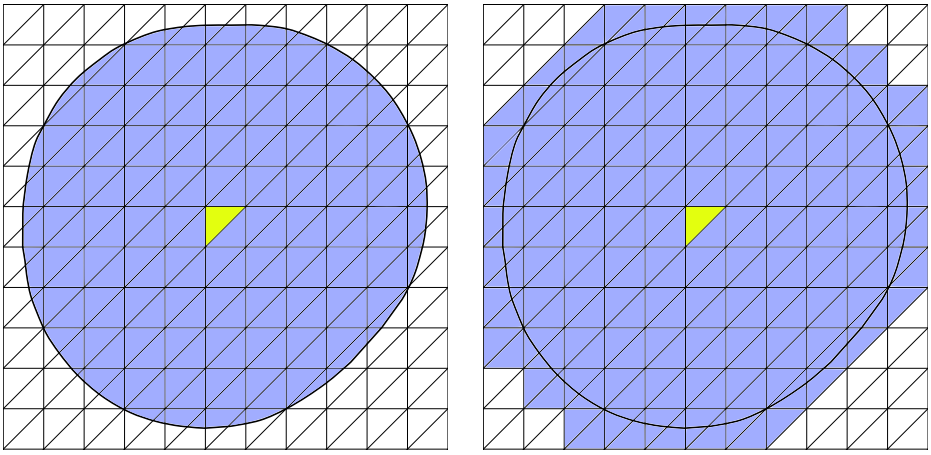


Figure 4.4: On the left-hand side the interaction domain $\mathcal{E}_I$ of the yellow triangle $\mathcal{E}$ with respect to Euclidean balls is depicted. On the right-hand side all triangles belonging to $S(\mathcal{E})$ are colored blueish, i.e., all triangles whose intersections with $\mathcal{E}_I$ have nonzero d-dimensional volume.

$$a_{kj} = \sum_{a=1}^{J_\Omega} \sum_{b \in S(\mathcal{E}_a)} \int_{\mathcal{E}_a} \varphi_j \int_{\mathcal{E}_b \cap S(\mathbf{x})} (\varphi_k \phi - \varphi_k{}' \phi') d\mathbf{y} d\mathbf{x}. \tag{4.20}$$

In Figure 4.4 the sets $\mathcal{E}_I$ and $S(\mathcal{E})$ are illustrated. We say that two elements $\mathcal{E}_a, \mathcal{E}_b \in \mathcal{T}^h$ *interact* with each other if $b \in S(\mathcal{E}_a)$ or equivalently $a \in S(\mathcal{E}_b)$ (due to (S2)), because they cause a nonzero contribution to the stiffness matrix in this case. Now, the two representations (4.11) and (4.20) suggest two different ways of implementing the assembly of the nonlocal stiffness matrix:

- ***Loop over basis functions:*** We can iterate through the indices $1 \leq k \leq K$ and $1 \leq j \leq K_\Omega$ and fully compute each entry $a_{kj} = (-\mathcal{L}\varphi_k, \varphi_j)_{L^2(\Omega)}$. Thus for a given pair (k, j) we have to determine the sets of element indices $E_k \subset \{1, \dots, J\}$ and $E_j \subset \{1, \dots, J_\Omega\}$ which build the supports $\mathcal{S}_k$ and $\mathcal{S}_j$ of the respective basis functions φ_k and φ_j, i.e., $\mathcal{S}_k = \bigcup_{a \in E_k} \mathcal{E}_a$ and similarly for φ_j. Then we can use the representations (4.13) and (4.18) to boil down the integrals over the supports to integrals over the elements. For example, let us consider (4.13), then we can write

$$\begin{aligned} a_{kj} = &\sum_{a \in E_k \cap E_j} \sum_{b \in S(\mathcal{E}_a)} \int_{\mathcal{E}_a} \varphi_j \varphi_k \int_{\mathcal{E}_b \cap S(\mathbf{x})} \phi \, d\mathbf{y} d\mathbf{x} \\ &- \sum_{a \in E_j} \sum_{b \in S(\mathcal{E}_a) \cap E_k} \int_{\mathcal{E}_a} \varphi_j \int_{\mathcal{E}_b \cap S(\mathbf{x})} \varphi_k{}' \phi' \, d\mathbf{y} d\mathbf{x}. \end{aligned}$$

  We can apply the same procedure to express the integrals in (4.18) in terms of integrals over the elements.

- ***Loop over elements:*** On the other hand, representation (4.20) advocates to iterate through the indices of the elements. In this case, for a given pair of elements $(\mathcal{E}_a, \mathcal{E}_b)$ with $1 \leq a \leq J_\Omega$ and $b \in S(\mathcal{E}_a) \subset \{1, \dots, J\}$, we need to determine the set of basis function indices

  $$H_{a,b} := \left\{ (k,j) \in \{1,\dots,K\} \times \{1,\dots,K_\Omega\} : \int_{\mathcal{E}_a} \varphi_j \int_{\mathcal{E}_b \cap S(\mathbf{x})} (\varphi_k \phi - \varphi_k' \phi') d\mathbf{y} d\mathbf{x} \neq 0 \right\}.$$

  Indices $(k,j) \in H_{a,b}$ then have a nonzero contribution to the entry a_{kj}. The representations (4.13) and (4.18) of a_{kj} can be utilized to determine the set $H_{a,b}$. For example let us again consider the representation (4.13) for integrable kernels and treat each of the two integrals separately. Then for the first integral in (4.13) we find

  $$\begin{aligned} H_{a,b}^1 &:= \left\{ (k,j) : \int_{\mathcal{E}_a} \varphi_k \varphi_j \int_{\mathcal{E}_b \cap S(\mathbf{x})} \phi \; d\mathbf{y} d\mathbf{x} \neq 0 \right\} \\ &= \left\{ (k,j) \colon \mathbf{x}_k \in \overline{\mathcal{E}}_a \text{ and } \mathbf{x}_j \in \overline{\mathcal{E}}_a \right\}. \end{aligned}$$

  We observe that the set $H_{a,b}^1$ only depends on a and thereby only accounts for local interactions as the double integral of interest is a weighted L^2-product. For the second integral in (4.13) we find

  $$\begin{aligned} H_{a,b}^2 &:= \left\{ (k,j) \colon \int_{\mathcal{E}_a} \varphi_j \int_{\mathcal{E}_b \cap S(\mathbf{x})} \varphi_k' \phi' \; d\mathbf{y} d\mathbf{x} \neq 0 \right\} \\ &= \left\{ (k,j) \colon \mathbf{x}_k \in \overline{\mathcal{E}}_b \text{ and } \mathbf{x}_j \in \overline{\mathcal{E}}_a \right\}. \end{aligned}$$

  Here, $H_{a,b}^2$ depends on both elements $\mathcal{E}_a$ and $\mathcal{E}_b$ and therefore accounts for nonlocal interactions. All in all, we find $H_{a,b} = H_{a,b}^1 \cup H_{a,b}^2$.

In any case, for the assembly of the nonlocal stiffness matrix we need to compute double integrals over pairs of elements. The order in which we treat such pairs is determined through the approach chosen from the two assembly possibilities just mentioned. Therefore, the crucial task is to deal with integrals of the form

$$\int_{\mathcal{E}_a} \int_{\mathcal{E}_b \cap S(\mathbf{x})} \psi(\mathbf{x}, \mathbf{y}) d\mathbf{y} d\mathbf{x} \tag{4.21}$$

where $\psi = \psi_{k,j,\phi} \colon (\Omega \cup \Omega_I) \times (\Omega \cup \Omega_I) \to \mathbb{R}$ shall serve as placeholder for the various integrands appearing in (4.13) and (4.18). In the following two subsections we discuss the quadrature of this double integral. More precisely, we subsequently integrate the outer integral and the inner integral. For this purpose we denote by $\left\{ \mathbf{x}_q^a, w_q^a \right\}_{q=1}^Q$ a quadrature rule for a physical element $\mathcal{E}_a$. As usual such quadrature rules are deduced through transformation from a quadrature rule defined on an appropriate reference element. We may employ quadrature rules of different types for the outer and inner integral, respectively, what we simply indicate by different orders $Q = Q_{out}$ and $Q = Q_{in}$, respectively.

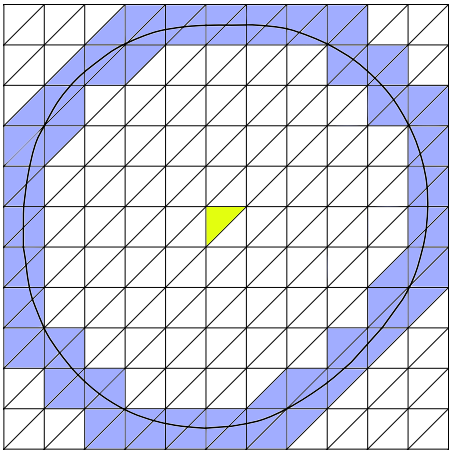


Figure 4.5: Illustration of the set $\partial S(\mathcal{E})$. Here, $\mathcal{E}$ is given by the yellow colored triangle. The set $\partial S(\mathcal{E})$ consists of all blueish colored triangles, which are intersected by the boundary of the interaction domain $\mathcal{E}_I$ (black line).

4.3.1 The outer integral

Let us denote the integrand of the outer integral in (4.21) by

$$\mathcal{I}_b(\mathbf{x}) := \int_{\mathcal{E}_b \cap S(\mathbf{x})} \psi(\mathbf{x}, \mathbf{y}) d\mathbf{y}, \tag{4.22}$$

so that the respective double integral can be written as

$$\int_{\mathcal{E}_a} \int_{\mathcal{E}_b \cap S(\mathbf{x})} \psi(\mathbf{x}, \mathbf{y}) d\mathbf{y} d\mathbf{x} = \int_{\mathcal{E}_a} \mathcal{I}_b(\mathbf{x}) d\mathbf{x}.$$

For a fixed element $\mathcal{E}_a$ we distinguish two different cases for $\mathcal{E}_b$ in (4.22). For identifying those cases let us define the set

$$\partial S(\mathcal{E}) := \left\{ b \in S(\mathcal{E}) \colon \mathcal{E}_I \cap \overline{\mathcal{E}_b} \neq \mathcal{E}_b \right\},$$

which is pictured in Figure 4.5. The first case is $b \in S(\mathcal{E}_a) \backslash \partial S(\mathcal{E}_a)$ in which $\mathcal{I}_b(\mathbf{x}) \neq 0$ for *all* $\mathbf{x} \in \mathcal{E}_a$. This is the simple case, since we can employ the standard element quadrature rule $\left\{ \mathbf{x}_q^a, w_q^a \right\}_{q=1}^{Q_{out}}$ for the discretization of the outer integral. Now let us treat the second case $b \in \partial S(\mathcal{E}_a)$. Here we find that the following subset

$$\mathcal{E}_a^0 := \{ \mathbf{x} \in \mathcal{E}_a \colon \mathcal{E}_b \cap S(\mathbf{x}) = \emptyset \} \subset \mathcal{E}_a$$

has nonzero volume in $\mathbb{R}^d$. Consequently $\mathcal{I}_b(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathcal{E}_a^0$. In this case $\mathcal{I}_b$ numerically behaves like a discontinuous function and we have to employ a carefully designed quadrature rule on $\mathcal{E}_a$. In fact, if we employ a standard element quadrature rule, then a significant number of quadrature points may be located in $\mathcal{E}_a^0$ leading to a

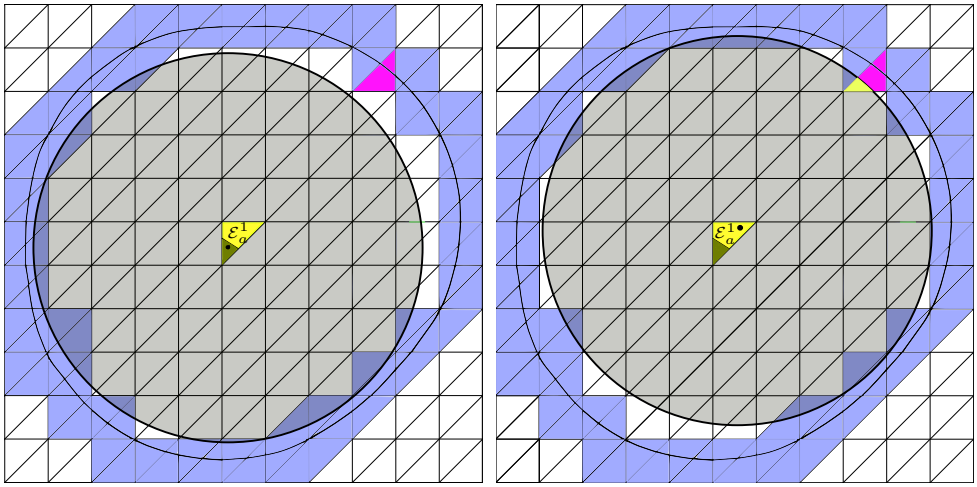


Figure 4.6: Illustration of the subsets $\mathcal{E}_a^1$ and $\mathcal{E}_a^0$. In both graphics $\mathcal{E}_a$ is given by the triangle in the center and $\mathcal{E}_b$ by the magenta colored triangle in $\partial S(\mathcal{E}_a)$. The shaded part of $\mathcal{E}_a$ indicates the subset $\mathcal{E}_a^0$, whereas the yellow part represents $\mathcal{E}_a^1$. In the left picture, the shady ball centered at a point in the shaded area $\mathcal{E}_a^0$ does not intersect with the magenta triangle $\mathcal{E}_b$. On the contrary, for balls centered at points in $\mathcal{E}_a^1$ this intersection is nonempty as illustrated in the right picture.

loss of accuracy. In Figure 4.6 this situation is depicted; we further define $\mathcal{E}_a^1 := \mathcal{E}_a \backslash \mathcal{E}_a^0 = \text{supp}(\mathcal{I}_b)$. Let us denote by $\left\{\mathbf{x}_q^a, w_q^a\right\}_{q=1}^{Q_{out}^{special}}$ a quadrature routine which considers the partition $\mathcal{E}_a = \mathcal{E}_a^0 \dot\cup \mathcal{E}_a^1$. Then we finally obtain a semi-discretization of (4.21), given by

$$\int_{\mathcal{E}_a} \mathcal{I}_b(\mathbf{x}) d\mathbf{x} \approx \begin{cases} \sum_{q=1}^{Q_{out}} \mathcal{I}_b(\mathbf{x}_q^a) w_q^a & : b \in S(\mathcal{E}_a) \backslash \partial S(\mathcal{E}_a) \\ \sum_{q=1}^{Q_{out}^{special}} \mathcal{I}_b(\mathbf{x}_q^a) w_q^a & : b \in \partial S(\mathcal{E}_a). \end{cases} \tag{4.23}$$

Remark 4.3.1. *For simplicity we use an adaptive quadrature rule which automatically takes care of the partition $\mathcal{E}_a = \mathcal{E}_a^0 \dot\cup \mathcal{E}_a^1$. In the case of norm induced interaction sets $S(\mathbf{x}) := B_{\delta,\bullet}(\mathbf{x})$ we activate the adaptive quadrature rule for pairs $(\mathcal{E}_a, \mathcal{E}_b)$ with $\|\mathbf{x}_a^{bary} - \mathbf{x}_b^{bary}\|_\bullet > \delta - h$. Here, $\mathbf{x}_a^{bary}$ denotes the barycenter of an element $\mathcal{E}_a$ and $h > 0$ the diameter of the largest element. With this heuristic criterion we might apply this special treatment to pairs $(\mathcal{E}_a, \mathcal{E}_b)$ with $b \notin \partial S(\mathcal{E}_a)$ (see the orange triangles in Figure 4.7). However, the adaptive quadrature rule detects such cases rather automatically and avoids a deep refinement there. Alternatively, we may also think of quadrature rules for the outer triangle $\mathcal{E}_a$ for which the quadrature points lie close or even on the boundary $\partial\mathcal{E}_a$ of the element and thereby avoiding that we miss out interacting triangles $\mathcal{E}_b$ for the inner integral with $b \in \partial S(\mathcal{E}_a)$. However, in this case a rigorous analysis for the minimum precision has to be conducted and is left to future work.*

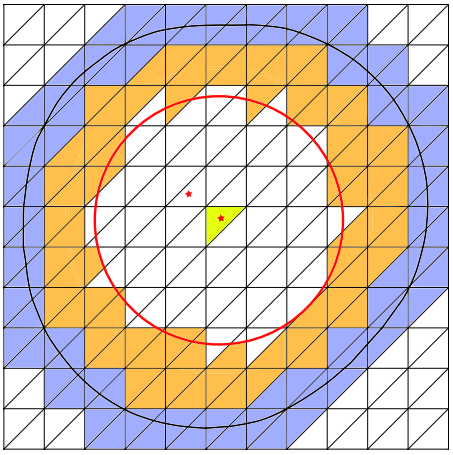


Figure 4.7: Illustration of the heuristic criterion discussed in Remark 4.3.1. The fixed element $\mathcal{E}_a$ of the outer integral is given by the yellow colored triangle with associated set $\partial S(\mathcal{E}_a)$ that consists of all blueish colored triangles. The red line indicates the boundary of the Euclidean ball of radius $\delta - h$ centered at the barycenter of $\mathcal{E}_a$ given by the red star. Although only necessary for the blueish triangles, we apply an adaptive quadrature on $\mathcal{E}_a$ whenever $\|\mathbf{x}_a^{bary} - \mathbf{x}_b^{bary}\|_2 > \delta - h$, thereby additionally including the orange colored triangles.

4.3.2 The inner integral

Let us continue with (4.23) to obtain a full discretization of the double integral (4.21). After we have discretized the outer integral, the next task is to apply an appropriate quadrature rule to the inner integral of (4.23), namely

$$\mathcal{I}_b(\mathbf{x}_q^a) = \int_{\mathcal{E}_b \cap S(\mathbf{x}_q^a)} \psi(\mathbf{x}_q^a, \mathbf{y}) d\mathbf{y}. \tag{4.24}$$

For each quadrature point $\mathbf{x}_q^a \in \mathcal{E}_a$ of the outer integral we have to integrate over the intersection $\mathcal{E}_b \cap S(\mathbf{x}_q^a)$. In analogy to the outer integral, there are two cases to consider. First, the intersection region may fully cover the element, i.e., $\mathcal{E}_b \cap S(\mathbf{x}_q^a) = \mathcal{E}_b$. In this situation we can employ a standard element quadrature rule on $\mathcal{E}_b$. Second, the intersection region may only partially cover the element, i.e., $\mathcal{E}_b \cap S(\mathbf{x}_q^a) \subsetneqq \mathcal{E}_b$. If we employ a standard element quadrature rule designed for the whole element $\mathcal{E}_b$ we would cause a significant loss of accuracy due to the discontinuous indicator function $\chi_{S(\mathbf{x}_q^a)}(\cdot)$. An accurate quadrature rule for the inner integral results from re-triangulating the intersection region into smaller *subelements*. Then we can derive a composite quadrature rule for $\mathcal{E}_b \cap S(\mathbf{x}_q^a)$ by applying a quadrature rule to each of these subelements. For the re-triangulation it is desirable to apply the same kind of elements as in $\mathcal{T}^h$, since we can then use the same type of element quadrature rules. However, the intersection region $\mathcal{E}_b \cap S(\mathbf{x}_q^a)$ does not have to be polyhedral. Therefore, in general we have to consider a partition of the intersection region into a polyhedral part and a nonpolyhedral part. The polyhedral part can then be covered exactly by appropriate polyhedral

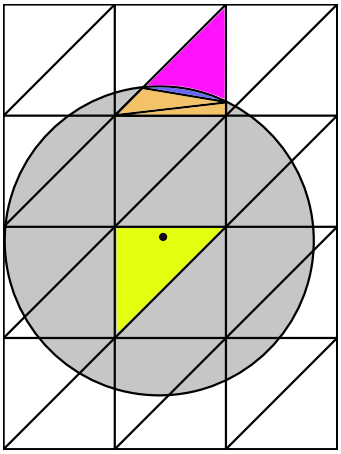


Figure 4.8: Subdivision of a partially covered triangle into $J_S = 3$ subelements. The fixed triangle $\mathcal{E}_a$ of the outer integral is depicted in yellow. The interaction set, here an Euclidean ball, centered at a point $\mathbf{x} \in \mathcal{E}_a$ (black dot) only partially covers the magenta colored triangle $\mathcal{E}_b$. The intersection region is subdivided into an orange polygonal part consisting of $J_S^{poly} = 2$ triangles and a blue nonpolygonal part consisting of $(J_S - J_S^{poly}) = 1$ circular segment.

elements (e.g., triangles); preferably those used in $\mathcal{T}^h$. The possibly curved or empty nonpolyhedral part requires different types of subelements. More precisely, let us assume that we have a triangulation of $\mathcal{E}_b \cap S(\mathbf{x}_q^a)$ into $J_S = J_{S,a,q,b} \geq 0$ subelements, where the first $J_S^{poly} = J_{S,a,q,b}^{poly} \leq J_S$ cover the delimited polyhedral part and the remaining $(J_S - J_S^{poly}) \geq 0$ cover the nonpolyhedral part. This leads to the following partition of the intersection region

$$\mathcal{E}_b \cap S(\mathbf{x}_q^a) = \bigcup_{c=1}^{J_S} \mathcal{E}_c = \left(\bigcup_{c=1}^{J_S^{poly}} \mathcal{E}_c \right) \dot{\cup} \left(\bigcup_{c=J_S^{poly}+1}^{J_S} \tilde{\mathcal{E}}_c \right). \tag{4.25}$$

This situation is depicted in Figure 4.8. Now let $\left\{ \mathbf{y}_p^c, w_p^c \right\}_{p=1}^{Q_{in}}$ and $\left\{ \mathbf{y}_p^c, w_p^c \right\}_{p=1}^{Q_{in}^{nonpoly}}$ denote a quadrature rule for subelements covering the polyhedral part and nonpolyhedral part, respectively. With this notation at hand we eventually obtain a discretized version of

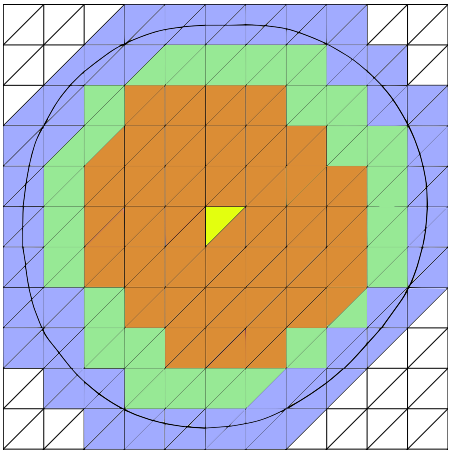


Figure 4.9: Illustration of the partition $S(\mathcal{E}) = \mathring{S}(\mathcal{E}) \dot{\cup} \tilde{S}(\mathcal{E}) \dot{\cup} \partial S(\mathcal{E})$ given in (4.27). The element $\mathcal{E}$ is indicated by the yellow triangle in the center of the image. The orange triangles belong to $\mathring{S}(\mathcal{E})$, the green ones to $\tilde{S}(\mathcal{E})$ and the remaining blue ones belong to $\partial S(\mathcal{E})$.

(4.24), given by

$$\begin{aligned}
\mathcal{I}_b(\mathbf{x}_q^a) &= \int_{\mathcal{E}_b \cap S(\mathbf{x}_q^a)} \psi(\mathbf{x}_q^a, \mathbf{y}) d\mathbf{y} \\
&= \sum_{c=1}^{J_S^{poly}} \int_{\mathcal{E}_c} \psi(\mathbf{x}_q^a, \mathbf{y}) d\mathbf{y} + \sum_{c=J_S^{poly}+1}^{J_S} \int_{\tilde{\mathcal{E}}_c} \psi(\mathbf{x}_q^a, \mathbf{y}) d\mathbf{y} \\
&\approx \sum_{c=1}^{J_S^{poly}} \sum_{p=1}^{Q_{in}} \psi(\mathbf{x}_q^a, \mathbf{y}_p^c) w_q^a w_p^c + \sum_{c=J_S^{poly}+1}^{J_S} \sum_{p=1}^{Q_{in}^{nonpoly}} \psi(\mathbf{x}_q^a, \mathbf{y}_p^c) w_q^a w_p^c.
\end{aligned} \tag{4.26}$$

4.3.3 Fully discretized stiffness matrix entry

In order to ease the identification of the considerable cases for $(\mathcal{E}_a, \mathcal{E}_b)$ discussed in the latter two subsections we introduce the set

$$\mathring{S}(\mathcal{E}) := \{b \in S(\mathcal{E}) : \quad \mathcal{E}_b \cap S(\mathbf{x}) = \mathcal{E}_b \quad \text{for all } \mathbf{x} \in \mathcal{E}\} \subset S(\mathcal{E}).$$

If $b \in \mathring{S}(\mathcal{E}_a)$, then neither a special treatment of the outer integral (note that $\mathcal{E}_a^1 = \mathcal{E}_a$) nor a subdivision of the intersection region (note that $\mathcal{E}_b \cap S(\mathbf{x}_q^a) = \mathcal{E}_b$) needs to be considered. By further defining $\tilde{S}(\mathcal{E}) := \left(S(\mathcal{E}) \backslash \left(\mathring{S}(\mathcal{E}) \cup \partial S(\mathcal{E})\right)\right)$ we obtain the partition

$$S(\mathcal{E}) = \mathring{S}(\mathcal{E}) \,\dot{\cup}\, \tilde{S}(\mathcal{E}) \,\dot{\cup}\, \partial S(\mathcal{E}), \tag{4.27}$$

which is pictured in Figure 4.9. This partition corresponds to the following three different cases for $(\mathcal{E}_a, \mathcal{E}_b)$ each of which gets a different numerical treatment. Particularly, let $1 \leq a \leq J_\Omega$ be a fixed element index, then

(a) if $b \in \mathring{S}(\mathcal{E}_a)$ we can employ standard element quadrature rules on both, $\mathcal{E}_a$ and $\mathcal{E}_b$,

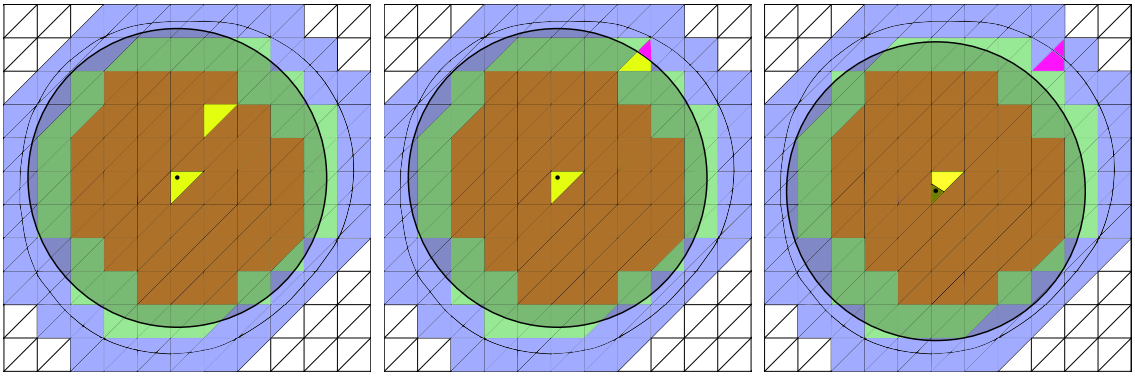
(b) if $b \in \tilde{S}(\mathcal{E})$ then there are $\mathbf{x} \in \mathcal{E}_a$ for which $\mathcal{E}_b \cap S(\mathbf{x}) \subsetneqq \mathcal{E}_b$ such that a subdivision of the intersection region may become necessary,

(c) and if $b \in \partial S(\mathcal{E}_a)$ we have to subdivide the intersection region $\mathcal{E}_b \cap S(\mathbf{x})$ and we have to take care of the partition $\mathcal{E}_a = \mathcal{E}_a^0 \dot{\cup} \mathcal{E}_a^1$.

In Figure 4.10 these cases are illustrated. Finally, invoking (4.20), (4.23) and (4.26) we arrive at the fully discretized matrix entry

$$
\begin{aligned}
a_{kj} =& \sum_{a=1}^{J_\Omega} \sum_{b \in \mathring{S}(\mathcal{E}_a)} \int_{\mathcal{E}_a} \varphi_j \int_{\mathcal{E}_b} (\varphi_k \phi - \varphi_k{}' \phi') d\mathbf{y} d\mathbf{x} \\
&+ \sum_{a=1}^{J_\Omega} \sum_{b \in \tilde{S}(\mathcal{E})} \int_{\mathcal{E}_a} \varphi_j \int_{\mathcal{E}_b \cap S(\mathbf{x})} (\varphi_k \phi - \varphi_k{}' \phi') d\mathbf{y} d\mathbf{x} \\
&+ \sum_{a=1}^{J_\Omega} \sum_{b \in \partial S(\mathcal{E}_a)} \int_{\mathcal{E}_a} \varphi_j \int_{\mathcal{E}_b \cap S(\mathbf{x})} (\varphi_k \phi - \varphi_k{}' \phi') d\mathbf{y} d\mathbf{x} \\
&\approx \sum_{a=1}^{J_\Omega} \sum_{b \in \mathring{S}(\mathcal{E}_a)} \sum_{q=1}^{Q_{out}} \sum_{p=1}^{Q_{in}} \psi_{kj}(\mathbf{x}_q^a, \mathbf{y}_p^b) w_q^a w_p^b \\
&+ \sum_{a=1}^{J_\Omega} \sum_{b \in \tilde{S}(\mathcal{E})} \left(\sum_{q=1}^{Q_{out}} \left(\sum_{c=1}^{J_S^{poly}} \sum_{p=1}^{Q_{in}} \psi_{kj}(\mathbf{x}_q^a, \mathbf{y}_p^c) w_q^a w_p^c \right.\right. \\
&\qquad \left.\left. + \sum_{c=1+J_S^{poly}}^{J_S} \sum_{p=1}^{Q_{in}^{nonpoly}} \psi_{kj}(\mathbf{x}_q^a, \mathbf{y}_p^c) w_q^a w_p^c \right) \right) \\
&+ \sum_{a=1}^{J_\Omega} \sum_{b \in \partial S(\mathcal{E}_a)} \left(\sum_{q=1}^{Q_{out}^{special}} \left(\sum_{c=1}^{J_S^{poly}} \sum_{p=1}^{Q_{in}} \psi_{kj}(\mathbf{x}_q^a, \mathbf{y}_p^c) w_q^a w_p^c \right.\right. \\
&\qquad \left.\left. + \sum_{c=1+J_S^{poly}}^{J_S} \sum_{p=1}^{Q_{in}^{nonpoly}} \psi_{kj}(\mathbf{x}_q^a, \mathbf{y}_p^c) w_q^a w_p^c \right) \right).
\end{aligned}
\tag{4.28}
$$

Remark 4.3.2 (***Quadrature rules***)**.** *We implement two (possibly different) quadrature rules for the reference element, say* $\{\hat{\mathbf{x}}, \hat{w}_q\}_{q=1}^{Q_{out}}$ *and* $\{\hat{\mathbf{y}}, \hat{w}_p\}_{p=1}^{Q_{in}}$*. The first rule* $\{\hat{\mathbf{x}}, \hat{w}_q\}_{q=1}^{Q_{out}}$ *is used for the integration of the outer integral, whereas* $\{\hat{\mathbf{y}}, \hat{w}_p\}_{p=1}^{Q_{in}}$ *is used for the elements of the polyhedral subdivision which appears along the integration of the inner integral. For the remaining nonpolyhedral segments of the inner integral we may need to implement more specialized quadrature rules (e.g., quadrature rules for circular segments; see [46, 45]) or consider polyhedral approximations thereof.*



(a) (b) (c)

Figure 4.10: The three cases for $\mathcal{E}_b$ in (4.28) for a fixed element $\mathcal{E}_a$. In all of the three images, $\mathcal{E}_a$ of the outer integral is represented by the yellow triangle in the center of each image containing a black dot. This dot indicates a point $\mathbf{x} \in \mathcal{E}_a$ and is the center of the shady interaction set $S(\mathbf{x})$, here an Euclidean ball. Image (a) corresponds to the simplest case $\mathcal{E}_b \in \mathring{S}(\mathcal{E}_a)$, where we integrate over whole triangles (yellow triangles). In (b), the case $\mathcal{E}_b \in \tilde{S}(\mathcal{E})$ is depicted. Here, we only integrate over the yellow colored part of the magenta triangle $\mathcal{E}_b$. Finally, in (c) we illustrate the case $\mathcal{E}_b \in \partial S(\mathcal{E}_a)$. More precisely, we consider a point $\mathbf{x} \in \mathcal{E}_a^0$, so that the intersection $S(\mathbf{x}) \cap \mathcal{E}_b$ is empty.

In our 2d implementations with P_1 elements, we use a symmetric Gaussian quadrature rule for triangles [41]. The two rules $\{\hat{\mathbf{x}}, \hat{w}_q\}_{q=1}^{Q_{out}}$ and $\{\hat{\mathbf{y}}, \hat{w}_p\}_{p=1}^{Q_{in}}$ then only differ in the quadrature order. As pointed out in Subsection 4.3.1 it is preferable to employ higher order quadrature rules for the outer integral. For the cases $1 \leq a \leq J_\Omega$ and $b \in \partial S(\mathcal{E}_a)$ we even use an adaptive quadrature rule for the outer integral over $\mathcal{E}_a$; see also Remark 4.3.1. Specifically, we employ Ronald Cools and Ann Haegemans 5-7 embedded rule [26] to save function evaluations.
For our finite element implementations with Q_1 elements, we "tensorize" 1d Gaussian quadrature rules to obtain quadrature rules on the d-dimensional elements. If a singular kernel is involved as in the study cases of Chapter 5, we use adaptive quadrature rules. In this case we "tensorize" the embedded 1d Gauss-Kronrod rule to obtain embedded adaptive quadrature rules on d-dimensional Q_1 elements; see also the Appendices A.2.1 and A.2.2.
Throughout the experiments conducted in this thesis the quadrature rules and especially the order of the quadrature rules have been chosen rather heuristically in such a way that expected convergence rate are achieved.

4.3.4 Challenges

After the detailed discussion on the assembly of the nonlocal stiffness matrix, we now want to illuminate and summarize some obvious and some less obvious difficulties related to this task. We figure out why finite element implementations of nonlocal equations

contrast strongly with standard finite element discretizations of analogue local problems. For this purpose let us first establish a comparison context and briefly recall a standard finite element implementation of the local homogeneous Poisson equation

$$\begin{cases} -\Delta u(\mathbf{x}) = f(\mathbf{x}) & \mathbf{x} \in \Omega \\ u(\mathbf{x}) = 0 & \mathbf{x} \in \partial\Omega. \end{cases} \tag{4.29}$$

The associated bilinear form is given by

$$A^{loc}(u, v) := \int_{\Omega} \nabla u^T \nabla v d\mathbf{x}.$$

We invoke the triangulation $\mathcal{T}_{\Omega}^{h}$ and define the entries of the local stiffness matrix as

$$a_{kj}^{loc} = A^{loc}(\varphi_k, \varphi_j) = \sum_{a=1}^{J_\Omega} \int_{\mathcal{E}_a \cap (\mathcal{S}_k \cap \mathcal{S}_j)} \nabla {\varphi_k}^T \nabla \varphi_j d\mathbf{x}. \tag{4.30}$$

We now list the most dominant difficulties of finite element methods for nonlocal equations.

Densely populated stiffness matrices

For a pair of basis functions (φ_k, φ_j) with nonoverlapping supports, i.e. $\mathcal{S}_k \cap \mathcal{S}_j = \emptyset$, we find that the last double integral in (4.13) (or equivalently in (4.18)) may contribute nonzero values to the stiffness matrix and for both types of kernels the entry reduces to

$$a_{kj} = -\int_{\mathcal{S}_j} \varphi_j \int_{\mathcal{S}_k \cap S(\mathbf{x})} {\varphi_k}' \phi' d\mathbf{y} d\mathbf{x}. \tag{4.31}$$

In Figure 4.11 this situation is illustrated in 2d for a triangular mesh. This stands in stark contrast to the stiffness matrix in the analogue local setting resulting from (4.30) where $a_{kj}^{loc} = 0$ if $\mathcal{S}_k \cap \mathcal{S}_j = \emptyset$. The discretization $\mathbf{A}^h$ of the nonlocal operator $-\mathcal{L}$ is therefore densely populated, which is the discrete analogue to nonlocality. The number of nonzero entries depends on the relations between the size of the interaction sets $\{S(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$, the size of the domain Ω and the grid size h. Specifically, if $\Omega = \Omega \cap S(\mathbf{x})$ for all $\mathbf{x} \in \Omega$, then $\mathbf{A}_{\Omega\Omega}^{h}$ is a dense matrix. See Figure 4.12 for an illustration thereof. Additional memory storage and appropriate solvers for the resulting linear system (4.9) need to be considered. In Chapter 5 we show how to circumvent these problems in the special case of regular grids and translation invariant kernels.

Singular kernels

Another challenge occurs if singular kernels are involved (see Figure 4.13 for an example). For instance, various fractional derivative models and the peridynamics model for solid mechanics are important applications where we have to cope with this type of kernels. In this case the necessity of sophisticated numerical quadrature rules arises. The implementation itself becomes more involved and results in additional computational costs

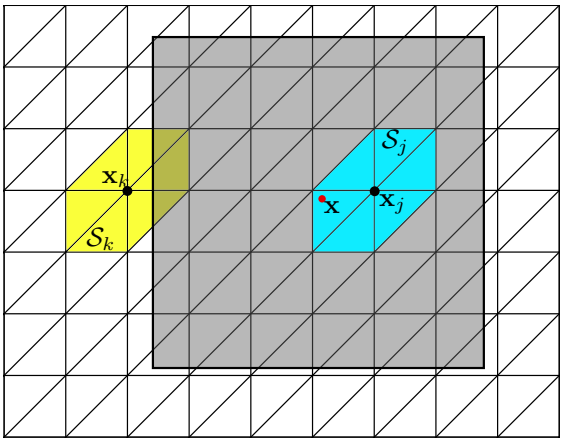


Figure 4.11: Nonzero contributions to the stiffness matrix despite nonoverlapping supports. We see the yellow colored support $\mathcal{S}_k$ of the basis function φ_k centered at $\mathbf{x}_k$ and the cyan colored support $\mathcal{S}_j$ of the basis function φ_j centered at $\mathbf{x}_j$. There are points $\mathbf{x} \in \mathcal{S}_j$ (see, e.g., the red dot depicted in the image) for which the integration domain $\mathcal{S}_k \cap S(\mathbf{x})$ (here $S(\mathbf{x}) = B_{\delta,\infty}(\mathbf{x})$) of the inner integral in (4.31) is nonempty and therefore $a_{kj} \neq 0$.

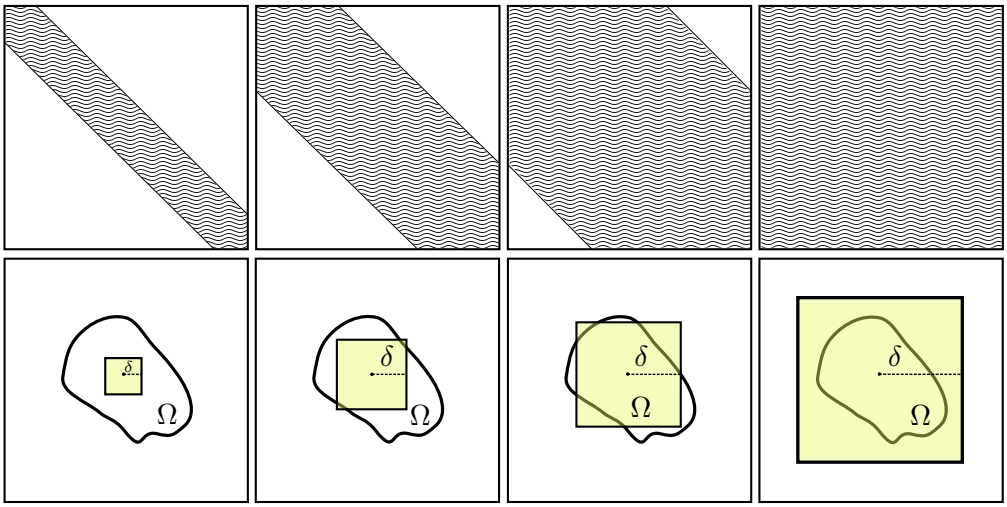


Figure 4.12: We obtain dense finite element stiffness matrices as the size of the interaction sets increases with respect to the (fixed) size of the domain Ω. We note that, in general and particularly in higher dimensions, the nonlocal stiffness matrix does not necessarily admit the structure of a banded matrix. In fact, the density pattern depends on the numbering of the finite element nodes.

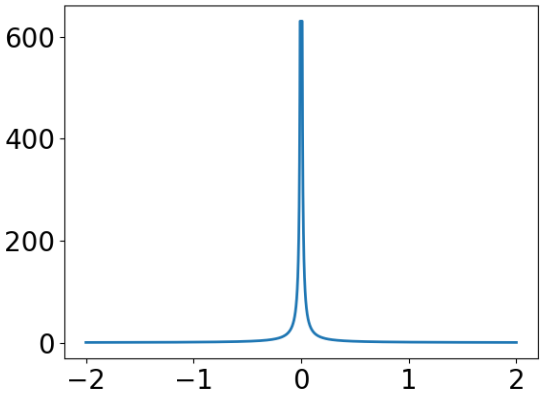


Figure 4.13: Challenging numerical integration due to the presence of singular kernels. Here, the fractional kernel (3.18) is plotted in 1d for a fraction $s = 0.4$ and $x = 0$. A carefully designed quadrature rule is needed to accurately resolve the singularity at the origin.

which can tremendously affect the assembly time. The simplest approach to handle singularities is to use adaptive quadrature rules, which avoid evaluations of the kernel at $(\mathbf{x}, \mathbf{x})$; see, e.g., [33] for a 1d implementation. For maintaining sufficient accuracy an appropriate refinement strategy is essential. However, such a refinement typically requires many function evaluations close to the singularity and therefore increases computation times. A second approach is to remove the singularities by hand. Such an approach is presented in [1] for the two-dimensional fractional Poisson equation; see (6.21). Therein, the authors apply the Duffy transformation [40] to the integrals over each finite element, such that the singularity is canceled out by the determinant term resulting from the transformation formula. This manual preprocessing may become very tedious and, in contrast to a naive adaptive quadrature approach, has to be performed for each kernel individually. However, the resulting nonsingular integrals are numerically tractable for standard quadrature rules such as Gaussian rules.

Infinite interactions

In case of unbounded interaction sets, i.e., $S(\mathbf{x}) = \mathbb{R}^d$, the subdivision task discussed for the inner integral in Subsection 4.3.2 is obsolete, but we have to cope with unbounded integration domains, which leads to fully populated stiffness matrices. We either have to appropriately truncate these domains in the occurring integrals while preventing a significant loss of accuracy, or we treat special cases in which such integrals can be computed by hand (see related discussions in Section 6.2), or transformed into a (local) boundary integral (see, e.g., [5, Section 2]). This situation arises for the fractional diffusion problem (6.21) and further details are provided in Section 6.2.

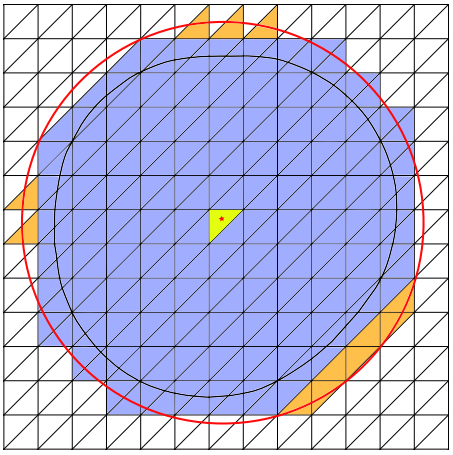


Figure 4.14: Illustration of $S^{approx}(\mathcal{E}_a)$ defined in (4.32). The fixed element $\mathcal{E}_a$ of the outer assembly loop is given by the yellow colored triangle with associated set $S(\mathcal{E}_a)$ that consists of all blueish colored triangles. The red line indicates the boundary of the Euclidean ball of radius $\delta + h$ centered at the barycenter of $\mathcal{E}_a$ given by the red star. The orange colored triangles indicate all additional triangles in $S^{approx}(\mathcal{E}_a) \backslash S(\mathcal{E}_a)$.

Double loop assembly

Considering representation (4.20) it is not sufficient to touch each element only once as it is the case for local problems such as (4.30). Clearly, this unfolds once again the nonlocal nature of the problem causing that not only basis functions with overlapping support contribute nonzero values to the stiffness matrix. More precisely, for each triangle $\mathcal{E}_a \in \mathcal{T}_\Omega^h$ we have to take into account all triangles $\mathcal{E}_b$ with $b \in S(\mathcal{E}_a)$, where $S(\mathcal{E}_a)$ is defined in (4.19). However, for interaction sets $S(\mathbf{x})$ which are "smaller" than the domain Ω, we have that $S(\mathcal{E}_a)$ may be significantly smaller than $\mathcal{T}^h$. We can exploit this fact in a mesh-preprocessing step by computing for every $\mathcal{E}_a \in \mathcal{T}_\Omega^h$ an approximation $S^{approx}(\mathcal{E}_a) \supset S(\mathcal{E}_a)$ and store the corresponding indices into a list. For example, for norm induced interaction sets $S(\mathbf{x}) := B_{\delta,\bullet}(\mathbf{x})$ we could compute $S^{approx}(\mathcal{E}_a)$ as follows

$$S^{approx}(\mathcal{E}_a) := \left\{ b \in \{1, \dots, J\} : \|\mathbf{x}_a^{bary} - \mathbf{x}_b^{bary}\|_\bullet < \delta + h \right\}, \tag{4.32}$$

where $\mathbf{x}_a^{bary}$ denotes the barycenter of $\mathcal{E}_a$, and similarly for b, and h denotes the diameter of the largest element. Clearly, for fine meshes such a list exceeds the memory capacities and rather has to be computed on the fly. In Figure 4.14 this criterion is illustrated.

Quadrature of the outer integral

As alluded to in Subsection 4.3.1 for pairs of elements $(\mathcal{E}_a, \mathcal{E}_b)$ with $b \in \partial S(\mathcal{E}_a)$ we need carefully designed quadrature rules for the outer integral over $\mathcal{E}_a$. This also adds to the costs created along the assembly process. In this regard we want to illuminate a particular situation where the negligence of these cases leads to highly inaccurate

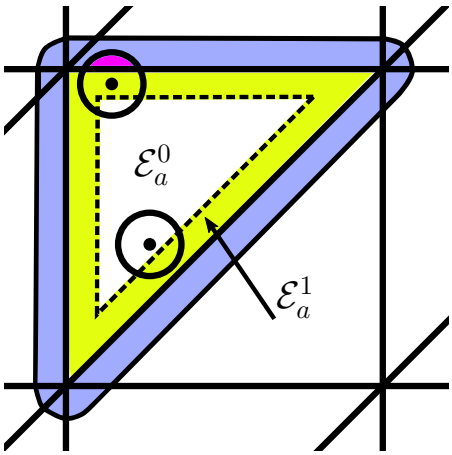


Figure 4.15: Illustration of the critical case in which the interaction horizon is smaller than the grid size. For $\delta \ll h$ we find $\mathcal{E}_a^0 \approx \mathcal{E}_a$ so that the integrand $\mathcal{I}_b(\mathbf{x})$ of the outer integral (4.22) is zero in most parts of $\mathcal{E}_a$ and quadrature rules have to be carefully chosen.

results. Let us consider the case of norm induced interaction sets $S(\mathbf{x}) := B_{\delta,\bullet}(\mathbf{x})$ with an interaction horizon δ *smaller* than the grid size h (see Figure 4.15 for an illustration). In this case the matrix entry (4.28) reduces to

$$a_{kj} = \sum_{a=1}^{J_\Omega} \sum_{b \in \partial B_{\delta,\bullet}(\mathcal{E}_a)} \int_{\mathcal{E}_a} \varphi_j \int_{\mathcal{E}_b \cap B_{\delta,\bullet}(\mathbf{x})} (\varphi_k \phi - \varphi_k{}' \phi') d\mathbf{y} d\mathbf{x}.$$

This situation appears, e.g., during the numerical investigation of the local limit $\delta \to 0$; see Chapter 7. Thus we exclusively have to deal with partially covered triangles and standard quadrature rules for the outer integral fail to accurately approximate the stiffness matrix entries. An exception for the latter statement arises in the trivial case where each ball $B_{\delta,\bullet}(\mathbf{x}_q^a)$ centered at the quadrature point $\mathbf{x}_q^a$ of the outer integral is fully contained in $\mathcal{E}_a$. In this case, each triangle only interacts with itself, meaning that we have a local model.

Quadrature of the inner integral

In Subsection 4.3.2 we explain how we deal with the inner integral

$$\mathcal{I}_b(\mathbf{x}_q^a) = \int_{\mathcal{E}_b \cap S(\mathbf{x}_q^a)} \psi(\mathbf{x}_q^a, \mathbf{y}) d\mathbf{y}.$$

The fundamental difficulty is to numerically integrate over the intersection $\mathcal{E}_b \cap S(\mathbf{x}_q^a)$ between a finite element $\mathcal{E}_b$ and the interaction set $S(\mathbf{x}_q^a)$ for $\mathbf{x}_q^a$ being a quadrature point in $\mathcal{E}_a$. In (4.25) we consider a re-triangulation of this intersection region into subelements, where we distinguish between a polyhedral part and the (possibly empty)

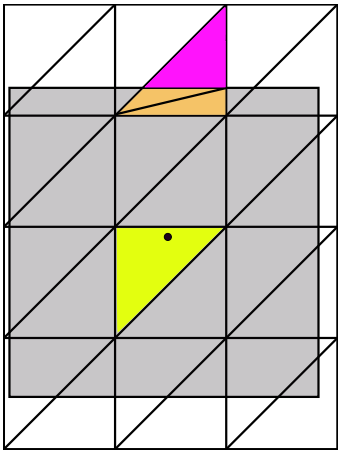
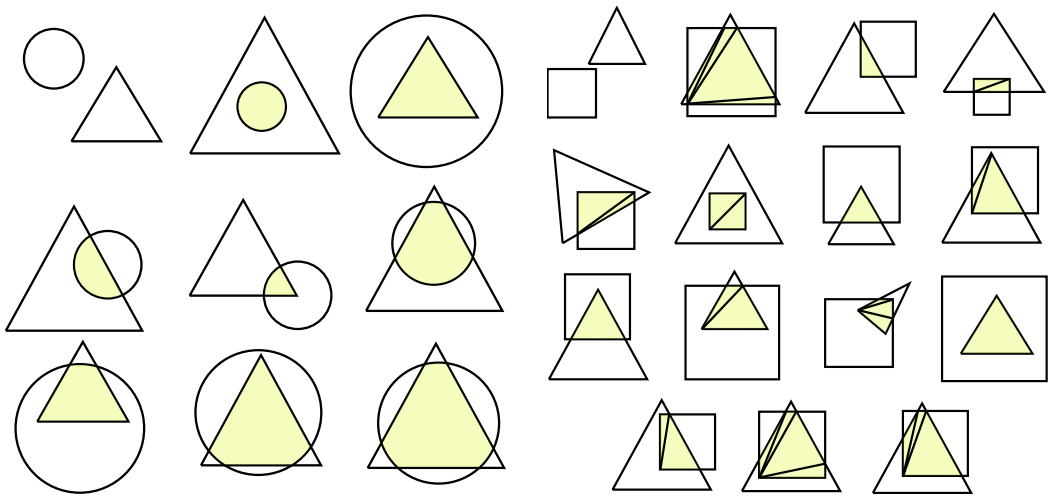


Figure 4.16: Simplified subdivision of a partially covered triangle due to the use of polygonal interaction sets. As in Figure 4.8, the fixed triangle $\mathcal{E}_a$ of the outer integral is depicted in yellow. The $\|\cdot\|_\infty$-ball centered at a point $\mathbf{x} \in \mathcal{E}_a$ (black dot) only partially covers the magenta colored triangle $\mathcal{E}_b$. However, the intersection region can *accurately* be subdivided into an orange polygonal part consisting of $J_S^{poly} = 2$ triangles.

nonpolyhedral part. Thus, we need to implement a routine, which for a given pair $(\mathcal{E}, S(\mathbf{x}))$ first *identifies the intersection region* $\mathcal{E} \cap S(\mathbf{x})$ (note that $\mathcal{E} \cap S(\mathbf{x}) = \emptyset$ and $\mathcal{E} \cap S(\mathbf{x}) = \mathcal{E}$ are possible) and then *outputs a subdivision* into appropriate subelements taking into account the delimitation of a polyhedral part as in (4.25). The necessity for such a routine does tremendously affect the efficiency of finite element implementations for nonlocal models with truncated kernels and is a key factor contributing to the high computational costs compared to those of analogue local finite element discretizations. For the sake of illustration let us now consider the most standard setting which is ubiquitous in the related literature; see, e.g., [36, 39, 39, 33]. We consider nonlocal models in $\mathbb{R}^2$ with kernels which are truncated by the standard interaction set $S(\mathbf{x}) = B_{\delta,2}(\mathbf{x})$ and we employ P_1 elements. The left image in Figure 4.17 shows the eight possible intersection scenarios occurring in this setting. We note that each of these intersection regions can be subdivided into a polygonal region and at most three circular segments. Thus, besides a quadrature rule for triangles, we need to implement a routine for integrating over general circular segments. This approach is realized in [95].

In Chapter 3 we figure out, that the well-posedness theory for problem (2.1) is not limited to just Euclidean balls. Instead, we have extended the established results in the referenced literature to a large variety of interaction sets. In regard to the discussed intersection problem within a finite element discretization it is therefore worth considering polyhedral interaction sets, such as balls induced by the supremum or Manhattan norm. In this case the intersection region $\mathcal{E}_b \cap S(\mathbf{x})$ is itself polyhedral and can accurately be covered by polyhedral subelements so that there is no need for integrating over complicated domains, such as circular segments (see Figure 4.16 and Figure 4.17 (b)). Also



(a) (b)

Figure 4.17: (a) The eight possible intersection scenarios for a disc and a triangle, as well as the nonintersecting case. (b) A selection of possible intersections between a square and a triangle with a subdivision of the polygonal intersection region (at most a hexagon).

note that the routine which accurately integrates over intersection regions resulting from polyhedral interaction sets is a subroutine of the routine employed for Euclidean balls.

In the three-dimensional case, the computation of integrals over regions resulting from intersecting finite elements, such as tetrahedra, and spherical interaction sets, is a highly ambitious task and out of the scope of this thesis. In contrast to that, imagine a three-dimensional regular grid composed of Q_1 (diced) elements and rectangular interaction sets such as $\|\cdot\|_\infty$-balls. The intersection between an element and a ball, would then be a simple cuboid. In combination with structure exploiting methods (see, e.g., Chapter 5) 3d implementations for truncated kernels thus approach technical feasibility.

4.4 Geometric approximations of interaction sets

We have in depth discussed the discretization of the nonlocal bilinear form which requires meticulously implemented numerical quadrature. In the view of Strang's first lemma, recalled below as Theorem 6.28, the overall finite element error consists of two parts. On the one hand we have the so-called *approximation error* (or *interpolation error*), which is the discretization error arising due to approximating the finite element solution $u \in V(\Omega \cup \Omega_I)$ by its projection $u^h \in V^h$ (see also (4.2)). On the other hand we have the so-called *consistency error* induced by approximating the analytic bilinear form (3.5) by means of numerical quadrature rules leading to a discretized bilinear form. Once the finite element space is chosen, such that the approximation error is fixed, one

aims to employ quadrature rules in such a way that the consistency error is smaller than the approximation error. With other words, the convergence rates of finite element solutions u^h to the analytic solution u as $h \to 0$ induced by the approximation quality of the ansatz spaces shall not be deteriorated by a poor numerical integration.
In Subsection 4.3.2 we derive a discretization of the inner integral of the nonlocal bilinear form. In the case of partially covered elements, we have to perform an intermediate step to partition the intersection region into subelements which are accessible to appropriate numerical quadrature rules. This step is very costly and may contribute a geometric component to the consistency error. In fact, it would be desirable to avoid this step at all by deliberately allowing for geometric errors which in the optimal case do not deteriorate the approximation error. For example, in the standard case of Euclidean balls in $2d$ we can think of neglecting the circular segments so that we were solely left with the polygonal part of the intersection region (see Figure 4.18). Even more brutal, in order to avoid the expensive subdivision task at all we could think of integrating not over the part $\mathcal{E}_b \cap S(\mathbf{x}_q^a)$ of the element which is covered but over the whole element $\mathcal{E}_b$. Such approaches lead to geometric approximations $S^h(\mathbf{x})$ of the interaction sets $S(\mathbf{x})$ and in the remainder we refer to such approximative approaches as *approximate interaction sets*. In Section 6.3 we propose different examples, invoke a result from Section 6.1 to analyze their effect on the finite element error and present some numerical tests.
Finally we want to remark that, besides the use of polyhedral interaction sets discussed at the end of the previous section, approximate interaction sets represent a second approach to ease finite element implementations for nonlocal models. However, the idea of approximate interaction sets can also be applied to polyhedral norm induced interaction sets; see Subsection 6.3.1.

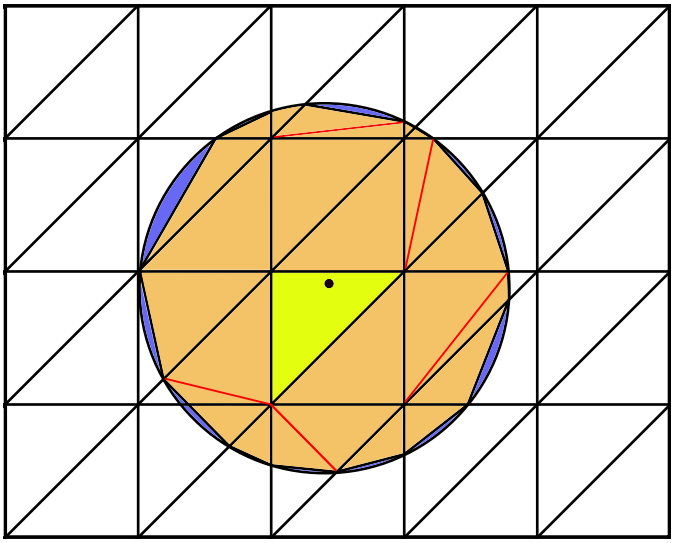


Figure 4.18: The orange colored domain represents one possible geometric approximation to the Euclidean ball centered at the black dot of the yellow triangle. This approximation results from neglecting the blue circular segments. The additional red lines indicate a subdivision of the partially covered elements of the mesh.

4.5 Concluding remarks and future work

We see great importance in the study of appropriate quadrature rules for the assembly of the nonlocal stiffness matrix involving truncated kernels. In order to reduce computational costs it is crucial to derive quantitative criteria which enable us to choose the lowest quadrature order possible such that the quadrature error does not deteriorate the approximation quality of the finite-dimensional subspaces V^h and thereby does not deteriorate convergence rates in h. To the best of our knowledge a thorough investigation of the latter for general truncated kernels and dimensions $d \geq 2$ cannot be found in literature. For the fractional kernel with infinite interactions related studies can be found, e.g., in [5, Appendix B].

Chapter 5

Exploiting multi-level Toeplitz structures on regular grids

As alluded to in Section 4.3, numerical implementations for nonlocal equations are challenging, especially in higher dimensions. In this chapter we consider a particular setting in which we can ease the assembly as well as the solving process in a structure exploiting fashion. As a consequence we are able to solve three-dimensional problems.

We study a finite element approximation of problem (2.1) with homogeneous volume constraints on an arbitrary d-dimensional hyperrectangle (parallel to the axis) for translation and reflection invariant kernels. More precisely, we analyze from a computational point of view a continuous Galerkin discretization with multilinear Q_1 elements for the following setting:

(A1) We set $\Omega := \prod_{i=0}^{d-1}(a_i, b_i)$, where (a_i, b_i) are open intervals on $\mathbb{R}$.

(A2) We assume that the kernel γ is *translation and reflection invariant*, such that

$$\gamma(\mathbf{b} + \mathbf{R}_i\mathbf{x}, \mathbf{b} + \mathbf{R}_i\mathbf{y}) = \gamma(\mathbf{x}, \mathbf{y})$$

for all $\mathbf{b} \in \mathbb{R}^d$ and all $0 \leq i \leq d$, where $\mathbf{R}_i(\mathbf{x}) := (x_0, \ldots, -x_i, \ldots, x_{d-1})$ and $R_d := Id$.

The assumptions made in (A2) imply that the kernel is symmetric. Therefore, we are considering nonlocal diffusion in this chapter. In fact, due to the translation invariance of the kernel, we find for $\mathbf{x}, \mathbf{y} \in \mathbb{R}$, that $\gamma(\mathbf{x}, \mathbf{y}) = \gamma(\mathbf{0}, \mathbf{y}-\mathbf{x})$ and $\gamma(\mathbf{y}, \mathbf{x}) = \gamma(\mathbf{0}, \mathbf{x}-\mathbf{y}) = \gamma(\mathbf{0}, -(\mathbf{y}-\mathbf{x}))$. Due to the reflection invariance we find that $\gamma(\mathbf{x}, \mathbf{y}) = \gamma(-\mathbf{x}, -\mathbf{y})$, which finally implies $\gamma(\mathbf{x}, \mathbf{y}) = \gamma(\mathbf{0}, \mathbf{y} - \mathbf{x}) = \gamma(\mathbf{0}, -(\mathbf{y} - \mathbf{x})) = \gamma(\mathbf{y}, \mathbf{x})$.
As a consequence, these structural assumptions on the underlying problem are reflected in the stiffness matrix. We obtain a symmetric d-level Toeplitz matrix, which has two crucial advantages. On the one hand, we only need to assemble (and store) the first row (or column) of the stiffness matrix. On the other hand, we can benefit from an efficient implementation of the matrix-vector product for solving the linear system. This result is presented in Theorem 5.1.2 and is crucial for this chapter, since it finally enables us

to solve the discretized system in an affordable way. For illustrative purposes we choose the fractional kernel (3.18).

This chapter is organized as follows. In Section 5.1 we give details about the precise finite element setting and proof our main result, that the stiffness matrix is multilevel Toeplitz. In Section 5.2 we explain the implementation of the solving procedure for the resulting structured system. Finally, in Section 5.3 we round off these considerations by presenting numerical results with application to space-fractional diffusion.

5.1 Finite element setting

In this section we study a continuous Galerkin discretization of the homogeneous nonlocal Dirichlet problem (2.1) under the assumptions (A1) and (A2) given above. Assumption (A1) allows for a simple triangulation of Ω, which we use to define a finite-dimensional energy space V_c^h. Together with (A2) we can show that this discretization yields the multilevel Toeplitz structure of the stiffness matrix, where the order of the matrix is determined by the number of grid points in each respective space dimension.

5.1.1 Definition of the finite-dimensional energy space

We rigorously construct Q_1 elements for our setting and point out some properties which are necessary to prove our main result in the next subsection.
For this purpose we decompose the domain

$$\Omega = \prod_{i=0}^{d-1} (a_i, b_i)$$

into d-dimensional hypercubes with sides of length $h > 0$ in each respective dimension. Note that we can omit a discretization of Ω_I since we assume homogeneous Dirichlet volume constraints; see Remark 4.2.2. Let

$$\mathbf{N} = (N_i)_{0 \le i < d} := \left(\frac{b_i - a_i}{h} \right)_{0 \le i < d}$$

and

$$\mathbf{L} = (L_i)_{0 \le i < d} := (N_i - 1)_{0 \le i < d},$$

then for the interior of Ω this procedure results in

$$\mathbf{L}^d := \prod_{i=0}^{d-1} L_i$$

degrees of freedom. Due to the simple structure of the domain we can choose a canonical numeration for the resulting grid

$$\prod_{i=0}^{d-1} (a_i + h \{0, \dots, L_i - 1\})$$

of interior points. More precisely, we employ the map

$$E^{\mathbf{n}}\colon \mathbb{N}^d \to \mathbb{N}, \quad E^{\mathbf{n}}(\mathbf{z}) := \sum_{i=0}^{d-1} z_i p_i(\mathbf{n}), \tag{5.1}$$

where

$$p_i(\mathbf{n}) := \prod_{j>i} n_j,$$

for establishing an order on a structured grid

$$\prod_{i=0}^{d-1} \{0, \dots n_i - 1\}, \quad \text{where} \quad \mathbf{n} = (n_0, \dots, n_{d-1}) \in \mathbb{N}^d.$$

Its inverse is given by

$$E^{-\mathbf{n}}\colon \mathbb{N} \to \mathbb{N}^d, \quad E^{-\mathbf{n}}(k) = \left(\left\lfloor \frac{k}{p_i(\mathbf{n})} \right\rfloor - \left\lfloor \frac{k}{p_{i-1(\mathbf{n})}} \right\rfloor n_i \right)_{0 \le i < d},$$

where $\lfloor \cdot \rfloor$ denotes the floor function given by

$$\lfloor m \rfloor := \max \{ z \in \mathbb{Z} : z \le m \}.$$

Let

$$\mathbf{e} := (1, \dots, 1) \in \mathbb{R}^d \quad \text{and} \quad \mathbf{a} := (a_0, \dots, a_{d-1}),$$

then we define the ordered array of interior grid points

$$(\mathbf{x}_k)_{0 \le k < \mathbf{L}^d} \in \mathbb{R}^{\mathbf{L}^d \times d}, \quad \text{where} \quad \mathbf{x}_k := \mathbf{a} + h(E^{-\mathbf{L}}(k) + \mathbf{e}). \tag{5.2}$$

In the view of (4.3) we have $K_\Omega = \mathbf{L}^d$. We further define physical elements

$$\mathcal{E}_j := \mathbf{b}_j + h\square, \quad \text{for} \quad 0 \le j < \mathbf{N}^d =: J_\Omega$$

where

$$\mathbf{b}_j = \left(b_k^i \right)_{0 \le i < d} := \mathbf{a} + hE^{-\mathbf{N}}(j) \quad \text{and} \quad \square := [0,1]^d,$$

so that we obtain a triangulation $\mathcal{T}_\Omega^h = \{\mathcal{E}_j\}_{j=0}^{J_\Omega - 1}$ of Ω with $\overline{\Omega} = \bigcup_{j=0}^{J_\Omega - 1} \mathcal{E}_j$. Next we aim to define appropriate element basis functions on the reference element $\square = [0,1]^d$. Therefore we denote by

$$(\mathbf{v}_k)_{0 \le k < 2^d} \in \mathbb{R}^{2^d \times d}$$

the vertices of the unit cube $\square$ ordered according to

$$\mathbf{v}_k := E^{-2\mathbf{e}}(k).$$

Then for each vertex $\mathbf{v}_k$, $0 \le k < 2^d$, we define an *element basis function* $\hat{\psi}_k\colon \square \to [0,1]$ by

$$\hat{\psi}_k(\mathbf{x}) = \left(\prod_{i=0, v_k^i = 0}^{d-1} (1 - x_i) \right) \left(\prod_{i=0, v_k^i = 1}^{d-1} x_i \right),$$

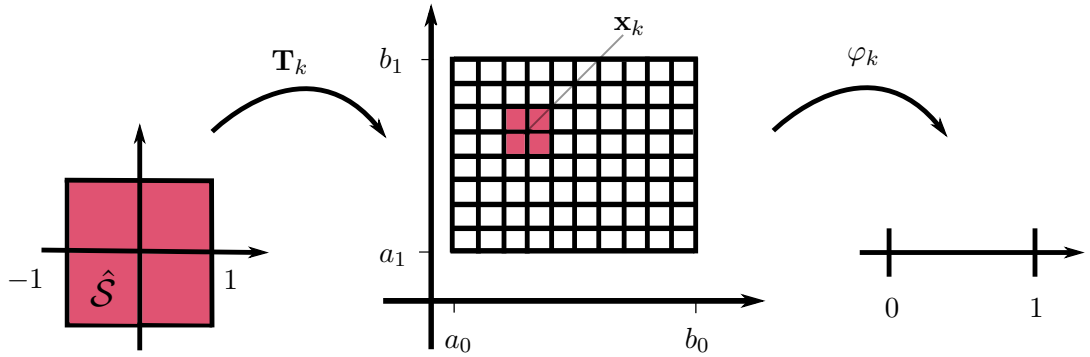


Figure 5.1: Transformation of the reference basis function, where $d = 2$.

where v_k^i denotes the i-th entry of the vector $\mathbf{v}_k = \left(v_k^0, \ldots, v_k^{d-1}\right)$. For dimensions $d \in \{1, 2, 3\}$ respectively, these are the usual linear, bilinear and trilinear element basis functions (see, e.g., [49, Chapter 1]). They are defined in such a way that $0 \leq \hat{\psi}_k \leq 1$ and $\hat{\psi}_k(\mathbf{v}_j) = \delta_{kj}$. Moreover, we define the *reference basis function* $\hat{\varphi}\colon \mathbb{R}^d \to [0, 1]$ by

$$\hat{\varphi}(\mathbf{x}) := \begin{cases} \hat{\psi}_i(\mathbf{v}_i + \mathbf{x}) & : \mathbf{x} \in (\square - \mathbf{v}_i) \\ 0 & : else. \end{cases} \tag{5.3}$$

We note that

$$\hat{\mathcal{S}} := [-1, 1]^d = \dot{\bigcup}_{i=0}^{2^d} (\square - \mathbf{v}_i) \quad \text{(disjoint union)},$$

such that $\hat{\varphi}$ is well defined and $\operatorname{supp}(\hat{\varphi}) = \hat{\mathcal{S}}$. Now let the physical support be defined as

$$\mathcal{S}_k := \bigcup \left\{ \mathcal{E}_j \colon \mathbf{x}_k \in \mathcal{E}_j, 0 \leq j < \mathbf{N}^d \right\},$$

which is a patch of the elements touching the node $\mathbf{x}_k$. We associate to each element $\mathcal{E}_k$ the transformation

$$\mathbf{T}_k \colon \hat{\mathcal{S}} \to \mathcal{S}_k, \quad \mathbf{T}_k(\mathbf{v}) := \mathbf{x}_k + h\mathbf{v}.$$

We note that $\det d\mathbf{T}_k(\mathbf{x}) = h^d$. Then for each node $\mathbf{x}_k$ we define a basis function $\varphi_k \colon \Omega \cup \Omega_I \to [0, 1]$ by

$$\begin{aligned} \varphi_k(\mathbf{x}) &:= \begin{cases} \hat{\varphi}(\mathbf{T}_k^{-1}(\mathbf{x})) & : \mathbf{x} \in \mathcal{S}_k \\ 0 & : else \end{cases} \\ &= \begin{cases} \hat{\psi}_i(\mathbf{v}_i + \mathbf{T}_k^{-1}(\mathbf{x})) & : \mathbf{T}_k^{-1}(\mathbf{x}) \in (\square - \mathbf{v}_i) \\ 0 & : else, \end{cases} \end{aligned}$$

which satisfies $0 \leq \varphi_k \leq 1$ and $\varphi_k(\mathbf{x}_j) = \delta_{kj}$. Figure 5.1 illustrates the latter considerations for $d = 2$. Finally, we can define a constrained finite element space by

$$V_c^h := \operatorname{span} \left\{ \varphi_k \colon 0 \leq k < \mathbf{L}^d \right\}, \tag{5.4}$$

such that each linear combination consisting of a set of these basis functions fulfills the homogeneous Dirichlet volume constraints. Notice that we parametrize these spaces by the grid size h indicating the dimension $\mathbf{L}^d$, which is by definition a function of h. Finally, we close this subsection with the following observations, which we exploit in the remainder.

Remark 5.1.1. *Let $\mathbf{x} \in \mathbb{R}^d$, then:*

i) $\hat{\varphi}(\mathbf{x}) = \hat{\varphi}(|\mathbf{x}|)$, *where* $|\mathbf{x}| := (|x_i|)_i$.

ii) Let $\mathbf{R}_i \colon \mathbb{R}^d \to \mathbb{R}^d$, *for* $0 \leq i < d$, *denote the reflection*

$$\mathbf{R}_i(\mathbf{x}) = (x_0, \dots, -x_i, \dots, x_{d-1}),$$

then

$$\hat{\varphi}(\mathbf{x}) = \hat{\varphi}(|\mathbf{x}|) \Leftrightarrow \hat{\varphi}(\mathbf{x}) = \hat{\varphi}(\mathbf{R}_i(\mathbf{x})) \quad \text{for all } 0 \leq i < d. \tag{5.5}$$

iii) $\hat{\varphi}(\mathbf{x}) = \hat{\varphi}((x_{\sigma(i)})_i)$ *for all permutations* $\sigma \colon \{0, \dots, d-1\} \to \{0, \dots, d-1\}$.

Proof. We first show i). Since $\hat{\varphi}(\mathbf{x}) = 0 = \hat{\varphi}(|\mathbf{x}|)$ for $\mathbf{x} \in \operatorname{int}(\hat{\mathcal{S}})^c$, let $\mathbf{x} \in \operatorname{int}(\hat{\mathcal{S}})$. Thus, there exists an index $0 \leq k < d$ such that $\mathbf{x} \in \operatorname{int}(\square) - \mathbf{v}_k$, which implies that $x_i < 0$ if and only if $v_k^i = 1$. Hence, we can conclude that

$$\begin{aligned}
\hat{\varphi}(\mathbf{x}) = \hat{\psi}_k(\mathbf{x} + \mathbf{v}_k) &= \left(\prod_{i=0, v_k^i = 0}^{d-1} (1 - x_i) \right) \left(\prod_{j=0, v_k^i = 1}^{d-1} (1 + x_i) \right) \\
&= \left(\prod_{i=0, v_k^i = 0}^{d-1} (1 - |x_i|) \right) \left(\prod_{j=0, v_k^i = 1}^{d-1} (1 - |x_i|) \right) \\
&= \prod_{i=0}^{d-1} (1 - |x_i|) = \hat{\psi}_0(|\mathbf{x}| + v_0) = \hat{\varphi}(|\mathbf{x}|).
\end{aligned}$$

Then for ii), on the one hand, we have that $|\mathbf{x}| = |\mathbf{R}_i(\mathbf{x})|$ and therefore $\hat{\varphi}(\mathbf{x}) = \hat{\varphi}(\mathbf{R}_i(\mathbf{x}))$ for all $0 \leq i < d$. On the other hand, we note that the operation $|\cdot|$ is a composition of reflections $\mathbf{R}_i$, more precisely

$$|\mathbf{x}| = \left(\prod_{x_i < 0} \mathbf{R}_i \right) (\mathbf{x}).$$

Thus, we obtain the equivalence stated in ii). Statement iii) follows from the representation

$$\hat{\varphi}(\mathbf{x}) = \hat{\varphi}(|\mathbf{x}|) = \prod_{i=0}^{d-1} (1 - |x_i|)$$

due to the commutativity of the product. □

5.1.2 Multilevel Toeplitz structure of the stiffness matrix

Now we aim to show that the stiffness matrix $\mathbf{A}^h$, as defined in (4.8), owns the structure of a d-level Toeplitz matrix. This is decisive for this work, since it finally enables us to solve the discretized system (4.9) in an affordable way. Since we assume homogeneous Dirichlet data, i.e., $g = 0$, it suffices to assemble the part $\mathbf{A}^h_{\Omega\Omega}$; see Remark 4.2.2. For simplicity, we will refer to $\mathbf{A}^h = \mathbf{A}^h_{\Omega\Omega}$ as nonlocal stiffness matrix in the remainder.

From now on the assumption (A2) on the kernel function becomes crucial. At this point we note that the indicator function

$$(\mathbf{x}, \mathbf{y}) \mapsto \chi_{B_{\delta,\bullet}(\mathbf{x})}(\mathbf{y})$$

is translation and reflection invariant for any discrete norm $\|\cdot\|_\bullet$. Note that the symmetry requirement (S2) on general interaction sets is weaker than being reflection invariant as required in (A2), which is why we specify to norm induced balls here. Hence, as in (K2), we can regard the kernel as a composition

$$\gamma(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x}, \mathbf{y})\chi_{B_{\delta,\bullet}(\mathbf{x})}(\mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^d,$$

for some translation and reflection invariant function ϕ. In order to analyze the multilevel structure of $\mathbf{A}^h \in \mathbb{R}^{\mathbf{L}^d \times \mathbf{L}^d}$ it is convenient to introduce an appropriate multi-index notation. To this end, we choose $E^{\mathbf{L}}$ from (5.1) as index bijection and we identify

$$a_{\mathbf{i}\,\mathbf{j}} = a_{E^{\mathbf{L}}(\mathbf{i}),E^{\mathbf{L}}(\mathbf{j})}.$$

We call the matrix $\mathbf{A}^h$ *d-level Toeplitz* if

$$a_{\mathbf{i}\,\mathbf{j}} = a(\mathbf{i} - \mathbf{j}).$$

If even

$$a_{\mathbf{i}\,\mathbf{j}} = a(|\,\mathbf{i} - \mathbf{j}\,|),$$

where the absolute value is understood componentwise, i.e.,

$$|\,\mathbf{i}\,| := (|i_k|)_{0 \le k < d},$$

then each level is symmetric and we can reconstruct the whole matrix from the first row (or column). For a more general and detailed consideration of multilevel Toeplitz matrices see, e.g., [88]. However, with this notation at hand we can now formulate

Theorem 5.1.2. *Let the domain Ω and the kernel γ fulfill assumptions* (A1) *and* (A2)*, respectively, and let the finite element space V_c^h be defined as in (5.4) for a grid size $h > 0$. Then the stiffness matrix $\mathbf{A}^h$ is d-level Toeplitz, where each level is symmetric.*

Proof. The key point in the proof is the relation

$$a_{kj} = a\left(\left|h^{-1}(\mathbf{x}_k - \mathbf{x}_j)\right|\right),$$

which we show in two steps. First we show that

$$a_{kj} = a\left(h^{-1}(\mathbf{x}_k - \mathbf{x}_j)\right)$$

and then we proof $a(\mathbf{z}) = a(|\mathbf{z}|)$. Therefore let us recall that the entry a_{kj} of the stiffness matrix $\mathbf{A}^h$ in the homogeneous case can be written as (see (3.10))

$$a_{kj} = \frac{1}{2}\int_{\Omega\cup\Omega_I}\int_{\Omega\cup\Omega_I}(\varphi_k(\mathbf{y}) - \varphi_k(\mathbf{x}))(\varphi_j(\mathbf{y}) - \varphi_j(\mathbf{x}))\gamma(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}.$$

In (4.16) we have examined the support of the integrand, so that

$$a_{kj} = \frac{1}{2}\int_{(\mathcal{S}_k^c\times\mathcal{S}_k^c)^c\cap(\mathcal{S}_j^c\times\mathcal{S}_j^c)^c}(\varphi_k(\mathbf{y}) - \varphi_k(\mathbf{x}))(\varphi_j(\mathbf{y}) - \varphi_j(\mathbf{x}))\gamma(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}.$$

Aiming to show $a_{kj} = a\left(h^{-1}(\mathbf{x}_k - \mathbf{x}_j)\right)$ we need to carry out some basic transformations of this integral. Since by definition $\varphi_j = \hat{\varphi}\circ\mathbf{T}_j^{-1}$ and also $\det d\mathbf{T}_j(\mathbf{x}) = h^d$, we find

$$\begin{aligned}
a_{kj} &= \frac{1}{2}\int_{(\mathcal{S}_k^c\times\mathcal{S}_k^c)^c\cap(\mathcal{S}_j^c\times\mathcal{S}_j^c)^c}(\varphi_k(\mathbf{y}) - \varphi_k(\mathbf{x}))(\varphi_j(\mathbf{y}) - \varphi_j(\mathbf{x}))\gamma(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}\\
&= \frac{h^{2d}}{2}\int_{\mathbf{T}_j^{-1}(\mathbb{R}^d)\times\mathbf{T}_j^{-1}(\mathbb{R}^d)}\left(1-\chi_{\mathcal{S}_j^c\times\mathcal{S}_j^c}(\mathbf{T}_j(\mathbf{v}),\mathbf{T}_j(\mathbf{w}))\right)\left(1-\chi_{\mathcal{S}_k^c\times\mathcal{S}_k^c}(\mathbf{T}_j(\mathbf{v}),\mathbf{T}_j(\mathbf{w}))\right)\\
&\qquad\left((\hat{\varphi}\circ\mathbf{T}_k^{-1})(\mathbf{T}_j(\mathbf{w})) - (\hat{\varphi}\circ\mathbf{T}_k^{-1})(\mathbf{T}_j(\mathbf{v}))\right)(\hat{\varphi}(\mathbf{w}) - \hat{\varphi}(\mathbf{v}))\\
&\qquad\gamma\left(\mathbf{T}_j(\mathbf{v}),\mathbf{T}_j(\mathbf{w})\right)d\mathbf{w}d\mathbf{v}.
\end{aligned}$$

Now we make a collection of observations. Due to assumption (A2) we obtain

$$\gamma(\mathbf{T}_j(\mathbf{v}),\mathbf{T}_j(\mathbf{w})) = \gamma(\mathbf{x}_j + h\mathbf{v},\mathbf{x}_j + h\mathbf{w}) = \gamma(h\mathbf{v},h\mathbf{w}).$$

Furthermore, by definition of the transformations $\mathbf{T}_j,\mathbf{T}_k$ we find that $\mathbf{T}_j^{-1}(\mathbb{R}^d) = \mathbb{R}^d$ as well as

$$(\mathbf{T}_k^{-1}\circ\mathbf{T}_j)(\mathbf{v}) = h^{-1}(\mathbf{x}_j + h\mathbf{v} - \mathbf{x}_k) = h^{-1}(\mathbf{x}_j - \mathbf{x}_k) + \mathbf{v}.$$

Since these transformations are bijective we also have that

$$\chi_{M^c\times M^c}(\mathbf{T}_j(\mathbf{x}),\mathbf{T}_j(\mathbf{y})) = \chi_{(\mathbf{T}_j^{-1}(M))^c\times(\mathbf{T}_j^{-1}(M))^c}(\mathbf{x},\mathbf{y})$$

for a set $M\subset\mathbb{R}^d$. Hence, defining

$$\mathbf{v}_{jk} := h^{-1}(\mathbf{x}_j - \mathbf{x}_k) = -\mathbf{v}_{kj} \tag{5.6}$$

and recognizing $\mathbf{T}_j^{-1}(\mathcal{S}_k) = \mathbf{v}_{kj} + \hat{\mathcal{S}}$ we finally obtain

$$a_{kj} = \frac{h^{2d}}{2}\int_{(\hat{\mathcal{S}}^c\times\hat{\mathcal{S}}^c)^c\cap((\mathbf{v}_{kj}+\hat{\mathcal{S}})^c\times(\mathbf{v}_{kj}+\hat{\mathcal{S}})^c)^c}$$

$$\begin{aligned}&\left(\hat{\varphi}(\mathbf{w}-\mathbf{v}_{kj})-\hat{\varphi}(\mathbf{v}-\mathbf{v}_{kj})\right)\left(\hat{\varphi}(\mathbf{w})-\hat{\varphi}(\mathbf{v})\right)\gamma(h\mathbf{w},h\mathbf{v})d\mathbf{w}d\mathbf{v}\\ =\ &a(\mathbf{v}_{kj}).\end{aligned}\tag{5.7}$$

Next, we proof that this functional relation fulfills $a(\mathbf{z}) = a(|\mathbf{z}|)$. Let us for this purpose define

$$F(\mathbf{x},\mathbf{y};\mathbf{z}) := (\hat{\varphi}(\mathbf{y}-\mathbf{z})-\hat{\varphi}(\mathbf{x}-\mathbf{z}))(\hat{\varphi}(\mathbf{y})-\hat{\varphi}(\mathbf{x}))\gamma(h\mathbf{y},h\mathbf{x})$$

such that

$$a(\mathbf{z}) = \frac{h^{2d}}{2}\int_{\left(\hat{\mathcal{S}}^c\times\hat{\mathcal{S}}^c\right)^c\cap\left((\mathbf{z}+\hat{\mathcal{S}})^c\times(\mathbf{z}+\hat{\mathcal{S}})^c\right)^c} F(\mathbf{x},\mathbf{y};\mathbf{z})d\mathbf{y}d\mathbf{x}.$$

Let $\mathbf{z}\in\left\{\mathbf{v}_{kj} : 0\le k,j<\mathbf{L}^d\right\}$. Then there exists a matrix $\mathbf{R}=\mathbf{R}(\mathbf{z})\in\mathbb{R}^{d\times d}$, which is a composition of reflections $\mathbf{R}_i$ from (5.5), such that $\mathbf{Rz}=|\mathbf{z}|$. Then from (5.5) and the assumption (A2) on the kernel, we obtain for $\mathbf{x},\mathbf{y}\in\mathbb{R}^d$ that

$$\begin{aligned}F(\mathbf{Rx},\mathbf{Ry};|\mathbf{z}|) &= \left(\hat{\varphi}(\mathbf{Ry}-\mathbf{Rz})-\hat{\varphi}(\mathbf{Rx}-\mathbf{Rz})\right)\left(\hat{\varphi}(\mathbf{Ry})-\hat{\varphi}(\mathbf{Rx})\right)\gamma(h\mathbf{Ry},h\mathbf{Rx})\\ &= \left(\hat{\varphi}(\mathbf{y}-\mathbf{z})-\hat{\varphi}(\mathbf{x}-\mathbf{z})\right)\left(\hat{\varphi}(\mathbf{y})-\hat{\varphi}(\mathbf{x})\right)\gamma(h\mathbf{y},h\mathbf{x})\\ &= F(\mathbf{x},\mathbf{y};\mathbf{z}).\end{aligned}$$

Since $\mathbf{R}(\hat{\mathcal{S}})=\hat{\mathcal{S}}$ and therefore

$$\mathbf{R}\left(\left(\hat{\mathcal{S}}^c\times\hat{\mathcal{S}}^c\right)^c\cap\left((\mathbf{z}+\hat{\mathcal{S}})^c\times(\mathbf{z}+\hat{\mathcal{S}})^c\right)^c\right) = \left(\hat{\mathcal{S}}^c\times\hat{\mathcal{S}}^c\right)^c\cap\left((|\mathbf{z}|+\hat{\mathcal{S}})^c\times(|\mathbf{z}|+\hat{\mathcal{S}})^c\right)^c,$$

we eventually obtain

$$\begin{aligned}a(|\mathbf{z}|) &= \frac{h^{2d}}{2}\int_{\left(\hat{\mathcal{S}}^c\times\hat{\mathcal{S}}^c\right)^c\cap\left((|\mathbf{z}|+\hat{\mathcal{S}})^c\times(|\mathbf{z}|+\hat{\mathcal{S}})^c\right)^c} F(\mathbf{x},\mathbf{y};|\mathbf{z}|)d\mathbf{y}d\mathbf{x}\\ &= \frac{h^{2d}}{2}\int_{\left(\hat{\mathcal{S}}^c\times\hat{\mathcal{S}}^c\right)^c\cap\left((\mathbf{z}+\hat{\mathcal{S}})^c\times(\mathbf{z}+\hat{\mathcal{S}})^c\right)^c} F(\mathbf{Rx},\mathbf{Ry};|\mathbf{z}|)d\mathbf{y}d\mathbf{x}\\ &= \frac{h^{2d}}{2}\int_{\left(\hat{\mathcal{S}}^c\times\hat{\mathcal{S}}^c\right)^c\cap\left((\mathbf{z}+\hat{\mathcal{S}})^c\times(\mathbf{z}+\hat{\mathcal{S}})^c\right)^c} F(\mathbf{x},\mathbf{y};\mathbf{z})d\mathbf{y}d\mathbf{x}\\ &= a(\mathbf{z}).\end{aligned}$$

Finally, we can show that $\mathbf{A}^h$ carries the structure of a d-level Toeplitz matrix. By having a closer look at the definitions of $E^{\mathbf{L}}$ and the grid points $\mathbf{x}_k$ (see (5.1) and (5.2)) we can conclude that

$$a_{\mathbf{i}\,\mathbf{j}} = a_{E^{\mathbf{L}}(\mathbf{i})E^{L}(\mathbf{j})} = a\left(\left|h^{-1}(\mathbf{x}_{E^{\mathbf{L}}(\mathbf{i})}-\mathbf{x}_{E^{\mathbf{L}}(\mathbf{j})})\right|\right) = a(|\,\mathbf{i}-\mathbf{j}\,|).$$

Thus, the entry $a_{\mathbf{i}\,\mathbf{j}}$ only depends on the difference $\mathbf{i}-\mathbf{j}$. □

With other words, the translation invariance of the kernel brings in the relation

$$a_{kj} = a\left(h^{-1}(\mathbf{x}_k-\mathbf{x}_j)\right).$$

The advantage of this observation relies on the usage of a regular grid leading to redundancy in the set

$$\left\{\mathbf{x}_k - \mathbf{x}_j \colon 0 \leq k, j < \mathbf{L}^d\right\}.$$

From the reflection invariance we can finally deduce

$$a_{kj} = a\left(\left|h^{-1}(\mathbf{x}_k - \mathbf{x}_j)\right|\right)$$

leading to symmetry in each level. We also point out that this result is independent of the relation between the interaction horizon δ and the diameter $\operatorname{diam}(\Omega)$ of the domain Ω.

As a consequence, in order to implement the matrix-vector product, it is sufficient to assemble solely the first row (or column)

$$\mathbf{R}^h := (a_{i0})_{0\leq i<\mathbf{L}^d} = \left(a\left(h^{-1}(\mathbf{x}_i - \mathbf{x}_0)\right)\right)_{0\leq i<\mathbf{L}^d} = \left(a\left(E^{\mathbf{L}}(i)\right)\right)_{0\leq i<\mathbf{L}^d}$$

of the stiffness matrix $\mathbf{A}^h$. Because with

$$i(k,j) := E^{\mathbf{L}}\left(h^{-1}(|\mathbf{x}_k - \mathbf{x}_j|)\right)$$

we obtain

$$a_{kj} = a\left(h^{-1}\left(|\mathbf{x}_k - \mathbf{x}_j|\right)\right) = a\left(E^{\mathbf{L}}\left(i(k,j)\right)\right) = \mathbf{R}^h_{i(k,j)}.$$

Note that $i(k,j)$ is well defined, since $h^{-1}\left(|\mathbf{x}_k - \mathbf{x}_j|\right)$ lies in the domain of definition of $E^{\mathbf{L}}$.

Remark 5.1.3. *Exploiting that $\hat{\varphi}$ is invariant under permutations, the same proof (by composing the reflection $\mathbf{R}$ with a permutation matrix) shows that $a(\mathbf{z}) = a((\mathbf{z}_{\sigma(i)})_i)$ for all permutations $\sigma\colon \{0,\dots,d-1\} \to \{0,\dots,d-1\}$. We use this observation to accelerate the assembling process. Also note in this regard, that a kernel of radial type, i.e., $\gamma(\mathbf{x},\mathbf{y}) = \phi(||\mathbf{x}-\mathbf{y}||_2)\chi_{B_{\delta,\bullet}(\mathbf{x})}(\mathbf{y})$, is also invariant under such permutations, independent of the norm $\|\cdot\|_\bullet$ used to define the ball $B_{\delta,\bullet}(\mathbf{x})$.*

5.2 Solving the discrete system

Now we discuss how to solve the densely populated multilevel Toeplitz system. The idea is to use an efficient implementation for the matrix-vector product of multilevel Toeplitz matrices, which is then delivered to the conjugate gradient (CG) method; see also [23].

Let us first illuminate the implementation of the matrix-vector product $\mathbf{Tx}$, for a general symmetric d-level Toeplitz matrix $\mathbf{T} \in \mathbb{R}^{\mathbf{L}^d \times \mathbf{L}^d}$ of order $\mathbf{L} = (L_0,\dots,L_{d-1})$ and $\mathbf{x}$ a vector in $\mathbb{R}^{\mathbf{L}^d}$. The crucial idea is to embed the Toeplitz matrix into a circulant matrix for which matrix-vector products can be efficiently computed with the help of the discrete Fourier transform (DFT) [97]. Here, by a d-level circulant matrix, we mean a matrix $\mathbf{C} \in \mathbb{R}^{\mathbf{L}^d \times \mathbf{L}^d}$, which satisfies

$$C_{\mathbf{ij}} = \mathbf{C}((\mathbf{i}-\mathbf{j}) \mod \mathbf{L}),$$

where

$$(\mathbf{i} \mod \mathbf{L}) := (i_k \mod L_k)_{0 \le k < d}.$$

In the real symmetric case, as it is present in our setting, $\mathbf{T}$ can be reconstructed from its first row $\mathbf{R} := (T_{0i})_i \in \mathbb{R}^{\mathbf{L}^d}$. Since circulant matrices are special Toeplitz matrices, the same holds for these matrices as well. Due to the multilevel structure it is convenient to represent $\mathbf{T}$ by a tensor

$$\mathbf{t} \in \mathbb{R}^{L_0 \times \cdots \times L_{d-1}},$$

which is composed of the values contained in $\mathbf{R}$. More precisely we define

$$\mathbf{t}(\mathbf{i}) := \mathbf{R}_{E^{\mathbf{L}}(\mathbf{i})}$$

for

$$\mathbf{i} \in \prod_{i=0}^{d-1} \{0, \dots, L_i - 1\}$$

with $E^{\mathbf{L}}$ from (5.1). Now $\mathbf{t}$ can be embedded into the tensor representation

$$\mathbf{c} \in \mathbb{R}^{2L_0 \times \cdots \times 2L_{d-1}}$$

of the associated d-level circulant matrix by

$$\mathbf{c}(\mathbf{i}) := \mathbf{t}(\hat{i}_0, \dots, \hat{i}_{d-1}),$$

where

$$\hat{i}_k := \begin{cases} i_k : & i_k < L_k, \\ 0 : & i_k = L_k, \\ 2L_k - i_k : & else. \end{cases}$$

We note that

$$\mathbf{t} = \mathbf{c}([0 : L_0 - 1], \dots, [0 : L_{d-1} - 1]).$$

Thus, we can use Algorithm 1 to compute the product $\mathbf{Tx}$, where the DFT is carried out by the fast Fourier transform (FFT). In the Python code we use the library pyFFTW (`https://hgomersall.github.io/pyFFTW/`) to perform a parallelized multidimensional DFT, which is a pythonic wrapper around the C subroutine library FFTW (`http://www.fftw.org/`). Furthermore, we want to note that a Message Passing Interface (MPI) implementation for solving multilevel Toeplitz systems in this fashion is presented in [23].
Finally, with this algorithm at hand, we employ the CG method, as it can be found, e.g., in [60], to obtain the solution of the symmetric discretized system (4.9).

Algorithm 1: Matrix-vector product for multilevel Toeplitz matrices

Input: $\mathbf{t} \in \mathbb{R}^{L_0 \times \cdots \times L_{d-1}}$ representing $\mathbf{T} \in \mathbb{R}^{\mathbf{L}^d \times \mathbf{L}^d}$ and $\mathbf{x} \in \mathbb{R}^{\mathbf{L}^d}$
Output: $\mathbf{y} = \mathbf{T}\mathbf{x}$

1 Construct $\mathbf{c} \in \mathbb{R}^{2L_0 \times \cdots \times 2L_{d-1}}$ by $\mathbf{c}(\mathbf{i}) := \mathbf{t}(\hat{i}_0, \ldots, \hat{i}_{d-1})$
2 Construct $\mathbf{x}' \in \mathbb{R}^{2L_0 \times \cdots \times 2L_{d-1}}$ by

$$\mathbf{x}'(\mathbf{i}) := \begin{cases} \mathbf{x}_{E^L(\mathbf{i})} : & \mathbf{i} \in \prod_{i=0}^{d-1} \{0, \ldots, L_i - 1\} \\ 0 : & else \end{cases}$$

3 Compute $\Lambda = FFT_{\mathbf{L}}(\mathbf{c})$
4 Compute $\mathbf{z} = FFT_{\mathbf{L}}(\mathbf{x}')$
5 Compute $\mathbf{w} = \Lambda \mathbf{z}$ (pointwise)
6 Compute $\mathbf{y}' = FFT_{\mathbf{L}}^{-1}(\mathbf{w})$
7 Construct $\mathbf{y} \in \mathbb{R}^{L_0 \times \cdots \times L_{d-1}}$ by $\mathbf{y}(\mathbf{i}) = \mathbf{y}'(\mathbf{i})$ for $\mathbf{i} \in \prod_{i=0}^{d-1} \{0, \ldots, L_i - 1\}$
8 Return $\mathbf{y}.reshape(\mathbf{L}^d)$

5.3 Numerical experiments

In this last section we want to complete the previous considerations by presenting numerical results in 1d, 2d and also in 3d. We now specify the nonlocal diffusion operator $-\mathcal{L}$ by choosing the truncated fractional Laplace kernel

$$\gamma(\mathbf{x}, \mathbf{y}) = \frac{c_{d,s}}{||\mathbf{y} - \mathbf{x}||^{d+2s}} \chi_{B_{\delta,\infty}(\mathbf{x})}(\mathbf{y}) \quad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^d,\ \mathbf{x} \neq \mathbf{y}), \tag{5.8}$$

where we employ $\|\cdot\|_\infty$-balls as interaction sets. In Chapter 6.2 we relate the nonlocal Dirichlet problem associated with this truncated kernel to the fractional Laplace operator $(-\Delta)^s$.

5.3.1 Numerical results

The implementation has been carried out in Python and the examples were run on a HP Workstation Z240 MT J9C17ET with Intel Core i7-6700 - 4 x 3.40GHz. Since we started from an arbitrary dimension throughout the whole analyzes, the codes for each dimension $d \in \{1, 2, 3\}$ own the same structure. In Appendix A we provide details on the implementation of the assembling process. The implementation of the solving procedure only consists of delivering Algorithm 1 to the CG method. Moreover, the codes are parallelized over 8 threads on the four Intel cores (see also the Appendix A.3).

In all examples we consider $\Omega = (0, 1)^d$ and the truncated fractional Laplace kernel (5.8) with a fraction $s = 0.4$. We consider a constant source term $f = 1$ and homogeneous Dirichlet data $g = 0$ in (2.1). Further we consider an interaction horizon $\delta = T + \lambda$ where $T = 2^{10}$ with coarsening parameter $q = 1.5$, minimum grid size $h_{min} = 10^{-2}$ and a parameter $\lambda > 2h$ for the box $\mathcal{B}$ (these parameters are explained in the Appendix

A.2.3). The CG method stops if a sufficient decrease of the residual $\frac{||\mathbf{A}^h\mathbf{x}_k-\mathbf{b}||}{||\mathbf{b}||} < 10^{-12}$ is reached. We present numerical examples for $d \in \{1,2,3\}$. For each grid size h we report on the number of grid points ("dofs") and the number of CG iterations ("cg its") as well as the CPU time ("CPU solving") needed for solving the discretized system. Furthermore we compute the energy error $||u_R^h - u_\infty||_{H^s(\Omega\cup\Omega_I)}$, where u_∞ is a numerical surrogate taken to be the finite element solution on the finest grid, and the rate of convergence.

1d Example

For the 1d example we choose $\lambda = 5$ as parameter for the box $\mathcal{B}$ and $n = 7$ Gauss points for the unit interval $[0,1]$. The results are presented in Figure 5.2 and Table 5.1.

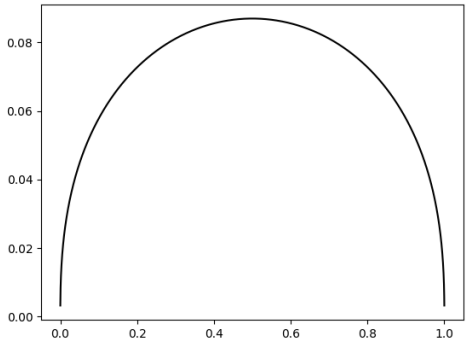


Figure 5.2: Plot of the 1d finite element solution u^h.

h	dofs	cg its	energy error	rate	CPU solving [s]
2^{-6}	63	16	2.43e-02	0.50	-
2^{-7}	127	24	1.72e-02	0.51	0.005
2^{-8}	255	34	1.21e-02	0.51	0.011
2^{-9}	511	46	8.47e-03	0.52	0.029
2^{-14}	16,383	191	-	-	5.261

Table 5.1: Results of the 1d test case.

2d Example

For the 2d example we choose $\lambda = 1$ as parameter for the box $\mathcal{B}$ and $n = 6$, i.e., 36 quadrature points for the unit square $[0,1]^2$. The results are presented in Figure 5.3 and Table 5.2.

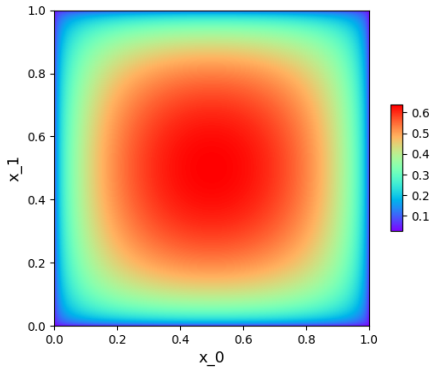


Figure 5.3: Contour plot of the 2d finite element solution u^h.

h	dofs	cg its	energy error	rate	CPU solving [s]
2^{-2}	9	3	3.11e-01	0.50	-
2^{-3}	49	10	2.17e-01	0.51	-
2^{-4}	225	16	1.53e-01	0.52	0.01
2^{-5}	961	20	1.06e-01	0.53	0.03
2^{-9}	261,121	58	-	-	3.15

Table 5.2: Results of the 2d test case.

3d Example

For the 3d example we choose $\lambda = 0.5$ as parameter for the box $\mathcal{B}$ and $n = 4$, i.e., 64 quadrature points for the unit cube $[0, 1]^3$. The results are presented in Figure 5.4 and Table 5.3.

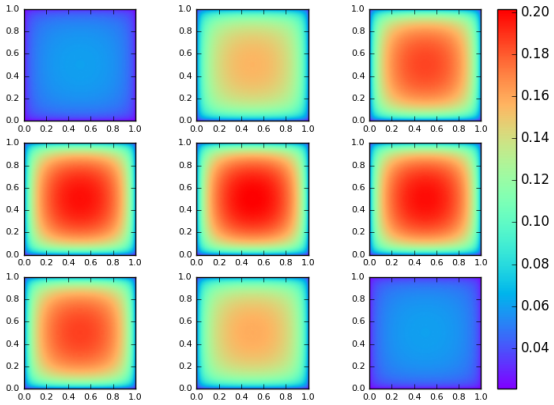


Figure 5.4: Plot of the finite element solution u^h. In order to illustrate the 3d solution we cut the domain $\Omega = (0,1)^3$ into nine slices along the third dimension ordered by increasing x_2-dimension.

h	dofs	cg its	energy error	rate	CPU solving [min]
2^{-3}	343	19	1.37e-01	0.51	-
2^{-4}	3,375	20	9.49e-02	0.52	-
2^{-5}	29,791	21	6.54e-02	0.54	0.01
2^{-6}	250,047	23	4.44e-02	0.59	0.03
2^{-9}	133,432,831	55	-	-	36.28

Table 5.3: Results of the 3d test case.

5.3.2 Discussion

Before we discuss the numerical results let us cite a result about the regularity of weak solutions and the convergence of finite element solutions.

Proposition 5.3.1 ([20, Theorem 3.5, Proposition 3.6])**.** *Let the domain $\Omega \subset \mathbb{R}^d$ have C^∞ boundary $\partial\Omega$ and let $f \in H^r(\Omega)$ for $r \geq 0$. Further let the kernel be of the form*

$$\gamma(\mathbf{x},\mathbf{y}) = \frac{c}{||\mathbf{x}-\mathbf{y}||_2^{d+2s}} \chi_{B_{\delta,2}(\mathbf{x})}(\mathbf{y}),$$

for a constant $c > 0$, such that (K1)*,* (K2) *and* (3.17) *are satisfied. Then for the solution $u \in V_c(\Omega \cup \Omega_I)$ of (3.7) the following regularity estimate holds*

$$|u|_{H^{s+\alpha}(\Omega\cup\Omega_I)} \leq C||f||_{H^r(\Omega)}, \quad C > 0,$$

where $\alpha = \min\{s + r, 1/2 - \epsilon\}$ for some arbitrarily small $\epsilon > 0$. Furthermore, by invoking this regularity estimate we obtain the following convergence result for piecewise linear finite element approximations:

$$||u - u^h||_{H^s(\Omega\cup\Omega_I)} \leq C' h^\alpha ||f||_{H^r(\Omega)}, \quad C' > 0. \tag{5.9}$$

To the best of our knowledge, for the truncated fractional kernel and less smooth domains corresponding results are not available.

Since we set $f = 1$, we find that our 1d results confirm the theoretical result given in (5.9), where

$$\alpha = \min\{s + r, 1/2 - \epsilon\} \approx 0.5.$$

Due to the numerical results for $d > 1$ we may conjecture, that this convergence result also holds for domains with less smooth boundary, such as hyperrectangles. The latter has already been shown for finite element approximations of the nontruncated problem (6.21) under a Hölder regularity assumption on the right-hand side f; see [2, Theorem 4.7].

Finally, we also note that the library pyFFTW needs a lot of memory for building the FFT object, such that we had to move the 3d computations for the finest grid to a machine with a larger RAM. We can circumvent this problem by using the sequential FFT implementation available in the NumPy library.

5.4 Concluding remarks

We want to note that the interaction horizon can also be smaller than the diameter of the domain. This complicates the numerical integration and with that the assembling procedure, but the stiffness matrix is no more fully populated and its structure still remains multilevel Toeplitz. Having that, we can model the transition to local diffusion and access a greater range of kernels.

Moreover, an aspect concerning the solving procedure, which is not examined above, is that of an efficient preconditioner for the discretized Galerkin system. A multigrid method might be a reasonable candidate due to the simple structure of the grid (see also [24, 63]). In general, a lot of effort has been put in the research of preconditioning structured matrices (see, e.g., [62, 21, 75]). Since we observe a moderate number of CG iterations in our numerical examples, a preconditioner has not been implemented.

The main drawback of our approach relies on the fact that the code is strictly limited to regular grids and is thus not applicable to more complicated domains. It is crucial that each element has the same geometry in order to achieve the multilevel Toeplitz structure of the stiffness matrix; meaning that only rectangular domains are reasonable. In contrast to that, the restriction to translation and reflection invariant kernels appears to be rather weak, since a lot of kernels treated in literature are even radial.

Chapter 6

Comparison of two nonlocal models and computational applications

In this chapter we compare the solutions of two nonlocal models posed on a bounded domain which solely differ in the kernel that is employed to evaluate nonlocal interactions. An important special case arises when the two different kernels only differ in their interaction sets which spatially limit the extent of nonlocal interactions.

The results of this chapter can then be applied to specific settings, particularly for the analysis of the effect of using interaction sets, and in particular balls, that are different from the Euclidean ball which is in standard use in nonlocal models. In addition to that, we discuss two other applications which lead to computational benefits. First, we approximate solutions corresponding to a nontruncated kernel by solutions corresponding to a truncated version of this kernel. Second, we study the effect that approximate interaction sets have on the finite element error.

More precisely, in Section 6.1 we derive the main result about the difference in solutions that correspond to two different kernels. An emphasis is then put on the special case where the two kernels possess a common $\|\cdot\|_2$-radial kernel function but one is truncated by standard Euclidean balls and the other by $\|\cdot\|_\infty$-balls. This special case is then further concretized in Section 6.2 by considering the fractional kernel. We find that we can approximate solutions of the fractional Poisson problem by solutions arising from the nonlocal model involving the fractional kernel which is truncated by $\|\cdot\|_\infty$-balls. Finally, in Section 6.3 we exploit the results from Section 6.1, especially an estimate for the difference in bilinear forms, to conduct a preliminary investigation of the effect that approximate interaction sets have on the global finite element error. The therein derived analytic results are supplemented by numerical study cases.

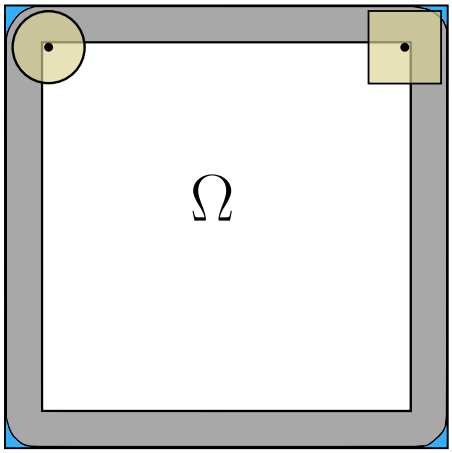


Figure 6.1: Illustration of Ω_I^{ab} for the case $S_a(\mathbf{x}) = B_{\delta,2}(\mathbf{x})$ and $S_b(\mathbf{x}) = B_{\delta,\infty}(\mathbf{x})$. The interaction domain $\Omega_I^b \supset \Omega_I^a$ additionally contains the blue colored corners, so that $\Omega_I^{ab} = \Omega_I^b$.

6.1 Comparison between models having different kernels

We consider two different kernels $\gamma_a, \gamma_b \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ of the form

$$\gamma_a(\mathbf{x},\mathbf{y}) = \phi_a(\mathbf{x},\mathbf{y})\chi_{S_a(\mathbf{x})}(\mathbf{y}) \qquad \text{and} \qquad \gamma_b(\mathbf{x},\mathbf{y}) = \phi_b(\mathbf{x},\mathbf{y})\chi_{S_b(\mathbf{x})}(\mathbf{y}) \tag{6.1}$$

with kernel functions $\phi_* \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ and families of interaction sets $\{S_*(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$ which we assume to satisfy the symmetry property (S2). We then have, for $* = \{a, b\}$, the interaction domains, operator equations, bilinear forms, spaces, weak formulations, and solutions as defined in Chapter 3. With respect to the Dirichlet data, we assume that we are given $g(\mathbf{x})$ defined for

$$\mathbf{x} \in \Omega_I^{ab} := \Omega_I^a \cup \Omega_I^b$$

so that for the weak formulations (3.7) corresponding to the two kernels, we have the data

$$g_a = g|_{\Omega_I^a} \quad \text{and} \quad g_b = g|_{\Omega_I^b}.$$

An example for Ω_I^{ab} is depicted in Figure 6.1. We assume that the forcing function f for $* = \{a, b\}$ is the same. We assume that the two kernels are chosen in such a way that the associated weak formulations are well-posed. In Chapter 3 we have presented two classes of truncated kernels satisfying this requirement. However, for the following considerations it is not necessary that interaction sets are bounded.

Because we want to compare the solutions u_a and u_b corresponding to the two kernels in (6.1), we have to take into account that those solutions are defined in different domains, i.e., $\Omega \cup \Omega_I^a$ and $\Omega \cup \Omega_I^b$ are not the same. However, we observe that the nonlocal energy (semi-)norm $|||\cdot|||$ defined in (3.6) only involves function evaluations in Ω. Thus, to effect

such a comparison we do not need to introduce extended bilinear forms for defining the following extended nonlocal energy spaces

$$\begin{cases} V_*^{ab}(\Omega\cup\Omega_I^{ab}) := \{u\in L^2(\Omega\cup\Omega_I^{ab}) \ : \ \|u\|_{V_*^{ab}(\Omega\cup\Omega_I^{ab})}<\infty\} \\ V_{*,c}^{ab}(\Omega\cup\Omega_I^{ab}) := \{u\in V_*^{ab}(\Omega\cup\Omega_I^{ab}) \ : \ u=0 \quad \text{on } \Omega_I^{ab}\} \\ \qquad\qquad\qquad = \{u\in L_c^2(\Omega\cup\Omega_I^{ab}) \ : \ |||u|||_*<\infty\}, \end{cases}$$

where

$$\|u\|_{V_*^{ab}(\Omega\cup\Omega_I^{ab})} := |||u|||_* + \|u\|_{L^2(\Omega\cup\Omega_I^{ab})} = \sqrt{A_*(u,u)} + \|u\|_{L^2(\Omega\cup\Omega_I^{ab})},$$

for $*\in\{a,b\}$. Note that although for both $*=a$ and $*=b$ these spaces are defined with respect to the same domains, we have that, in general $V_a^{ab}(\Omega\cup\Omega_I^{ab}) \neq V_b^{ab}(\Omega\cup\Omega_I^{ab})$ because the two kernel functions involved are different.

We now also define the difference kernel $\gamma_{ab} := \gamma_a - \gamma_b$ and the difference bilinear form $A_{ab} := A_a - A_b$. Then restating Lemma 3.3.4 for the difference objects just defined leads to the following result.

Corollary 6.1.1 (***Estimation of the difference in bilinear forms***)**.** *Let us assume*

$$C_{ab} = C_{ba} := \max\left\{\sup_{\mathbf{x}\in\Omega}\int_{\Omega\cup\Omega_I^{ab}} |\gamma_{ab}|\,d\mathbf{y}, \quad \sup_{\mathbf{x}\in\Omega}\int_{\Omega} |\gamma'_{ab}|\,d\mathbf{y}\right\} < +\infty, \tag{6.2}$$

then for any $u,v\in L_c^2(\Omega\cup\Omega_I^{ab})$ we find

$$|A_{ab}(u,v)| \le 2C_{ab}\|u\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}.$$

Proof. Let $u,v\in L_c^2(\Omega\cup\Omega_I^{ab})$, then by definition we have

$$\begin{aligned} A_{ab}(u,v) = (A_a - A_b)(u,v) &= \int_\Omega v\int_\Omega (u\gamma_a - u'\gamma'_a)d\mathbf{y}d\mathbf{x} + \int_\Omega vu\int_{\Omega_I}\gamma_a d\mathbf{y}d\mathbf{x} \\ &\quad - \int_\Omega v\int_\Omega (u\gamma_b - u'\gamma'_b)d\mathbf{y}d\mathbf{x} + \int_\Omega vu\int_{\Omega_I}\gamma_b d\mathbf{y}d\mathbf{x} \\ &= \int_\Omega v\int_\Omega (u\gamma_{ab} - u'\gamma'_{ab})d\mathbf{y}d\mathbf{x} + \int_\Omega vu\int_{\Omega_I}\gamma_{ab} d\mathbf{y}d\mathbf{x}. \end{aligned}$$

Since we require (6.6), we can apply Lemma 3.3.4 to the difference bilinear form A_{ab}. □

We now treat certain special cases for the choice of the kernels γ_a and γ_b and derive the resulting constant C_{ab} defined in (6.6).

Remark 6.1.2. *i) If $\phi_a = \phi_b = \phi$, then*

$$C_{ab} = C_{ba} = \max\left\{\sup_{\mathbf{x}\in\Omega}\int_{S_a(\mathbf{x})\triangle S_b(\mathbf{x})}\phi d\mathbf{y}, \ \sup_{\mathbf{x}\in\Omega}\int_{\Omega\cap(S_a(\mathbf{x})\triangle S_b(\mathbf{x}))}\phi' d\mathbf{y}\right\},$$

where $\triangle$ denotes the symmetric difference operator (see also Figure 6.2), i.e.,

$$S_a(\mathbf{x}) \triangle S_b(\mathbf{x}) := (S_a(\mathbf{x}) \cup S_b(\mathbf{x})) \setminus (S_a(\mathbf{x}) \cap S_b(\mathbf{x})) .$$

Note that, if $S_a(\mathbf{x}) \subset S_b(\mathbf{x})$, then

$$S_a(\mathbf{x}) \triangle S_b(\mathbf{x}) = S_b(\mathbf{x}) \backslash S_a(\mathbf{x}).$$

If further $\phi = \phi'$, we find

$$C_{ab} = C_{ba} = \sup_{\mathbf{x}\in\Omega} \int_{S_a(\mathbf{x}) \triangle S_b(\mathbf{x})} \phi d\mathbf{y}. \tag{6.3}$$

ii) If $\phi_a = \phi$ and $\phi_b = \mu\phi$ for some symmetric two-point function μ (e.g., a mollifier), then

$$C_{ab} = C_{ba} = \max \left\{ \sup_{\mathbf{x}\in\Omega} \int_{\Omega \cup \Omega_I^{ab}} \phi \left| \chi_{S_a(\mathbf{x})}(\mathbf{y}) - \mu \chi_{S_b(\mathbf{x})}(\mathbf{y}) \right| d\mathbf{y}, \right.$$
$$\left. \sup_{\mathbf{x}\in\Omega} \int_{\Omega} \phi' \left| \chi_{S_a(\mathbf{x})}(\mathbf{y}) - \mu \chi_{S_b(\mathbf{x})}(\mathbf{y}) \right| d\mathbf{y} \right\}.$$

If further, $\phi = \phi'$, we find

$$C_{ab} = C_{ba} = \sup_{\mathbf{x}\in\Omega} \int_{\Omega \cup \Omega_I^{ab}} \phi \left| \chi_{S_a(\mathbf{x})}(\mathbf{y}) - \mu \chi_{S_b(\mathbf{x})}(\mathbf{y}) \right| d\mathbf{y}.$$

Proof. We only treat i), because ii) is straightforward. By assumption we have $\phi_a = \phi_b = \phi$ so that

$$|\gamma_{ab}(\mathbf{x}, \mathbf{y})| = \left| \phi_a(\mathbf{x}, \mathbf{y}) \chi_{S_a(\mathbf{x})}(\mathbf{y}) - \phi_b(\mathbf{x}, \mathbf{y}) \chi_{S_b(\mathbf{x})}(\mathbf{y}) \right| = \phi(\mathbf{x}, \mathbf{y}) \left| \chi_{S_a(\mathbf{x})}(\mathbf{y}) - \chi_{S_b(\mathbf{x})}(\mathbf{y}) \right| .$$

Further, by definition of the indicator function, we find

$$\chi_{S_a(\mathbf{x})}(\mathbf{y}) - \chi_{S_b(\mathbf{x})}(\mathbf{y}) = \begin{cases} -1 & : \mathbf{y} \in S_b(\mathbf{x}) \backslash S_a(\mathbf{x}) \\ 0 & : \mathbf{y} \in (S_b(\mathbf{x}) \cap S_a(\mathbf{x})) \cup (S_a(\mathbf{x}) \cup S_b(\mathbf{x}))^c \\ 1 & : \mathbf{y} \in S_a(\mathbf{x}) \backslash S_b(\mathbf{x}). \end{cases}$$

Thus we have

$$|\chi_{S_b(\mathbf{x})}(\mathbf{y}) - \chi_{S_a(\mathbf{x})}(\mathbf{y})| = \chi_{S_a(\mathbf{x}) \triangle S_b(\mathbf{x})}(\mathbf{y}).$$

By the symmetry of the interaction sets, we also have

$$|\gamma_{ab}(\mathbf{y}, \mathbf{x})| = \phi(\mathbf{y}, \mathbf{x}) \chi_{S_a(\mathbf{x}) \triangle S_b(\mathbf{x})}(\mathbf{y}).$$

□

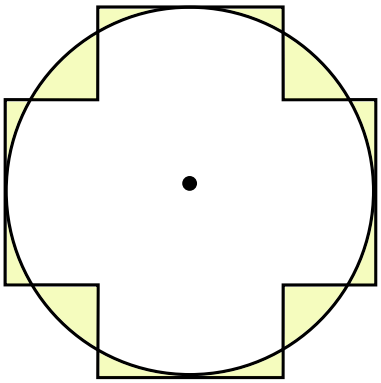


Figure 6.2: The symmetric difference between the two example interaction sets centered at the black dot is given by the yellow colored area.

We now also assume that, for $* = \{a, b\}$, the nonlocal Poincaré inequality

$$\|v\|_{L^2(\Omega)} \leq C_P^* |||v|||_* \quad \text{for all } v \in V_c^*(\Omega \cup \Omega_I^*) \tag{6.4}$$

holds with constant $C_P^* > 0$; see Chapter 3 for examples.

Lemma 6.1.3 (***Equivalence of the different energy spaces***)**.** *If $C_{ab} < +\infty$, then the extended constrained energy spaces*

$$(V_{a,c}^{ab}(\Omega \cup \Omega_I^{ab}), |||\cdot|||_a) \quad \textit{and} \quad (V_{b,c}^{ab}(\Omega \cup \Omega_I^{ab}), |||\cdot|||_b)$$

are equivalent, i.e., there exist positive constants $\underline{C}$ and $\overline{C}$ such that

$$\underline{C}|||v|||_b \leq |||v|||_a \leq \overline{C}|||v|||_b \qquad \textit{for all } \ v \in L_c^2(\Omega \cup \Omega_I).$$

Proof. Let $v \in L_c^2(\Omega \cup \Omega_I)$. By invoking Corollary 6.1.1 and the Poincaré inequality (6.4), we obtain

$$\begin{aligned} &|||v|||_a^2 - |||v|||_b^2 \\ &\qquad = A_a(v,v) - A_b(v,v) \leq |A_a(v,v) - A_b(v,v)| \\ &\qquad = |A_{ab}(v,v)| \leq C_{ab}\|v\|_{L^2(\Omega)}^2 \leq C_{ab}(C_P^b)^2 |||v|||_b^2. \end{aligned}$$

Thus, for $\overline{C} := \sqrt{1 + C_{ab}(C_P^b)^2}$ and $\underline{C} := 1/\sqrt{1 + C_{ab}(C_P^a)^2}$, it follows that

$$|||v|||_a \leq \overline{C}|||v|||_b \qquad \text{and} \qquad \underline{C}|||v|||_b \leq |||v|||_a,$$

where the second inequality follows by interchanging the roles of a and b.

□

Because the spaces $V_{a,c}^{ab}(\Omega \cup \Omega_I^{ab})$ and $V_{b,c}^{ab}(\Omega \cup \Omega_I^{ab})$ are equivalent, from now on we simply denote them by $V_c^{ab}(\Omega \cup \Omega_I^{ab})$. We can also consider the same space of trial functions in each weak formulation, so that $u_*\colon \Omega \cup \Omega_I^* \to \mathbb{R}$, for $* \in \{a, b\}$, solves

$$\begin{aligned} A_*(u_*, v) &= \int_\Omega v \int_\Omega \left(u\gamma_* - u'\gamma_*'\right) d\mathbf{y}d\mathbf{x} + \int_\Omega vu \int_{\Omega_I^*} \gamma_* d\mathbf{y}d\mathbf{x} \\ &= \int_\Omega fv d\mathbf{x} - \int_\Omega v \int_{\Omega_I^*} g'\gamma_*' d\mathbf{y}d\mathbf{x} = \ell_*(v) \end{aligned} \tag{6.5}$$

for all $v \in V_c^{ab}(\Omega \cup \Omega_I^{ab})$. We are finally in a position to derive an estimate for the difference in solutions.

Proposition 6.1.4 (***Estimation of the difference in solutions***)**.** *Given g defined on Ω_I^{ab} and f defined on Ω, for $* = \{a, b\}$, let $u_* \in V^*(\Omega \cup \Omega_I^*)$ denote the weak solution of the problem* (3.7) *with the forcing function f and boundary data $g|_{\Omega^*}$; thus solving* (6.5)*. If*

$$C_{ab} = C_{ba} := \max\left\{ \sup_{\mathbf{x}\in\Omega} \int_{\Omega\cup\Omega_I^{ab}} |\gamma_{ab}|\, d\mathbf{y}, \quad \sup_{\mathbf{x}\in\Omega} \int_{\Omega\cup\Omega_I^{ab}} |\gamma_{ab}'|\, d\mathbf{y} \right\} < +\infty, \tag{6.6}$$

then

$$|||u_b|_\Omega - u_a|_\Omega|||_a \le 2C_{ab}C_P^a \left(\|u_b\|_{L^2(\Omega)} + \|g\|_{L^2\left(\Omega_I^{ab}\right)} \right), \tag{6.7}$$

and

$$\|u_b|_\Omega - u_a|_\Omega\|_{L^2(\Omega)} \le 2C_{ab}(C_P^a)^2 \left(\|u_b\|_{L^2(\Omega)} + \|g\|_{L^2\left(\Omega_I^{ab}\right)} \right). \tag{6.8}$$

Note that we can interchange the roles of a and b, since $C_{ab} = C_{ba}$.

Proof. We first note that the norms $\|\cdot\|_{L^2(\Omega)}$ and $|||\cdot|||_*$ for $* \in \{a, b\}$ only involve function evaluations on Ω. Thus the norm evaluations in (6.7) and (6.8) are well defined, although the two solutions u_* are defined on possibly different domains $\Omega \cup \Omega_I^*$. Since u_* solves (6.5) we find that, for $v \in V_c^{ab}(\Omega \cup \Omega_I^{ab})$,

$$A_a(u_a, v) - A_b(u_b, v) = \int_\Omega v \int_{\Omega_I^b} g'\gamma_b' d\mathbf{y}d\mathbf{x} - \int_\Omega v \int_{\Omega_I^a} g'\gamma_a' d\mathbf{y}d\mathbf{x}.$$

Due to the definition (3.11) of the interaction domain Ω_I^* in terms of the interaction sets, and the fact that $\Omega_I^* \subset \Omega_I^{ab}$, we have

$$\int_{\Omega_I^{ab}\setminus\Omega_I^*} g'\gamma_*' d\mathbf{y} = 0,$$

so that

$$A_a(u_a, v) = A_b(u_b, v) + \int_\Omega v \int_{\Omega_I^{ab}} g'\gamma_{ba}'\, d\mathbf{y}d\mathbf{x}. \tag{6.9}$$

Let us define $v := u_a|_\Omega - u_b|_\Omega$ which we extend to Ω_I^{ab} by zero so that $v \in V_c^{ab}(\Omega \cup \Omega_I^{ab})$. Then by exploiting (6.9) and Corollary 6.1.1 we find

$$\begin{aligned}
|||v|||_a^2 = |A_a(v,v)| &= |A_a(u_a,v) - A_a(u_b,v)| \\
&= |A_b(u_b,v) + \int_\Omega v \int_{\Omega_I^{ab}} g' \gamma'_{ba} \, d\mathbf{y} d\mathbf{x} - A_a(u_b,v)| \\
&\leq |A_{ab}(u_b,v)| + \int_\Omega |v| \int_{\Omega_I^{ab}} |g'||\gamma'_{ab}| \, d\mathbf{y} d\mathbf{x} \\
&\leq 2C_{ab}\|u_b\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} + C_{ab}\|v\|_{L^2(\Omega)}\|g\|_{L^2\left(\Omega_I^{ab}\right)} \\
&\leq 2C_{ab}C_P^a \left(\|u_b\|_{L^2(\Omega)} + \|g\|_{L^2\left(\Omega_I^{ab}\right)}\right) |||v|||_a.
\end{aligned}$$

Inserting $v = u_a - u_b$ we finally arrive at

$$|||u_b - u_a|||_a \leq 2C_{ab}C_P^a \left(\|u_b\|_{L^2(\Omega)} + \|g\|_{L^2\left(\Omega_I^{ab}\right)}\right),$$

which proves the first estimate (6.7). Applying the nonlocal Poincaré inequality (6.4) for $* = a$ leads to the second estimate (6.8). □

6.1.1 Specialization to norm induced balls

We specialize the results of the previous subsection to norm induced interaction sets, and to the case $\phi_a = \phi_b = \phi$ where ϕ is assumed to be radial with respect to the $\|\cdot\|_2$-norm. In this case, the constant C_{ab} in (6.3) further simplifies to

$$C_{ab} = \sup_{\mathbf{x}\in\Omega} \int_{S_a(\mathbf{x}) \triangle S_b(\mathbf{x})} \phi(\|\mathbf{y}-\mathbf{x}\|_2) d\mathbf{y} = \int_{S_a(\mathbf{0}) \triangle S_b(\mathbf{0})} \phi(\|\mathbf{z}\|_2) d\mathbf{z}.$$

In specifying the two kernels (6.1), we consider balls based on both the $\|\cdot\|_2$- and $\|\cdot\|_\infty$-norm. We provide exact representations for C_{ab} in terms of integrals for several situations involving different concentric ball types and ball radii.

Lemma 6.1.5. *Let γ_a and γ_b be as in (6.1) with ϕ radial with respect to the $\|\cdot\|_2$-norm, i.e., $\phi(\mathbf{x},\mathbf{y}) = \phi(\|\mathbf{x}-\mathbf{y}\|_2)$. Then, for the following choices of balls, the constant*

$$C_{ab} = \int_{S_a(\mathbf{0}) \triangle S_b(\mathbf{0})} \phi(\|\mathbf{z}\|_2) d\mathbf{z}$$

can be defined exactly in terms of integrals.

i) ***$\|\cdot\|_2$-balls of different radii.*** *For $S_a(\mathbf{x}) := B_{\delta_a,2}(\mathbf{x})$ and $S_b(\mathbf{x}) := B_{\delta_b,2}(\mathbf{x})$, with $\delta_a < \delta_b < \infty$, we have*

$$C_{ab} = dV_d \int_{\delta_a}^{\delta_b} \phi(r) r^{d-1} dr, \tag{6.10}$$

where $V_d := \pi^{\frac{d}{2}}/\Gamma(\frac{d}{2}+1)$ denotes the volume of the unit $\|\cdot\|_2$-ball in $\mathbb{R}^d$ and $\Gamma(\cdot)$ denotes the Gamma function.

ii) ***Embedded*** $\|\cdot\|_2$***- and*** $\|\cdot\|_\infty$***-balls of different radii.*** *For* $S_a(\mathbf{x}) := B_{\delta_2,2}(\mathbf{x})$ *and* $S_b(\mathbf{x}) := B_{\delta_\infty,\infty}(\mathbf{x})$ *such that either* $S_b(\mathbf{x}) \subset S_a(\mathbf{x})$ *or conversely, we have*

$d = 2$:

if $B_{\delta_\infty,\infty}(\mathbf{x}) \subset B_{\delta_2,2}(\mathbf{x})$, *then* $\delta_2 \geq \sqrt{2}\delta_\infty$ *and*

$$C_{ab} = 8\int_0^{\pi/4}\int_{\delta_\infty/\cos\theta}^{\delta_2} \phi(r) r^{d-1} dr d\theta;$$

if $B_{\delta_a,2}(\mathbf{x}) \subset B_{\delta_b,\infty}(\mathbf{x})$, *then* $\delta_2 \leq \delta_\infty$ *and*

$$C_{ab} = 8\int_0^{\pi/4}\int_{\delta_2}^{\delta_\infty/\cos\theta} \phi(r) r^{d-1} dr d\theta. \tag{6.11}$$

$d = 3$:

if $B_{\delta_b,\infty}(\mathbf{x}) \subset B_{\delta_2,2}(\mathbf{x})$, *then* $\delta_2 \geq \sqrt{3}\delta_\infty$ *and*

$$\begin{aligned} C_{ab} =& 16\int_0^{\pi/4}\int_0^{\arctan(\cos(\theta_1))}\int_{\frac{\delta_\infty}{\cos(\theta_1)\cos(\theta_2)}}^{\delta_2} \phi(r) r^2\cos(\theta_2) dr d\theta_2 d\theta_1 \\ &+16\int_0^{\pi/4}\int_{\arctan(\cos(\theta_1))}^{\pi/2}\int_{\frac{\delta_\infty}{\sin(\theta_2)}}^{\delta_2} \phi(r) r^2\cos(\theta_2) dr d\theta_2 d\theta_1; \end{aligned}$$

if $B_{\delta_2,2}(\mathbf{x}) \subset B_{\delta_\infty,\infty}(\mathbf{x})$, *then* $\delta_2 \leq \delta_\infty$ *and*

$$\begin{aligned} C_{ab} =& 16\int_0^{\pi/4}\int_0^{\arctan(\cos(\theta_1))}\int_{\delta_2}^{\frac{\delta_\infty}{\cos(\theta_1)\cos(\theta_2)}} \phi(r) r^2\cos(\theta_2) dr d\theta_2 d\theta_1 \\ &+16\int_0^{\pi/4}\int_{\arctan(\cos(\theta_1))}^{\pi/2}\int_{\delta_2}^{\frac{\delta_\infty}{\sin(\theta_2)}} \phi(r) r^2\cos(\theta_2) dr d\theta_2 d\theta_1. \end{aligned} \tag{6.12}$$

iii) $\|\cdot\|_2$***- and*** $\|\cdot\|_\infty$***-balls of equal volume.*** *For* $S_a(\mathbf{x}) := B_{\delta_2,2}(\mathbf{x})$ *and* $S_b(\mathbf{x}) := B_{\delta_\infty,\infty}(\mathbf{x})$, *we have*

$d = 2$:

$$\begin{aligned} C_{ab} =& 8\int_0^{\arccos((\pi/4)^{1/2})}\int_{\delta_2(\pi/4)^{1/2}/\cos\theta}^{\delta_2} \phi(r) r dr d\theta \\ &+ 8\int_{\arccos((\pi/4)^{1/2})}^{\pi/4}\int_{\delta_2}^{\delta_2(\pi/4)^{1/2}/\cos\theta} \phi(r) r dr d\theta. \end{aligned} \tag{6.13}$$

$d = 3$:

$$\begin{aligned} &C_{ab} \\ &=16\int_0^{\arccos((\frac{\pi}{6})^{1/3})}\int_{\arccos((\frac{\pi}{6})^{1/3}/\cos(\theta_1))}^{\arctan(\cos(\theta_1))}\int_{\delta_2}^{\delta_\infty/(\cos(\theta_1)\cos(\theta_2))} \phi(r) r^2\cos(\theta_2) dr d\theta_2 d\theta_1 \end{aligned}$$

$$+16\int_{\arccos((\frac{\pi}{6})^{1/3})}^{\pi/4}\int_{0}^{\arctan(\cos(\theta_1))}\int_{\delta_2}^{\delta_\infty/(\cos(\theta_1)\cos(\theta_2))}\phi(r)r^2\cos(\theta_2)drd\theta_2d\theta_1$$
$$+16\int_{0}^{\pi/4}\int_{\arctan(\cos(\theta_1))}^{\arcsin((\frac{\pi}{6})^{1/3})}\int_{\delta_2}^{\delta_\infty/\sin(\theta_2)}\phi(r)r^2\cos(\theta_2)drd\theta_2d\theta_1$$
$$+16\int_{0}^{\pi/4}\int_{\arcsin((\frac{\pi}{6})^{1/3})}^{\pi/2}\int_{\delta_\infty/\sin(\theta_2)}^{\delta_2}\phi(r)r^2\cos(\theta_2)drd\theta_2d\theta_1.$$

Proof. i) In this case the symmetric difference becomes $S_a(\mathbf{0})\triangle S_b(\mathbf{0}) = B_{\delta_b,2}(\mathbf{0})\backslash B_{\delta_a,2}(\mathbf{0})$. Note that

$$\begin{aligned}&\int_{B_{\delta_b,2}(\mathbf{0})\backslash B_{\delta_a,2}(\mathbf{0})}\phi(\|\mathbf{z}\|_2)d\mathbf{z}\\ &\quad=\int_0^\pi\dots\int_0^\pi\int_0^{2\pi}\int_{\delta_a}^{\delta_b}\phi(r)r^{d-1}\Theta(\theta_1,\dots,\theta_{d-1})drd\theta_{d-1}\cdots d\theta_1\\ &\quad=\left(\int_{\delta_a}^{\delta_b}\phi(r)r^{d-1}dr\right)\left(\int_0^\pi\dots\int_0^\pi\int_0^{2\pi}\Theta(\theta_1,\dots,\theta_{d-1})d\theta_{d-1}\cdots d\theta_1\right)\end{aligned}$$

where we put all the remaining (sine and cosine) terms from the volume element of the d-dimensional spherical system into $\Theta(\theta_1,\dots,\theta_{d-1})$; note that these do not depend on r. The second factor equals

$$\frac{1}{\int_0^1 r^{d-1}dr}\left(\left(\int_0^1 r^{d-1}dr\right)\int_0^\pi\dots\int_0^\pi\int_0^{2\pi}\Theta(\theta_1,\dots,\theta_{d-1}d\theta_{d-1})\cdots d\theta_1\right)=dV_d$$

from which (6.10) easily follows.

ii) If $d=2$, the proofs are elementary so they are omitted here. If $d=3$, it is sufficient to consider the case where the $\|\cdot\|_2$-ball is embedded into the $\|\cdot\|_\infty$-ball (then we only have to reverse the integration bounds for the opposite case), i.e.,

$$C_{ba}=\int_{B_{\delta_\infty,\infty}(\mathbf{0})\backslash B_{\delta_2,2}(\mathbf{0})}\phi(||\mathbf{z}||_2)d\mathbf{z}.$$

We split the domain $B_{\delta_\infty,\infty}(\mathbf{0})\backslash B_{\delta_2,2}(\mathbf{0})$ into $16=2^{d+1}$ slices for which the integral is equal due to symmetry. We pick a particular slice which can be parametrized by

$$\Phi(\theta_1,\theta_2,r):=r\begin{pmatrix}\cos(\theta_1)\cos(\theta_2)\\ \sin(\theta_1)\cos(\theta_2)\\ \sin(\theta_2)\end{pmatrix},\tag{6.14}$$

where $\theta_1\in[0,\pi/4]$, $\theta_2\in[0,\pi/2]$, $r\geq\delta_2$ and

$$|\det d\Phi(\theta_1,\theta_2,r)|=r^2\cos(\theta_2).$$

Now, we have to find the precise intervals for the angles (θ_1,θ_2) and the radius $r=r(\theta_1,\theta_2)$ such that we correctly parametrize the domain of interest. The lower bound

for the radius $r = r(\theta_1, \theta_2)$ is given by the radius of the $\|\cdot\|_2$-ball δ_2, whereas the upper bound by the length of the vector which is given by the intersection of a "ray"

$$r(\theta_1, \theta_2) := \{\Phi(\theta_1, \theta_2, \mu) \colon \mu \geq 0\}$$

with the boundary of $B_{\delta_\infty,\infty}(\mathbf{0})$. Due to our choice of slice we know that such intersections can only occur at the following two parts of $\partial B_{\delta_\infty,\infty}(\mathbf{0})$: The "wall"

$$W := \{(\delta_\infty, \lambda_1, \lambda_2) \colon \lambda_i \in (0, \delta_\infty)\}$$

and the "roof"

$$R := \{(\lambda_1, \lambda_2, \delta_\infty) \colon \lambda_i \in (0, \delta_\infty)\}.$$

By equating the determining equations we straightforwardly find the intersection points

$$r(\theta_1, \theta_2) \cap W = \left\{ \frac{\delta_\infty}{\cos(\theta_1)\cos(\theta_2)} \begin{pmatrix} \cos(\theta_1)\cos(\theta_2) \\ \sin(\theta_1)\cos(\theta_2) \\ \sin(\theta_2) \end{pmatrix} \right\}$$

and

$$r(\theta_1, \theta_2) \cap R = \left\{ \frac{\delta_\infty}{\sin(\theta_2)} \begin{pmatrix} \cos(\theta_1)\cos(\theta_2) \\ \sin(\theta_1)\cos(\theta_2) \\ \sin(\theta_2) \end{pmatrix} \right\}.$$

Finally we have to partition the set $[0, \pi/4] \times [0, \pi/2]$ into those pairs (θ_1, θ_2) for which the ray $r(\theta_1, \theta_2)$ intersects W and those for which it intersects R. For the "wall" case the angles are determined by

$$\begin{aligned} 0 \leq \lambda_1 &= \delta_\infty \tan(\theta_1) \leq \delta_\infty \\ 0 \leq \lambda_2 &= \delta_\infty \tan(\theta_2)/\cos(\theta_1) \leq \delta_\infty \end{aligned}$$

leading to

$$\begin{aligned} &\{(\theta_1, \theta_2) \in [0, \pi/4] \times [0, \pi/2] \colon r(\theta_1, \theta_2) \cap W \neq \emptyset\} \\ =&\{(\theta_1, \theta_2) \in [0, \pi/4] \times [0, \pi/2] \colon \theta_2 \leq \arctan(\cos(\theta_1))\}. \end{aligned}$$

Consequently by considering the complement of this set we find for the "roof" case

$$\begin{aligned} &\{(\theta_1, \theta_2) \in [0, \pi/4] \times [0, \pi/2] \colon r(\theta_1, \theta_2) \cap R \neq \emptyset\} \\ =&\{(\theta_1, \theta_2) \in [0, \pi/4] \times [0, \pi/2] \colon \theta_2 > \arctan(\cos(\theta_1))\}. \end{aligned}$$

iii) If $d = 2$, the equal area assumption requires that $\delta_\infty = \sqrt{\frac{\pi}{4}}\delta_2$ and the symmetric difference $S_a(\mathbf{0}) \triangle S_b(\mathbf{0})$ consists of four circular segments and four corner regions; see Figure 6.4. The integrals of $\phi(\mathbf{z})$ over all four circular segments is the same, as are the integrals over all four corner regions. We first consider the circular segments. Let φ denote the corresponding central angle of a representative segment. Then $\delta_\infty = \delta_2 \cos(\frac{\varphi}{2})$

so that because $\delta_\infty = \sqrt{\frac{\pi}{4}}\delta_2$ as well, we obtain $\varphi = 2\arccos(\sqrt{\frac{\pi}{4}})$. Then, symmetry properties and simple algebraic manipulation to determine the limits of integration lead to the contribution

$$8\int_0^{\arccos(\pi/4)^{1/2}} \int_{\delta_2(\pi/4)^{1/2}/\cos\theta}^{\delta_2} \phi(r) r dr d\theta \tag{6.15}$$

to C_{ab} due to the four circular segments. Using similar simple algebraic and trigonometric manipulations as those used to arrive at (6.15) lead to the contribution

$$8\int_{\arccos(\pi/4)^{1/2}}^{\pi/4} \int_{\delta_2}^{\delta_2(\pi/4)^{1/2}/\cos\theta} \phi(r) r dr d\theta$$

to C_{ab} due to the four corner regions.

If $d = 3$, we want to compute

$$C_{ab} = \int_{B_{\delta_\infty,\infty}(\mathbf{0}) \triangle B_{\delta_2,2}(\mathbf{0})} \phi(||\mathbf{z}||_2) d\mathbf{z},$$

where both balls have equal volume, i.e., $\delta_\infty = (\frac{\pi}{6})^{1/3}\delta_2$. The symmetric difference $B_{\delta_\infty,\infty}(\mathbf{0}) \triangle B_{\delta_2,2}(\mathbf{0})$ consists of 8 corners and 6 spherical caps. We can consider the same slice as in the proof of *ii)* for $d = 3$ and proceed in an analogous manner in order to determine the appropriate intervals for the angles and the radius when using the parametrization (6.14). Let us first consider the two regions for which rays intersect the roof (see upper two regions in Figure 6.3) and let us start with the part of the cap. We have that rays $r(\theta_1, \theta_2)$ intersect the roof for all $\theta_1 \in (0, \pi/4)$. The interval for θ_2 and the radius is then simply determined by

$$r_{min}(\theta_1, \theta_2) = \delta_\infty / \sin(\theta_2) \leq \delta_2$$

from which it follows that $\theta_2 \in (\arcsin((\frac{\pi}{6})^{1/3}), \pi/2)$ and $r \in (\delta_\infty/\sin(\theta_2), \delta_2)$. Next, let us consider the part of the corner for which rays intersect with the roof. Here we have

$$\begin{aligned} \delta_\infty/\sin(\theta_2) &\geq \delta_2 \\ 0 \leq \lambda_1 = \delta_\infty \cos(\theta_1)/\tan(\theta_2) &\leq \delta_\infty \\ 0 \leq \lambda_2 = \delta_\infty \sin(\theta_1)/\tan(\theta_2) &\leq \delta_\infty \end{aligned}$$

from which it follows that $\theta_1 \in (0, \pi/4)$, $\theta_2 \in \left(\arctan(\cos(\theta_1)), \arcsin((\frac{\pi}{6})^{1/3})\right)$ and $r \in (\delta_2, \delta_\infty/\sin(\theta_2))$. Finally the part of the corner for which rays intersect with the wall is determined by the system

$$\begin{aligned} \delta_\infty/(\cos(\theta_1)\cos(\theta_2)) &\geq \delta_2 \\ 0 \leq \lambda_1 = \delta_\infty \tan(\theta_1) &\leq \delta_\infty \\ 0 \leq \lambda_2 = \delta_\infty \tan(\theta_2)/\cos(\theta_1) &\leq \delta_\infty. \end{aligned}$$

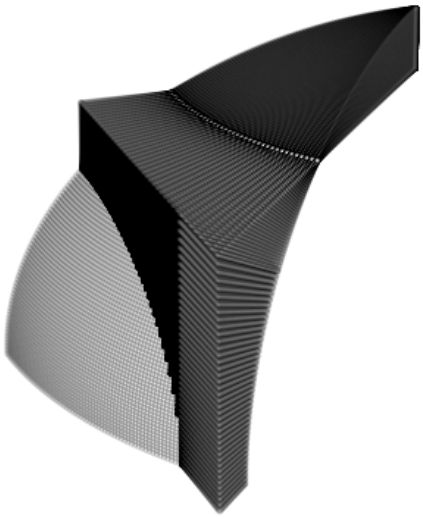


Figure 6.3: Equal volume case. The 1/16-slice which is divided into 4 disjoint sets. The lower left, light gray region is not part of the integration domain discussed in the proof of Lemma 6.1.5.

Due to the first requirement we find two disjoint sets determining the intervals (see lower two parts in Figure 6.3). One is given by

$$\begin{aligned}
\theta_1 &\in \left(0,\ \arccos\left((\tfrac{\pi}{6})^{1/3}\right)\right)\\
\theta_2 &\in \left(\arccos\left((\tfrac{\pi}{6})^{1/3}/\cos(\theta_1)\right),\ \arctan(\cos(\theta_1))\right)\\
r &\in (\delta_2,\ \delta_\infty/(\cos(\theta_1)\cos(\theta_2)))
\end{aligned}$$

the other one by

$$\begin{aligned}
\theta_1 &\in \left(\arccos((\tfrac{\pi}{6})^{1/3}),\ \pi/4\right)\\
\theta_2 &\in (0,\ \arctan(\cos(\theta_1)))\\
r &\in (\delta_2,\ \delta_\infty/(\cos(\theta_1)\cos(\theta_2)))\,.
\end{aligned}$$

Applying the parametrization with the afore derived intervals leads to the 4 integrals given in the lemma. □

In the two-dimensional case, $d = 2$, Lemma 6.1.5 can be summarized as follows. Let

$\delta_\infty = \lambda\delta_2$ with $\lambda > 0$, then

$$C_{ab}(\lambda) = \begin{cases} 8\int_0^{\pi/4}\int_{\lambda\delta_2/\cos(\theta)}^{\delta_2}\phi(r)rdrd\theta & : 0 < \lambda < \frac{1}{\sqrt{2}} \quad (i.e.,\ B_\infty \subset B_2) \\ 8\int_0^{\arccos(\lambda)}\int_{\lambda\delta_2/\cos(\theta)}^{\delta_2}\phi(r)rdrd\theta & \\ +8\int_{\arccos(\lambda)}^{\pi/4}\int_{\delta_2}^{\lambda\delta_2/\cos(\theta)}\phi(r)rdrd\theta & : \frac{1}{\sqrt{2}} \leq \lambda \leq 1 \\ 8\int_0^{\pi/4}\int_{\delta_2}^{\lambda\delta_2/\cos(\theta)}\phi(r)rdrd\theta & : 1 < \lambda \quad (i.e.,\ B_2 \subset B_\infty). \end{cases}$$

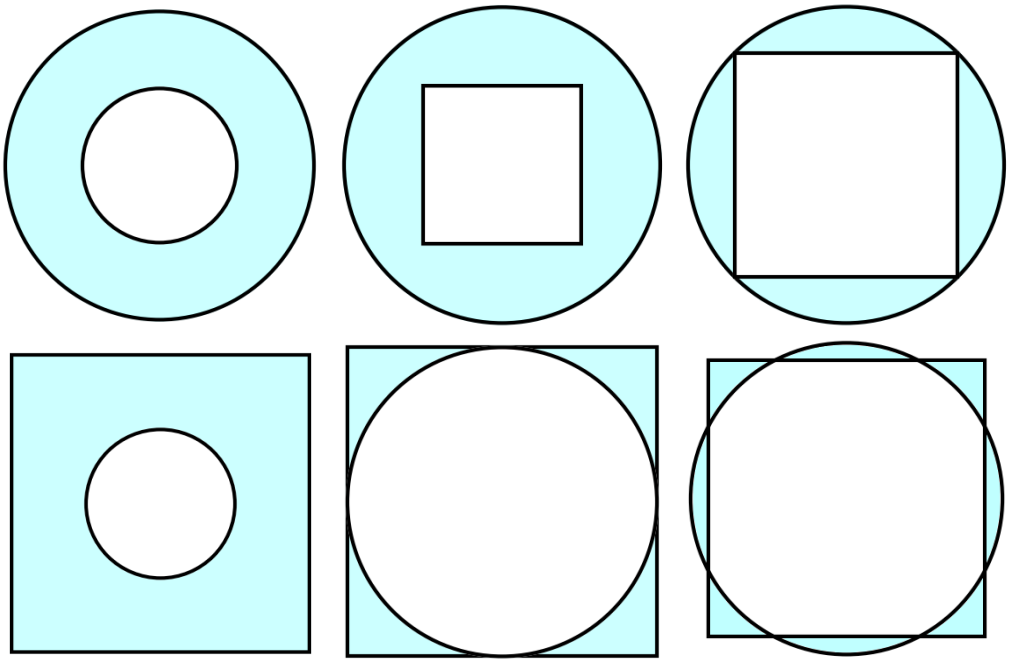


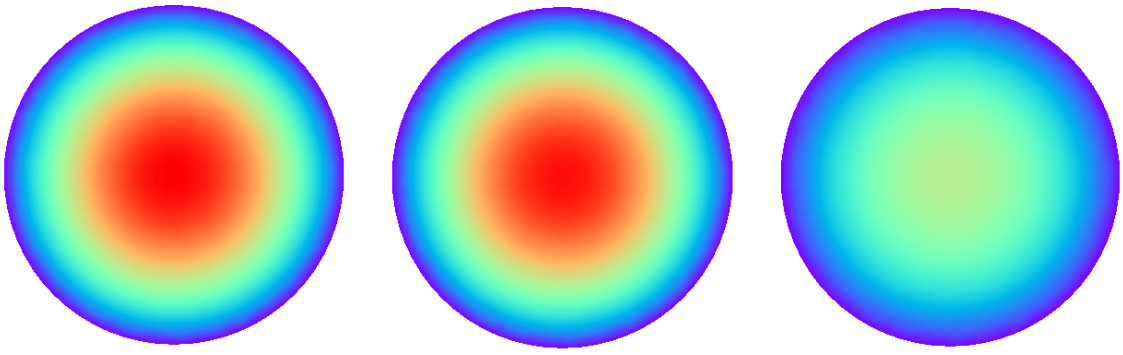
Figure 6.4: Examples in two dimensions for the different cases appearing in Lemma 6.1.5. The colored area indicates the symmetric difference of the two balls. The figure in the lower right corner corresponds to the case of equal volume, whereas all the others are examples of embedded balls.

6.1.2 A numerical example comparing two balls: $\|\cdot\|_2$ versus $\|\cdot\|_\infty$

In the following numerical experiment we compare the solution of a nonlocal model where the interaction sets are $\|\cdot\|_2$-balls to a nonlocal model associated with $\|\cdot\|_\infty$-balls. We set $\Omega = B_{1,2}(\mathbf{0})$ and

$$\gamma_*(\mathbf{x},\mathbf{y}) := C\chi_{B_{\delta,*}(\mathbf{x})}(\mathbf{y}), \quad \text{where } * \in \{2,\infty\}, C = 1000.$$

Furthermore we consider the forcing term $f = 1$ on Ω and Dirichlet data $g = 0$ on Ω_I^{ab}. In Figure 6.5 we report solutions for the $\|\cdot\|_2$-ball (left), $\|\cdot\|_\infty$-ball with the same volume (middle) and $\|\cdot\|_\infty$-ball with the same interaction radius (right). We note that using a ball with the same volume delivers a solution that is very close to the one obtained with standard balls. This is confirmed by the fact that $\|u_2 - u_\infty\|_{L^2(\Omega)} \approx 0.13$ (equal radius), whereas $\|u_2 - \tilde{u}_\infty\|_{L^2(\Omega)} \approx 0.0047$ (equal volume). Since the symmetric difference between the $\|\cdot\|_2$-ball and the $\|\cdot\|_\infty$-ball is smaller for the equal volume case than for the equal radius case, we find that this observation is in line with (6.13) and the estimate derived in (6.8) with $u_b = u_2$ and $u_a \in \{u_\infty, \tilde{u}_\infty\}$.



(a) (b) (c)

Figure 6.5: (a) Nonlocal solution u_2 for the case of a $\|\cdot\|_2$-ball ($* = 2$, $\delta = 0.1$). (b) Nonlocal solution $\tilde{u}_\infty$ for the case of a $\|\cdot\|_\infty$-ball of equal volume ($* = \infty$, $\delta = \frac{\sqrt{\pi}}{2}0.1$). (c) Nonlocal solution u_∞ for the case of a $\|\cdot\|_\infty$-ball of equal radius ($* = \infty$, $\delta = 0.1$). In all plots we use the same color scale, i.e., the same color indicates the same value.

6.2 Application to large horizons and the infinite horizon limit

In this section we assume that the interaction sets are large enough to include all points in Ω, i.e.,

$$\Omega \subset S(\mathbf{x}) \quad \text{for all} \quad \mathbf{x} \in \Omega.$$

In Figure 6.6 this situation is illustrated. The bilinear form (3.5) then becomes

$$A(u,v) = \int_\Omega \int_\Omega v(u\phi - u'\phi')d\mathbf{y}d\mathbf{x} + \int_\Omega uv\Big(\int_{S(\mathbf{x})\setminus\Omega} \phi d\mathbf{y}\Big)d\mathbf{x}. \tag{6.16}$$

We now consider two kernels that solely differ in the choice of interaction sets. More precisely, let

$$\gamma_a(\mathbf{x},\mathbf{y}) = \phi(\mathbf{x},\mathbf{y})\chi_{S_a(\mathbf{x})}(\mathbf{y}) \quad \text{and} \quad \gamma_b(\mathbf{x},\mathbf{y}) = \phi(\mathbf{x},\mathbf{y})\chi_{S_b(\mathbf{x})}(\mathbf{y}),$$

with interaction sets, that satisfy

$$\Omega \subset S_a(\mathbf{x}) \subset S_b(\mathbf{x}) \quad \text{for all} \quad \mathbf{x} \in \Omega.$$

Then from (6.16) we find for the difference bilinear form

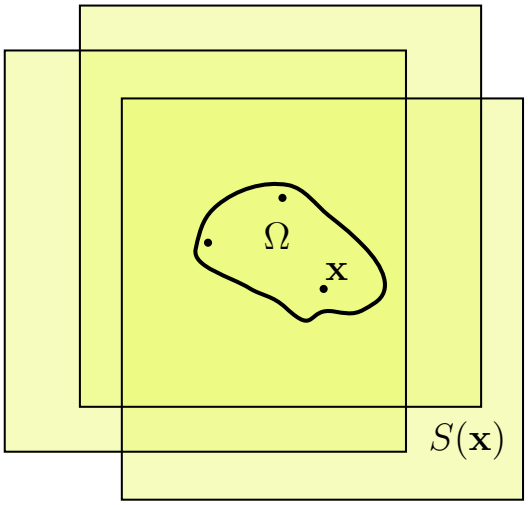


Figure 6.6: Illustration of a large interaction horizon. Here, $S(\mathbf{x}) = B_{\delta,\infty}(\mathbf{x})$ with $\delta > 0$ chosen large enough so that $\Omega \subset B_{\delta,\infty}(\mathbf{x})$ for all points $\mathbf{x} \in \Omega$.

$$A_b(u,v) - A_a(u,v) = \int_\Omega u(\mathbf{x})v(\mathbf{x}) \left(\int_{S_b(\mathbf{x})\setminus S_a(\mathbf{x})} \phi(\mathbf{x},\mathbf{y}) d\mathbf{y} \right) d\mathbf{x}.$$

Under the assumption that the integral term in brackets is constant in $\mathbf{x}$, so that

$$\int_{S_b(\mathbf{x})\setminus S_a(\mathbf{x})} \phi(\mathbf{x},\mathbf{y}) d\mathbf{y} = C_{ba},$$

we further find

$$A_b(u,v) - A_a(u,v) = C_{ba}(u,v)_{L^2(\Omega)}. \tag{6.17}$$

Note that this constant is precisely the special case (6.3) of the constant C_{ba} declared in (6.6).

Remark 6.2.1. *A practical consequence of* (6.17) *is that the difference between the finite element stiffness matrices corresponding to the two bilinear forms $A_*(u,v)$, $* = \{a,b\}$, is a constant multiple of the mass matrix. For instance, we can assemble the stiffness matrix corresponding to $\|\cdot\|_\infty$-balls and adjust it by a scalar multiple of the mass matrix in oder to obtain the stiffness matrix corresponding to $\|\cdot\|_2$-balls. See Remark 6.2.4 for a related discussion.*

In the following, we want to consider the size of the interaction set as a varying parameter and therefore focus on interaction sets that are norm induced balls $S(\mathbf{x}) = B_{\delta,\bullet}(\mathbf{x})$ with $\delta \geq \operatorname{diam}(\Omega)$ (or even $\delta \gg \operatorname{diam}(\Omega)$). Furthermore we want to rely on the results derived in Lemma 6.1.5 which is why we now also specialize to kernel functions that are radial with respect to the $\|\cdot\|_2$-norm, i.e., $\phi(\mathbf{x},\mathbf{y}) = \phi(\|\mathbf{y}-\mathbf{x}\|_2)$ so that

$$C_{ba} = \int_{S_b(\mathbf{0})\setminus S_a(\mathbf{0})} \phi(\|\mathbf{z}\|_2) d\mathbf{z}. \tag{6.18}$$

6.2.1 The truncated fractional Laplacian kernel and the infinite horizon limit

We consider the case of a large horizon δ and the limit as $\delta \to \infty$. Being the most studied setting for that limit, we only consider fractional diffusion.

In [33], the fractional Laplace operator is introduced as a special case of the nonlocal diffusion operator given in (2.1). There it is shown that for a specific choice of kernel γ, the weak solution of the nonlocal problem (3.7) converges, as $\delta \to \infty$, to the weak solution of a similar problem governed by the fractional Laplace equation. Here, we generalize this result by casting the problem of approximating fractional diffusion by nonlocal models into our framework and thus allowing for different types of balls for truncating the kernel function.

The fractional Laplace operator $(-\Delta)^s$ is the pseudo-differential operator with Fourier symbol [59]

$$\mathcal{F}((-\Delta)^s u)(\xi) = \|\xi\|_2^{2s}\widehat{u}(\xi) \quad \text{for } 0 < s < 1,$$

where $\widehat{u}$ denotes the Fourier transform of u. An equivalent characterization of the fractional Laplacian is given by [54]

$$(-\Delta)^s u = c_{d,s} \int_{\mathbb{R}^d} \frac{u(\mathbf{x}) - u(\mathbf{y})}{\|\mathbf{y}-\mathbf{x}\|_2^{d+2s}} d\mathbf{y} \qquad \text{for } 0 < s < 1, \tag{6.19}$$

where the normalizing constant $c_{d,s}$ is given by

$$c_{d,s} = \frac{2^{2s} s \Gamma(s + \frac{d}{2})}{\pi^{d/2}\Gamma(1-s)}. \tag{6.20}$$

We consider the homogeneous volume constrained *fractional Poisson problem*

$$\begin{cases} (-\Delta)^s u(\mathbf{x}) = f(\mathbf{x}) & \text{for } \mathbf{x} \in \Omega \\ u(\mathbf{x}) = 0 & \text{for } \mathbf{x} \in \Omega^c, \end{cases} \tag{6.21}$$

where $f \in L^2(\Omega)$. The associated constrained energy space is

$$H_c^s(\mathbb{R}^d) = \{u \in H^s(\mathbb{R}^d) \ : \ u = 0 \ \text{ on } \Omega^c\},$$

where $H^s(\mathbb{R}^d)$ denotes the Sobolev space of order s [3] given by

$$H^s(\mathbb{R}^d) := \left\{ u \in L^2(\mathbb{R}^d) \colon |u|_{H^s(\mathbb{R}^d)} := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{(u(\mathbf{x}) - u(\mathbf{y}))^2}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}} d\mathbf{y} d\mathbf{x} < \infty \right\}.$$

A weak formulation of the problem (6.21) is given as

$$\begin{aligned} &\textit{given } f \in L^2(\Omega), \ \textit{find } u_{d,s} \in H_c^s(\mathbb{R}^d) \textit{ such that} \\ &A_{d,s}(u_{d,s}, v) = \int_\Omega f v d\mathbf{x} \quad \textit{for all } v \in H_c^s(\mathbb{R}^d), \end{aligned} \tag{6.22}$$

with associated bilinear form

$$\begin{aligned} &A_{d,s}(u, v) \\ &:= c_{d,s} \int_\Omega v(\mathbf{x}) \int_\Omega \frac{(u(\mathbf{x}) - u(\mathbf{y}))}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}} d\mathbf{y} d\mathbf{x} + \int_\Omega u(\mathbf{x}) v(\mathbf{x}) \int_{\Omega^c} \frac{c_{d,s}}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}} d\mathbf{y} d\mathbf{x}. \end{aligned} \tag{6.23}$$

For any $u, v \in H_c^s(\mathbb{R}^d)$ we can also write (compare to (3.10))

$$A_{d,s}(u, v) = \frac{c_{d,s}}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{(u(\mathbf{x}) - u(\mathbf{y})(v(\mathbf{x}) - v(\mathbf{y}))}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}} d\mathbf{y} d\mathbf{x}.$$

The existence and uniqueness of such $u_{d,s} \in H_c^s(\mathbb{R}^d)$ is shown in [91] for all $s \in (0, 1)$. Note that the well-posedness theory developed in Chapter 3 does not apply to infinite interactions. Furthermore, a Poincaré inequality of the form

$$\|v\|_{L^2(\Omega)} \leq C_P \sqrt{A_{d,s}(v, v)} \quad \text{for all } v \in H_c^s(\mathbb{R}^d)$$

holds for bounded Lipschitz domains Ω with constant $C_P = C_P(\Omega, d, s)$ [2, Proposition 2.4].

Clearly, if the kernel γ in (2.1) is chosen as

$$\gamma(\mathbf{x}, \mathbf{y}) = \frac{c_{d,s}}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}} \qquad \text{for } \ \mathbf{x}, \mathbf{y} \in \mathbb{R}^d, \ \mathbf{x} \neq \mathbf{y}, \ \ 0 < s < 1, \tag{6.24}$$

then, the fractional Laplace operator $(-\Delta)^s$ is a special case of the nonlocal operator $-\mathcal{L}$ defined in (2.1) corresponding to this specific choice for the kernel γ. Furthermore, the strong and weak formulations (6.21) and (6.22) are special cases of (2.1) and (3.7) with $g = 0$, respectively.

We aim to approximate weak solutions of the nontruncated fractional diffusion problem (6.22) by solutions of the nonlocal diffusion problem (3.7) with a *truncated fractional Laplacian kernel*

$$\gamma(\mathbf{x}, \mathbf{y}) = \frac{c_{d,s}}{\|\mathbf{x} - \mathbf{y}\|_2^{d+2s}} \mathcal{X}_{B_{\delta,\bullet}(\mathbf{x})}(\mathbf{y}) \quad \text{for } \ \mathbf{x}, \mathbf{y} \in \mathbb{R}^d, \ \mathbf{x} \neq \mathbf{y}, \ 0 < s < 1,$$

for different choices of norm induced balls $B_{\delta,\bullet}(\mathbf{x})$. To this end, we make use of the following lemma that gives precise formulas for the constant C_{ba} in (6.18) for several combinations of interaction balls. It is a special case of Lemma 6.1.5 with the fractional kernel function

$$\phi(r) = \frac{c_{d,s}}{r^{d+2s}}.$$

Lemma 6.2.2. *Let $\delta \geq \operatorname{diam}(\Omega)$. Then, for the kernels*

$$\gamma_a(\mathbf{x},\mathbf{y}) := \frac{c_{d,s}}{\|\mathbf{x}-\mathbf{y}\|_2^{d+2s}}\mathcal{X}_{S_a(\mathbf{x})}(\mathbf{y}) \;\; \text{and} \;\; \gamma_b(\mathbf{x},\mathbf{y}) := \frac{c_{d,s}}{\|\mathbf{x}-\mathbf{y}\|_2^{d+2s}}\mathcal{X}_{S_b(\mathbf{x})}(\mathbf{y}), \tag{6.25}$$

where $c_{d,s}$ is defined in (6.20)*, the constant C_{ba} from* (6.18) *is given by*

$$C_{ba} = C_{d,s}\delta^{-2s},$$

where $C_{d,s}$, which does not depend on δ, is given as follows.

i) $S_a(\mathbf{x}) = B_{\delta,2}(\mathbf{x})$ *and* $S_b(\mathbf{x}) = \mathbb{R}^d$*:*

$$C_{d,s} = \frac{d}{2s}V_d c_{d,s},$$

where

$V_d = \frac{\sqrt{\pi^d}}{\Gamma\left(\frac{d}{2}+1\right)}$ = *volume of the d-dimensional unit $\|\cdot\|_2$-ball.*

ii) $S_a(\mathbf{x}) = B_{\delta,2}(\mathbf{x})$ *and* $S_b(\mathbf{x}) = B_{\delta,\infty}(\mathbf{x})$*, i.e., the $\|\cdot\|_2$-ball of radius δ is inscribed in the $\|\cdot\|_\infty$-ball having the same radius:*

- $d = 2$,
$$C_{2,s} = \tfrac{1}{s}c_{2,s}\left(\pi - 2\left(\widetilde{B}(s+\tfrac{1}{2},\tfrac{1}{2}) - \widetilde{B}_{\frac{1}{2}}\left(s+\tfrac{1}{2},\tfrac{1}{2}\right)\right)\right),$$
where $\widetilde{B}$ and $\widetilde{B}_{(\cdot)}$ denote the Beta function and the unregularized incomplete Beta function, respectively.
- $d = 3$,
$$C_{3,s} = c_{3,s}\frac{2}{s}\left(\frac{\pi}{1+2s} - p_1 + 2p_2 + \frac{4p_3}{1+2s}\right),$$
with
$$p_1 = \widetilde{B}\left(s+1,\tfrac{1}{2}\right)\left(\widetilde{B}\left(s+\tfrac{1}{2},\tfrac{1}{2}\right) - \widetilde{B}_{\frac{1}{2}}\left(s+\tfrac{1}{2},\tfrac{1}{2}\right)\right)$$
$$p_2 = \int_0^{\pi/4}\cos(\theta)^{2s}\widetilde{B}_{\frac{1}{1+\cos^2(\theta)}}(s+1,\tfrac{1}{2})d\theta$$
$$p_3 = \int_0^{\pi/4}\left(\frac{\cos(\theta)}{\sqrt{\cos^2(\theta)+1}}\right)^{1+2s}d\theta,$$
where p_1, p_2, and p_3 can be computed numerically with a desired accuracy since the integrands are smooth.

iii) $S_a(\mathbf{x}) = B_{\delta,\infty}(\mathbf{x})$ *and* $S_b(\mathbf{x}) = \mathbb{R}^d$*:*

- $d = 2$,
$$C_{2,s} = \frac{2}{s}c_{2,s}\left(\widetilde{B}\left(s+\tfrac{1}{2},\tfrac{1}{2}\right) - \widetilde{B}_{\frac{1}{2}}\left(s+\tfrac{1}{2},\tfrac{1}{2}\right)\right) > 0.$$

– $d = 3$,

$$C_{3,s} = \frac{2}{s} c_{3,s} \left(\frac{\pi}{1+2s} + p_1 - 2p_2 - \frac{4p_3}{1+2s} \right) > 0,$$

with p_i from ii).

Proof. i) Inserting the fractional kernel function into (6.10) leads to

$$C_{ba} = dV_d c_{d,s} \int_\delta^\infty \frac{r^{d-1}}{r^{d+2s}} dr = dV_d c_{d,s} \frac{\delta^{-2s}}{2s}.$$

ii) For $d = 2$, we insert the fractional kernel function into expression (6.11) and obtain

$$\begin{aligned}
C_{ab} &= 8 \int_0^{\pi/4} \int_\delta^{\frac{\delta}{\cos(\theta)}} \phi(r) r dr d\theta \\
&= 8 c_{2,s} \int_0^{\pi/4} \int_\delta^{\frac{\delta}{\cos(\theta)}} \frac{r}{r^{2+2s}} dr d\theta \\
&= 8 c_{2,s} \int_0^{\pi/4} \left(\tfrac{r^{-2s}}{-2s} \right) \Big|_\delta^{\frac{\delta}{\cos(\theta)}} d\theta \\
&= c_{2,s} \frac{4}{s} \left(\tfrac{\pi}{4} \delta^{-2s} - \int_0^{\pi/4} \left(\tfrac{\delta}{\cos(\theta)} \right)^{-2s} d\theta \right) \\
&= c_{2,s} 4 \frac{\delta^{-2s}}{s} \left(\tfrac{\pi}{4} - \int_0^{\pi/4} \cos^{2s}(\theta) d\theta \right).
\end{aligned}$$

We also have that

$$\begin{aligned}
\int_0^{\pi/4} \cos^{2s}(\theta) d\theta &= \int_1^{\sqrt{2}/2} t^{2s} \arccos'(t) dt \\
&= \int_1^{\sqrt{2}/2} \left(t^2 \right)^s \left(1 - t^2 \right)^{0.5} dt \\
&= - \int_1^{\sqrt{2}/2} \left(t^2 \right)^s \left(1 - t^2 \right)^{0.5} dt \\
&= \int_{\sqrt{2}/2}^1 \left(t^2 \right)^s \left(1 - t^2 \right)^{0.5} \frac{2t}{2\sqrt{t^2}} dt \\
&= \tfrac{1}{2} \int_{1/2}^1 t^{s-0.5} (1-t)^{0.5} dt \\
&= \tfrac{1}{2} \int_0^1 t^{s-0.5} (1-t)^{0.5} dt - \int_0^{1/2} t^{s-0.5} (1-t)^{0.5} dt \\
&= \tfrac{1}{2} \left(\widetilde{B}(s + \tfrac{1}{2}, \tfrac{1}{2}) - \widetilde{B}_{\frac{1}{2}}(s + \tfrac{1}{2}, \tfrac{1}{2}) \right).
\end{aligned}$$

Combining results yields the stated constant.

For $d = 3$, we start from (6.12) in order to obtain

$$\begin{aligned}
C_{ba} &= c_{3,s}16 \int_0^{\pi/4} \int_0^{\arctan(\cos(\theta_1))} \int_\delta^{\frac{\delta}{\cos(\theta_1)\cos(\theta_2)}} \frac{1}{2r^{3+2s}} r^2 \cos(\theta_2) dr d\theta_2 d\theta_1 \\
&+ c_{3,s}16 \int_0^{\pi/4} \int_{\arctan(\cos(\theta_1))}^{\pi/2} \int_\delta^{\frac{\delta}{\sin(\theta_2)}} \frac{1}{r^{2s+1}} \cos(\theta_2) dr d\theta_2 d\theta_1 \\
&= c_{3,s}16 \int_0^{\pi/4} \int_0^{\arctan(\cos(\theta_1))} \frac{r^{-2s}}{-2s}\Big|_\delta^{\frac{\delta}{\cos(\theta_1)\cos(\theta_2)}} \cos(\theta_2) d\theta_2 d\theta_1 \\
&+ c_{3,s}16 \int_0^{\pi/4} \int_{\arctan(\cos(\theta_1))}^{\pi/2} \frac{r^{-2s}}{-2s}\Big|_\delta^{\frac{\delta}{\sin(\theta_2)}} \cos(\theta_2) d\theta_2 d\theta_1 \\
&= c_{3,s}16 \tfrac{\delta^{-2s}}{2s} \int_0^{\pi/4} \int_0^{\arctan(\cos(\theta_1))} \left(1 - \cos^{2s}(\theta_1)\cos^{2s}(\theta_2)\right) \cos(\theta_2) d\theta_2 d\theta_1 \\
&+ c_{3,s}16 \tfrac{\delta^{-2s}}{2s} \int_0^{\pi/4} \int_{\arctan(\cos(\theta_1))}^{\pi/2} \left(1 - \sin^{2s}(\theta_2)\right) \cos(\theta_2) d\theta_2 d\theta_1 \\
&= c_{3,s}16 \tfrac{\delta^{-2s}}{2s} \int_0^{\pi/4} \int_0^{\pi/2} \cos(\theta_2) d\theta_2 d\theta_1 \\
&- c_{3,s}16 \tfrac{\delta^{-2s}}{2s} \int_0^{\pi/4} \int_0^{\arctan(\cos(\theta_1))} \cos^{2s}(\theta_1)\cos^{2s+1}(\theta_2) d\theta_2 d\theta_1 \\
&- c_{3,s}16 \tfrac{\delta^{-2s}}{2s} \int_0^{\pi/4} \int_{\arctan(\cos(\theta_1))}^{\pi/2} \sin^{2s}(\theta_2)\cos(\theta_2) d\theta_2 d\theta_1 \\
&= c_{3,s}8 \tfrac{\delta^{-2s}}{s} \left(\tfrac{\pi}{4} - p_1' - p_2'\right).
\end{aligned}$$

For p_1' we first recall from ii) that

$$\int_0^{\pi/4} \cos^{2s}(t) dt = \tfrac{1}{2}\left(\widetilde{B}\left(s + \tfrac{1}{2}, \tfrac{1}{2}\right) - \widetilde{B}_{\frac{1}{2}}\left(s + \tfrac{1}{2}, \tfrac{1}{2}\right)\right)$$

and in a similar way we obtain

$$\int_0^{a} \cos^{2s+1}(t) dt = \tfrac{1}{2}\left(\widetilde{B}\left(s + 1, \tfrac{1}{2}\right) - \widetilde{B}_{\cos^2(a)}\left(s + 1, \tfrac{1}{2}\right)\right).$$

Thus, we find

$$\begin{aligned}
p_1' &= \int_0^{\pi/4} \cos^{2s}(\theta_1) \int_0^{\arctan(\cos(\theta_1))} \cos^{2s+1}(\theta_2) d\theta_2 d\theta_1 \\
&= \int_0^{\pi/4} \cos^{2s}(\theta_1) \left(\tfrac{1}{2}\left(\widetilde{B}\left(s + 1, \tfrac{1}{2}\right) - \widetilde{B}_{\cos^2(\arctan(\cos(\theta_1)))}\left(s + 1, \tfrac{1}{2}\right)\right)\right) d\theta_1 \\
&= \tfrac{1}{2}\widetilde{B}\left(s + 1, \tfrac{1}{2}\right)\left(\tfrac{1}{2}\left(\widetilde{B}\left(s + \tfrac{1}{2}, \tfrac{1}{2}\right) - \widetilde{B}_{\frac{1}{2}}\left(s + \tfrac{1}{2}, \tfrac{1}{2}\right)\right)\right) \\
&- \tfrac{1}{2}\int_0^{\pi/4} \cos^{2s}(\theta_1) \widetilde{B}_{\frac{1}{1+\cos^2}}\left(s + 1, \tfrac{1}{2}\right) d\theta_1
\end{aligned}$$

$$= \frac{p_1}{4} - \frac{p_2}{2}.$$

Now, consider p_2'. We have

$$\begin{aligned}
p_2' &= \int_0^{\pi/4} \int_{\arctan(\cos(\theta_1))}^{\pi/2} \sin^{2s}(\theta_2)\cos(\theta_2) d\theta_2 d\theta_1 \\
&= \int_0^{\pi/4} \frac{\sin^{1+2s}(\theta_2)}{1+2s}\Big|_{\arctan(\cos(\theta_1))}^{\pi/2} d\theta_1 \\
&= \tfrac{1}{1+2s} \int_0^{\pi/4} \left(1 - \sin^{1+2s}(\arctan(\cos(\theta_1)))\right) d\theta_1 \\
&= \tfrac{1}{1+2s} \left(\frac{\pi}{4} - \int_0^{\pi/4} \sin^{1+2s}(\arctan(\cos(\theta_1))) d\theta_1\right) \\
&= \tfrac{1}{1+2s} \left(\frac{\pi}{4} - \int_0^{\pi/4} \left(\frac{\cos(\theta_1)}{\sqrt{\cos(\theta_1)^2+1}}\right)^{1+2s} d\theta_1\right) \\
&= \tfrac{1}{1+2s} \left(\tfrac{\pi}{4} - p_3\right).
\end{aligned}$$

Inserting the last two results into

$$C_{ba} = c_{3,s} 4 \tfrac{\delta^{-2s}}{s} (\tfrac{\pi}{4} - p_1' - p_2')$$

yields

$$C_{ba} = c_{3,s} 8 \tfrac{\delta^{-2s}}{2s} \left(\tfrac{\pi}{4} - p_1' - p_2'\right) = c_{3,s} 8 \tfrac{\delta^{-2s}}{2s} \left(\tfrac{\pi}{4} - \tfrac{p_1}{4} + \tfrac{p_2}{2} - \tfrac{1}{1+2s}\left(\tfrac{\pi}{4} - p_3\right)\right).$$

which immediately leads to the stated result.

iii) Because $B_{\delta,2}(0) \subset B_{\delta,\infty}(0)$, we have

$$\int_{\mathbb{R}^d \setminus B_{\delta,\infty}(0)} \phi(\|\mathbf{z}\|) d\mathbf{z} = \int_{\mathbb{R}^d \setminus B_{\delta,2}(0)} \phi(\|\mathbf{z}\|) d\mathbf{z} - \int_{B_{\delta,\infty}(0) \setminus B_{\delta,2}(0)} \phi(\|\mathbf{z}\|) d\mathbf{z}$$

so that, for $d = 2$,

$$C_{ba} = c_{2,s} \tfrac{\delta^{-2s}}{2s} \pi - c_{2,s} \tfrac{\delta^{-2s}}{2s} \left(\pi - 2\widetilde{B}(s+\tfrac{1}{2}, \tfrac{1}{2}) + 2\widetilde{B}_{\frac{1}{2}}(s+\tfrac{1}{2}, \tfrac{1}{2})\right)$$

which, canceling the common terms, is the stated result. For $d = 3$,

$$C_{ba} = c_{3,s} \tfrac{\delta^{-2s}}{s} \pi - c_{3,s} \tfrac{1}{s} \left(\tfrac{2s\pi}{1+2s} - p_1 + 2p_2 + \tfrac{4p_3}{1+2s}\right)$$

which immediately leads to the stated result. □

We now focus on $\|\cdot\|_\infty$-balls and specialize (6.17) to the two kernels in (6.25) with $S_a(\mathbf{x}) = B_{\delta,\infty}(\mathbf{x})$ and $S_b(\mathbf{x}) = \mathbb{R}^d$ so that $A_b(u,v) = A_{d,s}(u,v)$, where $A_{d,s}(u,v)$ is the bilinear form (6.23) corresponding to the fractional Laplacian operator (6.19).

Proposition 6.2.3. *For $d \in \{2,3\}$ and for $S_a(\mathbf{x}) = B_{\delta,\infty}(\mathbf{x})$ with $\delta \geq \operatorname{diam}(\Omega)$ and $S_b(\mathbf{x}) = \mathbb{R}^d$,*

$$A_{d,s}(u,v) = A_a(u,v) + C_{d,s}(u,v)_{L^2(\Omega)}\delta^{-2s}, \tag{6.26}$$

where the constant $C_{d,s}$ is given in Lemma 6.2.2 part (iii). Furthermore, the spaces $H_c^s(\mathbb{R}^d)$ and $V_c^a(\mathbb{R}^d)$ are equivalent and

$$|u_{d,s}|_\Omega - u_a|_\Omega|_{H^s(\Omega\cup\Omega_I)} \leq C_{d,s}C_P^a\|u_{d,s}\|_{L^2(\Omega)}\delta^{-2s}$$
$$\|u_{d,s}|_\Omega - u_a|_\Omega\|_{L^2(\Omega)} \leq C_{d,s}\left(C_P^a\right)^2\|u_{d,s}\|_{L^2(\Omega)}\delta^{-2s},$$

where C_P^a denotes the uniform Poincaré constant for the space $H_c^s(\Omega\cup\Omega_I)$ independent of δ; see Remark 3.3.8.

Proof. First we point out that in the notation of Section 6.1 and particularly Lemma 6.1.4, we here have that $\Omega\cup\Omega_I^{ab} = \mathbb{R}^d$ and $\Omega\cup\Omega_I^a = \Omega\cup\Omega_I$, so that

$$|||u|||_a = |u|_{H^s(\Omega\cup\Omega_I)} = \int_\Omega\int_\Omega \frac{c_{d,s}\left(u-u'\right)^2}{\|\mathbf{x}-\mathbf{y}\|_2^{d+2s}}d\mathbf{y}d\mathbf{x} + \int_\Omega u^2\int_{\Omega_I}\frac{c_{d,s}}{\|\mathbf{x}-\mathbf{y}\|_2^{d+2s}}d\mathbf{y}d\mathbf{x}.$$

The relation (6.26) follows immediately from (6.17). An immediate consequence of (6.26) is that

$$|A_{d,s}(u,v) - A_a(u,v)| \leq C_{d,s}\|u\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}\delta^{-2s}$$

so that the remaining assertions follow immediately as special cases of Lemma 6.1.3 and Lemma 6.1.4 with $g = 0$. □

Proposition 6.2.3 tells us that weak solutions of the truncated nonlocal problem (3.7) with $S_a(\mathbf{x}) = B_{\delta,\infty}(\mathbf{x})$ converge to the weak solution of the fractional Laplace problem (6.22) with a rate of $2s$ as the extent of nonlocal interactions increases, i.e., as $\delta \to \infty$. The same results are shown for $S_a(\mathbf{x}) = B_{\delta,2}(\mathbf{x})$ in [33].

Remark 6.2.4 (***Implementation recipe***)**.** *Related to Remark 6.2.1, Proposition 6.2.3 and in particular* (6.26) *suggest the following recipe for implementing finite element methods for the fractional Laplacian with homogeneous volume constraints. We choose the smallest δ possible, i.e., $\delta = \operatorname{diam}(\Omega)$, and compute the stiffness matrix of the (truncated) nonlocal diffusion problem. This stiffness matrix is then "corrected" by adding a scalar multiple of the mass matrix $(u,v)_{L^2(\Omega)}$ which is independent of δ and other properties peculiar to the problem such as the fraction s, other kernel properties, etc. This implies that, in practice, we do not have to be concerned with having to deal with a "large" ball in the implementation process. For this purpose it was necessary to compute the precise values for the constant C_{ba} in Lemma 6.2.2.*
We also want to mention that in [1] a related implementation recipe is considered. There, a fixed ball containing the domain Ω is considered and an auxiliary mesh is used for the intersection between this ball and the complement of Ω. Outside of the ball, spherical coordinates are employed.

Remark 6.2.5 (***Extension to inhomogeneous problems and other kernels***)**.**
i) The results of this section can be extended to the case of inhomogeneous volume constraints. In particular, the assembly process for the stiffness matrices for the fractional Laplacian problem and for the truncated nonlocal problem remain unchanged. One issue that does arise for inhomogeneous volume constraints is that the implementation of the second term of the right-hand side (3.4)*, i.e.,*

$$\int_{\Omega} v(\mathbf{x}) \int_{\Omega^c} g(\mathbf{y})\gamma(\mathbf{x},\mathbf{y}) d\mathbf{y} d\mathbf{x}$$

requires computing integrals of the Dirichlet data g over the complement domain. Of course, this issue arises in the implementation of any numerical method for the fractional Laplacian problem with inhomogeneous Dirichlet-type volume constraints. Because g is a given function, the needed integrals can be approximated using suitable quadrature rules and appropriate truncation for the fractional Laplace equation.
ii) The considerations of this section not only hold for the fractional Laplacian problem, but can be extended to other settings such as, e.g., other types of radial kernels; see, e.g., Lemma 6.1.5.

6.2.2 A numerical example

We illustrate the results of Proposition 6.2.3 for the case of $\|\cdot\|_\infty$-balls such that $\Omega \subset B_{\delta,\infty}(\mathbf{x})$, for all $\mathbf{x} \in \Omega$, for a sequence of increasing δ. Specifically, we consider $\Omega = (0,1)^2$, which we uniformly discretize with a fixed grid size $h = 2^{-8}$. Furthermore we choose $s = 0.4$, and

$$\delta_i = 2^i \delta_0, \quad \text{for} \quad i \in \{4, 5, \ldots, 8\}$$

with $\delta_0 := 1.4453125$ so that

$$\delta_0 > \operatorname{diam}(\Omega) = \sqrt{2} \quad \text{and} \quad \frac{\delta_0}{h} \in \mathbb{N}.$$

We use Q_1 elements and exploit the multilevel Toeplitz structure of the stiffness matrix; see Chapter 5. In Table 6.1, $u_{d,s}$ denotes a surrogate for the solution of the fractional problem (6.22) whereas u_a denotes the solution of (3.7) with the kernel $\gamma_a(\mathbf{x},\mathbf{y})$ given in (6.25) with $S_a(\mathbf{x}) = B_{\delta,\infty}(\mathbf{x})$. The results presented in Table 6.1 show that, as predicted by the results in Proposition 6.2.3, the convergence rate of $2s = 0.8$ is achieved for both the L^2-norm and the energy norm.

δ	$\|u_{d,s} - u_a\|_{L^2}$	rate	$\|u_{d,s} - u_a\|_{H^s}$	rate
$2^4\delta_0$	0.019	0.825	0.027	0.826
$2^5\delta_0$	0.011	0.815	0.015	0.814
$2^6\delta_0$	0.006	0.809	0.009	0.808
$2^7\delta_0$	0.003	0.807	0.005	0.804
$2^8\delta_0$	0.002	0.804	0.003	0.802

Table 6.1: For $s = 0.4$, comparisons of a numerical surrogate for the weak solution $u_{d,s}$ of the fractional Laplacian problem (6.22) with finite element approximations of the weak solutions u_a of (3.7) corresponding to finite $\|\cdot\|_\infty$-balls for several interaction horizons δ.

6.3 Application to approximate interaction sets

We now pick up on the discussion of Section 4.4 where we introduce the notion of approximate interaction sets. The motivation is to ease the subdivision task arising along the numerical integration of the inner integral in (4.21) by intentionally allowing for geometric errors, thereby leading to approximations $S^h(\mathbf{x})$ to the exact interaction set $S(\mathbf{x})$.

As in Section 4.2 we assume that $\Omega \cup \Omega_I$ is a polyhedral domain for which we have a triangulation $\mathcal{T}^h$ into finite elements. Let $\left\{V^h\right\}_{h>0}$ denote a sequence of finite-dimensional subspaces $V^h \subset V(\Omega \cup \Omega_I)$ consisting of piecewise polynomials defined with respect to $\mathcal{T}^h$. The numerical quadrature discussed in Section 4.3 for discretizing the nonlocal bilinear form A and the linear functional ℓ leads to approximate bilinear and linear forms

$$A^h : V_c^h(\Omega \cup \Omega_I) \times V_c^h(\Omega \cup \Omega_I) \to \mathbb{R}, \quad \ell^h : V_c^h(\Omega \cup \Omega_I) \to \mathbb{R}. \tag{6.27}$$

We want to discuss the effect of the discretization on the global finite element error with a focus on considering approximate interaction sets. A fundamental result in the finite element theory which proves useful to initiate such an investigation is given by Strang's first lemma, which we recall for convenience (see, e.g., [44, Lemma 2.17]).

Theorem 6.3.1 (***Strang's first lemma***)**.** *Let $(H, \|\cdot\|_H)$ be a Hilbert space and $A\colon H \times H \to \mathbb{R}$ a coercive and continuous bilinear form, not necessarily symmetric, and $\ell \in H^*$. Further we consider finite-dimensional subspaces $H^h \subset H$ and let $A^h\colon H^h \times H^h \to \mathbb{R}$ be uniformly H^h-elliptic, i.e.,*

$$A^h(u^h, u^h) \geq \beta \|u^h\|_H^2 \quad \textit{for all } h, \textit{ and all } u^h \in V_c^h(\Omega \cup \Omega_I),$$

for some $\beta > 0$ independent of h, and let ℓ^h be a linear functional on H^h. Let $u \in H$ be, such that $A(u, v) = \ell(v)$ for all $v \in H$ and analogously $u^h \in H^h$ be, such that

$A^h(u^h, v^h) = \ell^h(v^h)$ *for all* $v \in H^h$. *Then there exists a constant* $C_S > 0$ *independent of* h, *such that*

$$||u - u^h||_H \leq C_S \left(\inf_{v^h \in H^h} \left\{ \|u - v^h\|_H + \sup_{w^h \in H^h} \frac{\left|(A - A^h)(v^h, w^h)\right|}{\|w^h\|_H} \right\} + \sup_{w^h \in H^h} \frac{\left|(\ell - \ell^h)(w^h)\right|}{\|w^h\|_H} \right). \tag{6.28}$$

Strang's first lemma derives an upper bound for the global discretization error, which consists of two parts. On the one hand we have the so-called *approximation or interpolation error*

$$I(h) := \inf_{v^h \in H^h} \|u - v^h\|_H, \tag{6.29}$$

which is affected by the choice of the ansatz space V^h, and on the other hand we have the so-called *consistency errors*

$$\sup_{w^h \in H^h} \frac{\left|(A - A^h)(v^h, w^h)\right|}{\|w^h\|_H} \qquad \text{and} \qquad \sup_{w^h \in H^h} \frac{\left|(\ell - \ell^h)(w^h)\right|}{\|w^h\|_H},$$

which can be controlled by carefully chosen quadrature rules.

Let us cast the setting of the latter theorem into our framework. In Chapter 3 we have considered a class of integrable kernels for which the nonlocal bilinear form A is coercive and continuous on $\left(L_c(\Omega \cup \Omega_I), \|\cdot\|_{L^2(\Omega \cup \Omega_I)}\right)$ and also a class of singular kernels for which A shares these properties on the fractional Sobolev space $\left(H^s_c(\Omega \cup \Omega_I), |\cdot|_{H^s(\Omega \cup \Omega_I)}\right)$. Thus, these spaces are natural choices for $(H, \|\cdot\|_H)$ in Strang's first lemma. Furthermore we have deduced that these spaces are equivalent to the nonlocal constrained energy space $\left(V_c(\Omega \cup \Omega_I), |||\cdot||| = \sqrt{A(\cdot,\cdot)}\right)$, so that this can also be considered a candidate space for H. In the following we keep $H = H(\Omega \cup \Omega_I)$ as placeholder for the spaces just mentioned. For the finite-dimensional subspaces however we have considered the same piecewise polynomial finite element basis functions in each case, so that we can consider $H^h = V^h$.

For simplicity, we now assume that the quadrature rules which we employ to arrive at the discretized forms A^h and ℓ^h are exact for the integrands appearing in the respective integrals over the finite elements. For example, for a kernel γ and a forcing term f which are smooth within each element, we can choose a standard Gaussian quadrature rule of sufficiently high order; note that we are considering piecewise polynomial basis functions. Thus, under this assumption, we find that

$$\sup_{w^h \in H^h} \frac{\left|(\ell - \ell^h)(w^h)\right|}{\|w^h\|_H} = 0$$

and consequently, the consistency error is only effected by

$$\sup_{w^h \in H^h} \frac{\left|(A - A^h)(v^h, w^h)\right|}{\|w^h\|_H}$$

which arises due to considering geometric approximations to the interaction sets. In the view of Section 6.1, to compare A and A^h we can consider two kernels, the original kernel

$$\gamma_a(\mathbf{x}, \mathbf{y}) := \gamma(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x}, \mathbf{y})\chi_{S(\mathbf{x})}(\mathbf{y})$$

and the kernel involving the approximate interaction set, i.e.,

$$\gamma_b(\mathbf{x}, \mathbf{y}) := \gamma^h(\mathbf{x}, \mathbf{y}) := \phi(\mathbf{x}, \mathbf{y})\chi_{S^h(\mathbf{x})}(\mathbf{y}).$$

The result presented in Corollary 6.1.1 then enables us to derive an estimate for the difference in bilinear forms $(A - A^h)(v^h, w^h)$ appearing in (6.28). Since the two kernels solely differ in the choice of interaction sets, we are in the case of Remark 6.1.2 i), and can carry over the results as follows. If

$$C_{ba} = C_{ba}(h) = \max\left\{\sup_{\mathbf{x}\in\Omega}\int_{S(\mathbf{x})\triangle S^h(\mathbf{x})} \phi d\mathbf{y},\ \sup_{\mathbf{x}\in\Omega}\int_{S(\mathbf{x})\triangle S^h(\mathbf{x})} \phi' d\mathbf{y}\right\} < \infty, \tag{6.30}$$

then by invoking Corollary 6.1.1 we find

$$\left|(A - A^h)(u, v)\right| \leq 2C_{ab}(h)\|u\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} \tag{6.31}$$

for all $u, v \in L^2_c(\Omega \cup \Omega^{ab}_I)$.

We now combine Strang's first lemma with (6.31) in order to arrive at the following bound for the global finite element error in the setting just discussed.

Corollary 6.3.2 (***Approximate balls***)**.** *Let $(H(\Omega \cup \Omega_I), \|\cdot\|_H)$ be a Hilbert space with $\|\cdot\|_H$ satisfying a Poincaré type inequality*

$$\|u\|_H \geq C_P\|u\|_{L^2(\Omega)} \quad \textit{for all } u \in L^2_c(\Omega \cup \Omega_I). \tag{6.32}$$

We assume that the nonlocal bilinear form A is coercive and continuous on H and for a forcing term f and Dirichlet data g let $u \in H$ denote the unique solution of the weak problem (3.7) *posed on H. Further let A^h and ℓ^h be as in* (6.27)*. We assume that A^h is uniformly V^h_c-elliptic and for ℓ^h let u^h denote the solution of the discrete problem* (4.6)*. Finally we require $C_{ba}(h) < \infty$ with $C_{ba}(h)$ from* (6.30)*. Then it follows that*

$$\|u - u^h\|_H \leq C_S\left(1 + 2C_{ba}(h)C_P^2\right) I(h) + 2C_S C_{ba}(h)C_P^2\|u\|_H,$$

where $C_S > 0$ is the constant from Strang's Lemma and $I(h)$ the interpolation error defined in (6.29)*. Applying* (6.32) *also yields*

$$\|u - u^h\|_{L^2(\Omega)} \leq C_S\left(\frac{1}{C_P} + C_{ba}(h)C_P\right) I(h) + C_S C_{ba}(h)C_P\|u\|_H.$$

Proof. Since the requirements for Strang's first lemma (6.28) are fulfilled, we find

$$||u-u^h||_H \leq C_S \left(\inf_{v^h \in V_c^h} \left\{ ||u-v^h||_H + \sup_{w^h \in V_c^h} \frac{\left|(A-A^h)(v^h,w^h)\right|}{\|w^h\|_H} \right\} + \sup_{w^h \in V_c^h} \frac{\left|(\ell-\ell^h)(w^h)\right|}{\|w^h\|_H} \right).$$

Since we assume perfect quadrature, the last term vanishes identically, i.e.,

$$\sup_{w^h \in V_c^h} \frac{\left|(\ell-\ell^h)(w^h)\right|}{\|w^h\|_H} = 0.$$

Invoking (6.31) and (6.32) we find

$$\begin{aligned} \sup_{w^h \in V_c^h} \frac{\left|(A-A^h)(v^h,w^h)\right|}{\|w^h\|_H} &\leq 2C_{ba}(h)C_P^2 \|v^h\|_H \\ &\leq 2C_{ba}(h)C_P^2 \left(\|u-v^h\|_H + \|u\|_H \right), \end{aligned}$$

where the second inequality follows from triangle inequality on H. We finally obtain the desired estimate

$$\begin{aligned} \|u-u^h\|_H &\leq C_S \inf_{v^h \in V_c^h} \left\{ \|u-v^h\|_H + 2C_{ba}(h)C_P^2 \left(\|u-v^h\|_H + \|u\|_H \right) \right\} \\ &= C_S \left(1 + 2C_{ba}(h)C_P^2 \right) \left(\inf_{v^h \in V_c^h} \|u-v^h\|_H \right) + 2C_S C_{ba}(h) C_P^2 \|u\|_H. \end{aligned}$$

□

Hence, we can conclude that the finite element error is dominated by either the interpolation error $I(h)$ or the geometric error $C_{ab}(h)$. In the next subsection we concretize these thoughts and draw conclusions for specific examples.

6.3.1 Numerical examples for different approximate balls

We now specify to norm induced interaction sets $S(\mathbf{x}) = B_{\delta,\bullet}(\mathbf{x})$ for a norm $\|\cdot\|_\bullet$ and a fixed interaction horizon $\delta > 0$. We aim to define *approximate balls* $S^h(\mathbf{x}) = B^h_{\delta,\bullet}(\mathbf{x})$ which ease the quadrature of the inner integral in (4.21). In the following we only discuss approaches which completely circumvent the subdivision task discussed in Section 4.3.2.

In restricting to norm induced balls we find that the constant of interest given in (6.30) becomes

$$C_{ab}(h) = \max \left\{ \sup_{\mathbf{x}\in\Omega} \int_{B_{\delta,\bullet}(\mathbf{x}) \triangle B^h_{\delta,\bullet}(\mathbf{x})} \phi d\mathbf{y}, \quad \sup_{\mathbf{x}\in\Omega} \int_{B_{\delta,\bullet}(\mathbf{x}) \triangle B^h_{\delta,\bullet}(\mathbf{x})} \phi' d\mathbf{y} \right\} < \infty.$$

We assume that $\delta > h$ and

$$B_{\delta,\bullet}(\mathbf{x}) \triangle B^h_{\delta,\bullet}(\mathbf{x}) \subset B_{(\delta+h),\bullet}(\mathbf{x}) \backslash B_{(\delta-h),\bullet}(\mathbf{x}). \tag{6.33}$$

The latter requirement (6.33) is fulfilled by the examples of approximate balls $B^h_{\delta,\bullet}(\mathbf{x})$ considered in the numerical study below. Now let us estimate $C_{ab}(h)$ in oder to determine the order of the consistency error. By (6.33) we first find that

$$C_{ab}(h) \leq \left(\sup_{\mathbf{x}\in\Omega} \sup_{\mathbf{y}\in B_{\delta,\bullet}(\mathbf{x}) \triangle B^h_{\delta,\bullet}(\mathbf{x})} \max\{\phi, \phi'\} \right) \left| B_{(\delta+h),\bullet}(\mathbf{0}) \backslash B_{(\delta-h),\bullet}(\mathbf{0}) \right|. \tag{6.34}$$

Clearly, the second factor satisfies

$$\left| B_{(\delta+h),\bullet}(\mathbf{0}) \backslash B_{(\delta-h),\bullet}(\mathbf{0}) \right| \leq Mh$$

for some constant $M = M(d, \delta, \bullet) > 0$. Hence, it remains to analyze how the first factor in (6.34) depends on h, which relies on the choice of the kernel function ϕ. For this purpose we restrict ourselves to radial, thereby symmetric, and monotonically decreasing kernel functions

$$\phi(\mathbf{x}, \mathbf{y}) = \phi(\|\mathbf{x} - \mathbf{y}\|_\sharp). \tag{6.35}$$

Typical examples are constant kernel functions $\phi(\mathbf{x}, \mathbf{y}) = const$, the truncated fractional kernel function $\phi(\mathbf{x}, \mathbf{y}) = \frac{c_{d,s}}{\|\mathbf{x}-\mathbf{y}\|_2^{d+2s}}$, the peridynamics kernel function $\phi(\mathbf{x}, \mathbf{y}) = \frac{1}{\|\mathbf{x}-\mathbf{y}\|_2}$ as well as Gaussian type kernel functions of the form $\phi(\mathbf{x}, \mathbf{y}) = \exp\left(-c\frac{\|\mathbf{x}-\mathbf{y}\|_\sharp}{\delta^2}\right)$ all of which are ubiquitous in the relevant literature; see, e.g., [36, 78, 27, 33] and the references therein. Furthermore, let $h_0 \in (0, \delta)$ be a maximum grid size such that $h \in (0, h_0]$. Then, due to the equivalence of norms in $\mathbb{R}^d$, we find a constant $K = K(d, \bullet, \sharp) > 0$ independent of h such that for $\mathbf{y} \in B_{(\delta+h),\bullet}(\mathbf{x}) \backslash B_{(\delta-h),\bullet}(\mathbf{x})$ we have

$$\|\mathbf{x} - \mathbf{y}\|_\sharp \geq K\|\mathbf{x} - \mathbf{y}\|_\bullet \geq K(\delta - h) \geq K(\delta - h_0).$$

Thus, since $\phi(\mathbf{x}, \mathbf{y}) = \phi(\|\mathbf{x} - \mathbf{y}\|_\sharp)$ is assumed to be monotonically decreasing we find

$$\sup_{\mathbf{y}\in B_{(\delta+h),\bullet}(\mathbf{x}) \backslash B_{(\delta-h),\bullet}(\mathbf{x})} \phi(\mathbf{x}, \mathbf{y}) \leq \phi(K(\delta - h_0))$$

which is independent of the grid size $h \in (0, h_0]$. All in all we can conclude that there exists a constant $L = L(\delta, d, \bullet, \sharp) > 0$ independent of $h \in (0, h_0]$, such that

$$C_{ba}(h) \leq Lh$$

and by Corollary 6.3.2 we can at most expect a first-order convergence for kernel functions of type (6.35) and approximate interaction sets satisfying (6.33).

We now present three different preliminary approaches to approximate the exact ball. All of them satisfy assumption (6.33) and avoid the subdivision task for the inner integral which we have detailedly discussed in Section 4.3.2. In Figure 6.7 these examples are depicted. Let in the following $\mathbf{x} \in \mathcal{E} \in \mathcal{T}^h$.

- **Example 1:** We integrate over all elements $\mathcal{E}'$ which are intersected by the exact ball, i.e.,
$$B^h_{\delta,\bullet}(\mathbf{x}) := \bigcup \left\{ \mathcal{E}' \in \mathcal{T}^h \colon B^h_{\delta,\bullet}(\mathbf{x}) \cap \mathcal{E}' \neq \emptyset \right\}.$$
- **Example 2:** We integrate over all elements $\mathcal{E}'$ whose barycenter $\mathbf{x}^{bary}_{\mathcal{E}'}$ lies in the exact ball, i.e.,
$$B^h_{\delta,\bullet}(\mathbf{x}) := \bigcup \left\{ \mathcal{E}' \in \mathcal{T}^h \colon \mathbf{x}^{bary}_{\mathcal{E}'} \in B^h_{\delta,\bullet}(\mathbf{x}) \right\}.$$
- **Example 3**: We integrate over all elements $\mathcal{E}'$ whose barycenter is at most a distance of δ away from the barycenter of $\mathcal{E}$, i.e.,
$$B^h_{\delta,\bullet}(\mathbf{x}) := B^h_{\delta,\bullet} := \bigcup \left\{ \mathcal{E}' \in \mathcal{T}^h \colon \|\mathbf{x}^{bary}_{\mathcal{E}} - \mathbf{x}^{bary}_{\mathcal{E}'}\|_\bullet < \delta \right\}.$$

We want to point out that the third example allows the use of vectorized quadrature, since we consider the same ball $B^h_{\delta,\bullet}$ for all $\mathbf{x} \in \mathcal{E}$.

In the numerical experiment we use the following problem setting. We consider problem (3.7) on $\Omega = (0,1)^2$, with homogeneous Dirichlet data $g = 0$, and with the following kernel, which is truncated by $\|\cdot\|_\infty$-balls, given by

$$\gamma(\mathbf{x},\mathbf{y}) = c_\delta \chi_{B_{\delta,\infty}(\mathbf{x})}(\mathbf{y}), \quad \text{where } \delta = 0.1,\ c_\delta = \frac{3}{4\delta^4}.$$

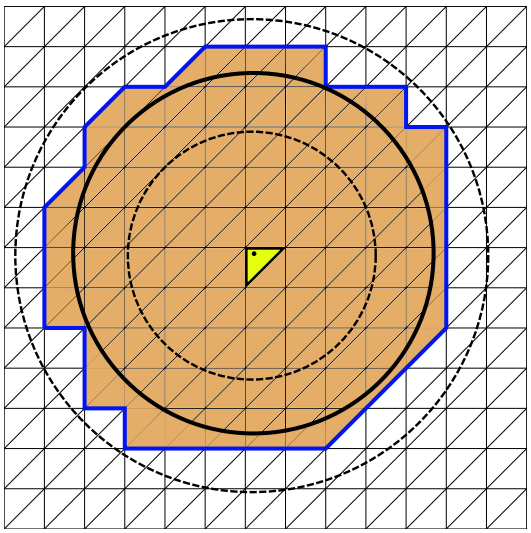
For computing convergence rates as $h \to 0$, we construct a manufactured solution. Therefore, we consider $u(\mathbf{x}) := p(x_1)p(x_2)$ for some function $p : \mathbb{R} \to \mathbb{R}$ with anti-derivative $P : \mathbb{R} \to \mathbb{R}$ and compute the corresponding right-hand side $f(\mathbf{x}) := -\mathcal{L}u(\mathbf{x})$. For homogeneous volume constraints, i.e., $g = 0$, and $\mathbf{x} = (x_1, x_2) \in \Omega$ we find

$$\begin{aligned}
f(\mathbf{x}) =& \int_{B_{\delta,\infty}(\mathbf{x})} (u(\mathbf{x}) - u(\mathbf{y}))\, \gamma(\mathbf{x},\mathbf{y}) d\mathbf{y} \\
=& c_\delta u(\mathbf{x}) |B_{\delta,\infty}(\mathbf{x})| - c_\delta \int_{B_{\delta,\infty}(\mathbf{x})} u(\mathbf{y}) d\mathbf{y} \\
=& c_\delta u(\mathbf{x}) 4\delta^2 - c_\delta \int_{\max(x_1-\delta,0)}^{\min(x_1+\delta,1)} \int_{\max(x_2-\delta,0)}^{\min(x_2+\delta,1)} p(y_1)p(y_2) dy_2 dy_1 \\
=& c_\delta \Big(u(\mathbf{x}) 4\delta^2 - \big(P(\min(x_1+\delta,1)) - P(\max(x_1-\delta,0))\big) \\
& \qquad \cdot \big(P(\min(x_2+\delta,1)) - P(\max(x_2-\delta,0))\big) \Big).
\end{aligned}$$

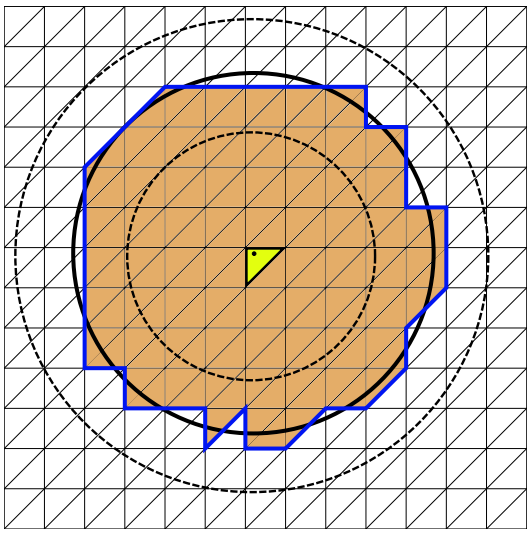
Now we set $p(x) := \sin(2\pi x)$, such that $P(x) = -\frac{1}{2\pi}\cos(2\pi x)$ and therefore

$$u(\mathbf{x}) = \begin{cases} \sin(2\pi x_1)\sin(2\pi x_2) & : \mathbf{x} \in \Omega \\ 0 & : \mathbf{x} \in \Omega_I. \end{cases}$$

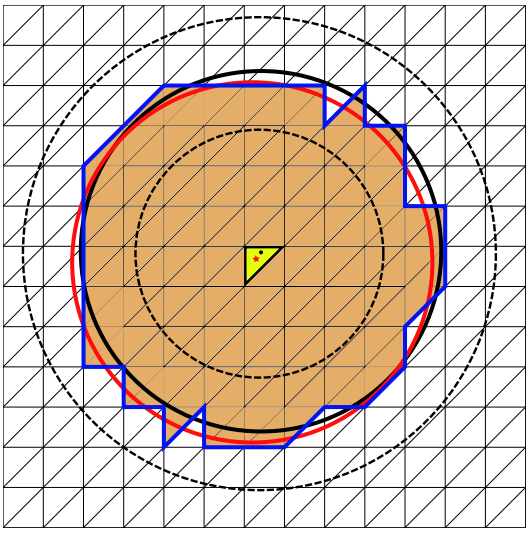
We assemble the stiffness matrix with the exact $\|\cdot\|_\infty$-ball by executing the re-triangulating task for the inner integral, and with the three examples of approximate



Example 1



Example 2



Example 3

Figure 6.7: The three examples of approximate balls considered in the numerical experiments of Table 6.2. In all images the thick solid black line indicates the exact interaction set, here $S(\mathbf{x}) = B_{\delta,2}(\mathbf{x})$, centered at the black dot $\mathbf{x}$ in the yellow triangle $\mathcal{E}$. The dashed lines indicate the enclosing balls $B_{(\delta-h),2}(\mathbf{x})$ and $B_{(\delta+h),2}(\mathbf{x})$ from (6.33). The respective approximate interaction sets $B^h_{\delta,2}(\mathbf{x})$ are represented by the blue lines. All three approaches are based on the underlying grid and contain whole triangles, which are colored orange. In addition to that, the red line in the bottom image of Example 3 indicates the ball $B_{\delta,2}(\mathbf{x}^{bary})$ centered at the red star-shaped barycenter $\mathbf{x}^{bary}$ of the yellow triangle. In this case, the approximate ball is the same for all points in the yellow triangle.

	Exact		**Example 1**		**Example 2**		**Example 3**	
h	$\|\|\|\cdot\|\|\|$	rate	$\|\|\|\cdot\|\|\|$	rate	$\|\|\|\cdot\|\|\|$	rate	$\|\|\|\cdot\|\|\|$	rate
h_1	0.291	1.91	3.184	0.46	0.375	0.89	1.107	0.86
h_2	0.077	1.94	2.319	0.65	0.202	0.87	0.610	1.17
h_3	0.020	1.96	1.481	0.79	0.110	0.92	0.271	1.03
h_4	0.005	1.91	0.857	0.88	0.058	0.94	0.133	1.00
h_5	0.001	-	0.464	-	0.030	-	0.067	-

Table 6.2: We consider grids of size $h_i = 0.2 \cdot 2^{-i}$, $1 \leq i \leq 5$ and report on the convergence rates with respect to the given exact solution measured in the energy norm $|||\cdot|||$.

balls introduced above. The results are presented in Table 6.2. As predicted by the theory, we find that the use of first-order ball approximations deteriorates the nearly second-order convergence rates observed of the exact ball. However, the assembly times for the approximate balls were up to three times faster than the assembly time needed for the exact ball.

6.3.2 Concluding remarks and future work

Approximate interaction sets represent a promising approach to ease finite element implementations for nonlocal models involving truncated kernels; particularly truncations induced by the Euclidean ball. The examples presented are all based on "whole triangle"-approaches and only lead to first-order approximations. Thus, in future work other types which might lead to higher order geometric approximations need to be considered and investigated more in depth. For instance, let us consider Euclidean balls in 2d. Yet another geometric approximation would arise by simply neglecting the circular segments or by considering a re-triangulation thereof into several polyhedral subelements. This should lead to a better, but still cost saving, approximation and is worth being analyzed. In this regard, also a rigorous analysis of the error arising from the quadrature rules as well as the validness of the uniform V_c^h-ellipticity has to be conducted. All in all, this section serves as the basis for further considerations and the idea of approximate interaction sets may potentially contribute to the realization of tractable 3d finite element implementations for nonlocal models.

Chapter 7

Vanishing nonlocality for general norm induced balls

As in Section 6.2, we treat the size of the interaction set as a varying parameter and therefore specialize to norm induced balls in this chapter. Opposed to the infinite limit $\delta \to \infty$ considered in the aforementioned section, we now consider the local limit as the horizon vanishes, i.e., $\delta \to 0$. In this case we asymptotically change the nature of the model from a nonlocal to a local one and different analytical tools have to be consulted. It is a well known fact [85, 38, 37, 84, 10], that the nonlocal weak solution corresponding to an appropriately scaled $\|\cdot\|_2$-radial kernel with bounded second moment converges as the interaction horizon vanishes to the weak solution of the classical Poisson equation (4.29) involving the Laplace operator. We extend this result to kernels which are radial with respect to any norm in $\mathbb{R}^d$ and truncated by any norm induced ball, but satisfy an analogue scaling condition.

After setting up notation in Section 7.1, we establish necessary convergence results for the nonlocal operator and the nonlocal bilinear form in Section 7.2. Then in Section 7.3 we estimate the nonlocal energy norm associated with a kernel satisfying our more general setting against the energy norm of the standard setting, which is crucial for the analysis presented in this chapter. In fact, combining this estimate with the results of the previous section brings us into a position to invoke established results from the standard case. Finally, in Section 7.4 we provide the exact scaling constants for several example kernels which we then use in Section 7.5 for numerical tests confirming the theoretical results of this chapter.

7.1 Preliminaries

For small horizon, i.e., $\delta \ll 1$, and for the local limit the case of Euclidean interaction balls in combination with radial kernel functions $\phi(\mathbf{x}, \mathbf{y}) = \phi(\|\mathbf{y}-\mathbf{x}\|_2)$ has been studied, e.g., in [85], where it is shown that if

$$\int_{B_{\delta,2}(\mathbf{0})} z_1^2 \phi(\|\mathbf{z}\|_2) d\mathbf{z} = C < \infty,$$

then for sufficiently smooth u, v that vanish on the interaction domain Ω_I we find

$$A(u,v) = \int\limits_{\Omega\cup\Omega_I} \int\limits_{\Omega\cup\Omega_I\cap B_{\delta,2}(\mathbf{x})} \big(u(\mathbf{x})-u(\mathbf{y})\big)\big(v(\mathbf{x})-v(\mathbf{y})\big)\phi(\|\mathbf{y}-\mathbf{x}\|_2)d\mathbf{y}d\mathbf{x}$$
$$\xrightarrow[\delta\to 0]{} C\int_\Omega \nabla u(\mathbf{x})\cdot\nabla v(\mathbf{x})d\mathbf{x} = CA^{loc}(u,v)$$

and, in addition, that weak solutions of (3.7) with $g = 0$ converge, as $\delta \to 0$, to weak solutions of a corresponding (local) partial differential equation.

We proceed along the lines of the aforementioned papers to show that these results hold independently of the norm $\|\cdot\|_\bullet$ chosen for the ball or the norm $\|\cdot\|_\sharp$ chosen for the kernel function, i.e., they hold for kernels (K2) of the form

$$\gamma_{\delta,\bullet,\sharp}(\mathbf{x},\mathbf{y}) = \phi(\|\mathbf{x}-\mathbf{y}\|_\sharp)\chi_{B_{\delta,\bullet}(\mathbf{x})}, \tag{7.1}$$

where the ball is defined as $B_{\delta,\bullet}(\mathbf{x}) := \{\mathbf{y}\in\mathbb{R}^d : \|\mathbf{y}-\mathbf{x}\|_\bullet < \delta\}$ for $\mathbf{x}\in\mathbb{R}^d$ and some norm $\|\cdot\|_\bullet$ and $\phi(\mathbf{x},\mathbf{y}) = \phi(\|\mathbf{x}-\mathbf{y}\|_\sharp)$ is radial with respect to some norm $\|\cdot\|_\sharp$. We define

$$\begin{aligned}
-\mathcal{L}_{\delta,\bullet,\sharp}u(\mathbf{x}) &:= 2\int_{(\Omega\cup\Omega_I)\cap B_{\delta,\bullet}(\mathbf{x})} \big(u(\mathbf{x})-u(\mathbf{y})\big)\phi(\|\mathbf{y}-\mathbf{x}\|_\sharp)d\mathbf{y},\\
A_{\delta,\bullet,\sharp}(u,v) &:= \int_{\Omega\cup\Omega_I}\int_{(\Omega\cup\Omega_I)\cap B_{\delta,\bullet}(\mathbf{x})} \big(u(\mathbf{x})-u(\mathbf{y})\big)\big(v(\mathbf{x})-v(\mathbf{y})\big)\phi(\|\mathbf{y}-\mathbf{x}\|_\sharp)d\mathbf{y}d\mathbf{x},\\
K_{\delta,\bullet,\sharp} &:= \int_{B_{\delta,\bullet}(\mathbf{0})} z_1^2\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z},\\
A_0(u,v) &:= \int_\Omega \nabla u(\mathbf{x})\cdot\nabla v(\mathbf{x})d\mathbf{x}.
\end{aligned} \tag{7.2}$$

For simplicity, in the rest of this chapter we omit the subscript $\bullet,\sharp$ and abbreviate the objects above by γ_δ, $\mathcal{L}_\delta$, A_δ, and K_δ.

Remark 7.1.1. *Without loss of generality we scale the nonlocal operator $\mathcal{L}$ by a factor of 2 in (7.2). This cancels the factor of $\frac{1}{2}$ appearing in the Taylor expansion up to the second derivative. Furthermore, we consider homogeneous Dirichlet data $g = 0$ in this chapter. Since we also consider radial and thereby symmetric kernels, we find that the nonlocal bilinear form defined in (3.5) can be written as (3.10) which coincides with the definition in given in (7.2) considering the scaled operator.*

The following lemma proves useful in the other proofs of this chapter.

Lemma 7.1.2.
i) Let $h\colon \mathbb{R}^d \to \mathbb{R}$ denote an odd function, i.e., $h(-\mathbf{z}) = -h(\mathbf{z})$. Then,

$$\int_S h(\mathbf{z})d\mathbf{z} = 0.$$

for all $S \subset \mathbb{R}^d$ with $S = -S$, e.g., $S = B_{\delta,\bullet}(\mathbf{0})$.

ii) Let $\gamma_\delta(\mathbf{x}, \mathbf{y})$ be given as in (7.1). Then, with $\mathbf{z} = \mathbf{y} - \mathbf{x}$, we have

$$\int_{B_{\delta,\bullet}(\mathbf{x})} (y_i - x_i)^2 \, \gamma_\delta(\mathbf{x}, \mathbf{y}) d\mathbf{z} = \int_{B_{\delta,\bullet}(\mathbf{0})} z_1^2 \, \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} \qquad \text{for } i = 1, \dots, d$$

$$\text{and} \quad \int_{B_{\delta,\bullet}(\mathbf{0})} z_i z_k \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} = 0 \qquad \text{for } i, k = 1, \dots, d, \; k \neq i.$$

iii) For the scaling constant K_δ defined in (7.2) we have

$$K_\delta{}^{-1} \int_{B_{\delta,\bullet}(\mathbf{0})} z_1^2 \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} = 1, \qquad K_\delta{}^{-1} \int_{B_{\delta,\bullet}(\mathbf{0})} \mathbf{z}^T \mathbf{z} \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} = d$$

$$\text{and} \qquad K_\delta{}^{-1} \int_{B_{\delta,\bullet}(\mathbf{0})} \mathbf{z}\mathbf{z}^T \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} = \mathbf{Id}\,.$$

Proof.
i) We have

$$\int_S h(\mathbf{z}) d\mathbf{z} = \int_S h(-\mathbf{z}) |(-1)^d| d\mathbf{z} = -\int_S h(\mathbf{z}) d\mathbf{z} = 0.$$

ii) The first result follows by considering the transformation $\mathbf{S}_i(\mathbf{z})$ which swaps the first and i-th component of the vector $\mathbf{z}$ and for which $\mathbf{S}_i(B_{\delta,\bullet}(\mathbf{0})) = B_{\delta,\bullet}(\mathbf{0})$, $\|\mathbf{S}_i(\mathbf{z})\|_\sharp = \|\mathbf{z}\|_\sharp$ and $|\det d\mathbf{S}_i(\mathbf{z})| \equiv 1$. The second follows by considering the transformation $\mathbf{R}_i(\mathbf{z}) := (z_1, ..., -z_i, ..., z_d)$ for which $\mathbf{R}_i(B_{\delta,\bullet}(\mathbf{0})) = B_{\delta,\bullet}(\mathbf{0})$, $\|\mathbf{R}_i(\mathbf{z})\|_\sharp = \|\mathbf{z}\|_\sharp$ and $|\det d\mathbf{R}_i(\mathbf{z})| \equiv 1$.
iii) Note that

$$\int_{B_{\delta,\bullet}(\mathbf{0})} \mathbf{z}^T \mathbf{z} \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} = \sum_{1 \leq i \leq d} \int_{B_{\delta,\bullet}(\mathbf{0})} z_i^2 \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z}$$

and

$$\int_{B_{\delta,\bullet}(\mathbf{0})} \mathbf{z}\mathbf{z}^T \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} = \left(\int_{B_{\delta,\bullet}(\mathbf{0})} z_i z_j \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} \right)_{1 \leq i \leq d, 1 \leq j \leq d}.$$

Thus, the three equalities are immediate consequences of ii). □

7.2 Convergence results for operators

We have the following result concerning the convergence of the nonlocal operator and nonlocal bilinear form to their local analogs. The proofs of i) and ii) are an immediate consequence of Taylor expansion and the scaling requirement, whereas iii) is closely related to the considerations in [38, Section 3.3] for the standard setting.

Proposition 7.2.1. *Let the kernel $\gamma_\delta(\mathbf{x},\mathbf{y})$ be as in (7.1). Then,*

i) $K_\delta^{-1}\mathcal{L}_\delta u(\mathbf{x}) = \Delta u(\mathbf{x}) + \mathcal{O}(\delta^2)$
for all $\mathbf{x}\in\Omega$ and $u\in C^\infty(\Omega\cup\Omega_I)$ with support in Ω;

ii) $K_\delta^{-1}A_\delta(u,v) = A_0(u,v) + \mathcal{O}(\delta^2)$
for all $u,v\in C^\infty(\Omega\cup\Omega_I)$ with support in Ω;

iii) $|K_\delta^{-1}A_\delta(u,v) - A_0(u,v)| = \mathcal{O}(\delta)$

where for (iii), $\mathcal{T}^h=\{\mathcal{E}_j\}_j$ refers to a triangulation of $\Omega\cup\Omega_I$ where we assume $v,u|_{\mathcal{E}_j}\in C^q(\mathcal{E}_j)$ for $\mathcal{E}_j\in\mathcal{T}^h$ with $q\geq 2$ and $u,v\in C^0(\Omega\cup\Omega_I)\cap H^1(\Omega\cup\Omega_I)$ with support in Ω. □

Proof. i) For any $\mathbf{x}\in\Omega$ we have

$$-\mathcal{L}_\delta u(\mathbf{x}) = 2\int_{B_{\delta,\bullet}(\mathbf{0})}(u(\mathbf{x})-u(\mathbf{x}+\mathbf{z}))\phi(\|\mathbf{y}\|_\sharp)d\mathbf{y}.$$

Because u is assumed to be smooth, we can approximate the difference $(u(\mathbf{x})-u(\mathbf{x}+\mathbf{z}))$ by the Taylor polynomial

$$-(u(\mathbf{x})-u(\mathbf{x}+\mathbf{z})) = \sum_{|\boldsymbol{\alpha}|\leq n}\frac{\mathbf{z}^{\boldsymbol{\alpha}}}{\boldsymbol{\alpha}!}d^{\boldsymbol{\alpha}}u(\mathbf{x})+\mathcal{O}(\delta^n),$$

where $\boldsymbol{\alpha}$ denotes a multi-index and $n\in\mathbb{N}$. We then obtain

$$-\mathcal{L}_\delta u(\mathbf{x}) = -2\sum_{|\boldsymbol{\alpha}|\leq n}\int_{B_{\delta,\bullet}(\mathbf{0})}\left(\frac{\mathbf{z}^{\boldsymbol{\alpha}}}{\boldsymbol{\alpha}!}d^{\boldsymbol{\alpha}}u(\mathbf{x})+\mathcal{O}(\delta^n)\right)\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z}.$$

Recognizing that all polynomials $\mathbf{z}^{\boldsymbol{\alpha}}$ for odd $|\boldsymbol{\alpha}|$ are odd functions, we find that their product with the kernel vanish identically under the integral by statement i) in Lemma 7.1.2. Also, for the kernel $\phi(\|\mathbf{z}\|_\sharp)$,

$$-2\mathcal{O}(\delta^n)\int_{B_{\delta,\bullet}(\mathbf{0})}\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z} = -2K_\delta\mathcal{O}(\delta^n)\mathcal{O}(\delta^{-2}) = K_\delta\mathcal{O}(\delta^{n-2}).$$

For $|\boldsymbol{\alpha}|=2$, we obtain from parts i) and iii) of Lemma 7.1.2 that

$$\begin{aligned}-\sum_{i=1}^d(H_u(\mathbf{x}))_{ii}&\int_{B_{\delta,\bullet}(\mathbf{0})}z_i^2\ \phi(\|\mathbf{z}\|_\sharp)d\mathbf{z}\\ &=\left(\int_{B_{\delta,\bullet}(\mathbf{0})}z_1^2\ \phi(\|\mathbf{z}\|_\sharp)d\mathbf{z}\right)\sum_{i=1}^d-(H_u(\mathbf{x}))_{ii} = -K_\delta\Delta u(\mathbf{x}),\end{aligned}$$

noting that all first order terms in $\{\mathbf{z}^{\boldsymbol{\alpha}}:|\boldsymbol{\alpha}|=2\}$ are odd functions. For other even orders $|\boldsymbol{\alpha}|=m$ we find that

$$-2\sum_{|\boldsymbol{\alpha}|=m}\int_{B_{\delta,\bullet}(\mathbf{0})}\frac{\mathbf{z}^{\boldsymbol{\alpha}}}{\boldsymbol{\alpha}!}d^{\boldsymbol{\alpha}}u(\mathbf{x})\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z} = \mathcal{O}(\delta^{m-2}).$$

Thus, terms of order higher than $n = 4$ can be neglected and $\mathcal{O}(\delta^2)$ is the highest rate we can expect.

ii) For the nonlocal bilinear form $A_\delta(\cdot,\cdot)$, we can write

$$\begin{aligned} A_\delta(u,v) &= \int_{\Omega\cup\Omega_I}\int_{(\Omega\cup\Omega_I)\cap B_{\delta,\bullet}(\mathbf{x})}(u(\mathbf{x})-u(\mathbf{y}))(v(\mathbf{x})-v(\mathbf{y}))\phi(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x} \\ &= \int_{\Omega}\int_{B_{\delta,\bullet}(\mathbf{x})}(u(\mathbf{x})-u(\mathbf{y}))(v(\mathbf{x})-v(\mathbf{y}))\phi(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x} \\ &\quad + \int_{\Omega_I}\int_{\Omega\cap B_{\delta,\bullet}(\mathbf{x})}(u(\mathbf{x})-u(\mathbf{y}))(v(\mathbf{x})-v(\mathbf{y}))\phi(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}. \end{aligned}$$

Because u and v are supported in Ω second summand becomes

$$\begin{aligned} &\int_{\Omega_I}\int_{\Omega\cap B_{\delta,\bullet}(\mathbf{x})}(u(\mathbf{x})-u(\mathbf{y}))(v(\mathbf{x})-v(\mathbf{y}))\phi(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x} \\ &\quad = \int_{\Omega_I}\int_{B_{\delta,\bullet}(\mathbf{x})}(u(\mathbf{x})-u(\mathbf{y}))(v(\mathbf{x})-v(\mathbf{y}))\phi(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}, \end{aligned}$$

where we identity u, v with their zero extensions to all of $\mathbb{R}^d$. Thus, we obtain

$$\begin{aligned} A_\delta(u,v) &= \int_{\Omega\cup\Omega_I}\int_{B_{\delta,\bullet}(\mathbf{x})}(u(\mathbf{x})-u(\mathbf{y}))(v(\mathbf{x})-v(\mathbf{y}))\phi(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x} \\ &= \int_{\Omega\cup\Omega_I}\int_{B_{\delta,\bullet}(\mathbf{0})}(u(\mathbf{x})-u(\mathbf{z}+\mathbf{x}))(v(\mathbf{x})-v(\mathbf{z}+\mathbf{x}))\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z}d\mathbf{x}. \end{aligned}$$

Because u and v are assumed to be smooth, we can approximate the difference $-(u(\mathbf{x})-u(\mathbf{z}+\mathbf{x}))$ by the Taylor polynomial $-(u(\mathbf{x})-u(\mathbf{x}+\mathbf{z})) = \nabla u(\mathbf{x})\cdot\mathbf{z}+\frac{1}{2}\mathbf{z}^T H_u(\mathbf{x})\mathbf{z}+\mathcal{O}(\delta^2)$ and analogously for v. Thus, we obtain

$$\begin{aligned} A_\delta(u,v) = \int_{\Omega\cup\Omega_I}\int_{B_{\delta,\bullet}(\mathbf{0})} &\left(\nabla u(\mathbf{x})^T\mathbf{z}+\frac{1}{2}\mathbf{z}^T H_u(\mathbf{x})\mathbf{z}+\mathcal{O}(\delta^2)\right) \\ &\cdot\left(\nabla v(\mathbf{x})^T\mathbf{z}+\frac{1}{2}\mathbf{z}^T H_v(\mathbf{x})\mathbf{z}+\mathcal{O}(\delta^2)\right)\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z}d\mathbf{x}. \end{aligned}$$

We next expand the brackets and analyze the resulting terms. By part i) of Lemma 7.1.2 all the first-order terms as well as the first-order terms multiplied by second-order terms vanish identically. For example, we find

$$\mathcal{O}(\delta^2)\nabla u(\mathbf{x})^T\int_{B_{\delta,\bullet}(\mathbf{0})}\mathbf{z}\ \phi(\|\mathbf{z}\|_\sharp)d\mathbf{z} = 0,$$

and

$$\int_{B_{\delta,\bullet}(\mathbf{0})}\nabla u(\mathbf{x})^T\mathbf{z}\left(\frac{1}{2}\mathbf{z}^T H_v(\mathbf{x})\mathbf{z}\right)\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z} = 0.$$

Furthermore, we obtain

$$\mathcal{O}(\delta^2)\int_{\Omega\cup\Omega_I}\int_{B_{\delta,\bullet}(\mathbf{0})}\mathbf{z}^T H_u(\mathbf{x})\mathbf{z}\ \phi(\|\mathbf{z}\|_\sharp)d\mathbf{z}d\mathbf{x} = K_\delta\mathcal{O}(\delta^2)\int_{\Omega\cup\Omega_I}\Delta u(x)d\mathbf{x},$$

$$\mathcal{O}(\delta^2)\int_{\Omega\cup\Omega_I}\int_{B_{\delta,\bullet}(\mathbf{0})}\mathbf{z}^T H_v(\mathbf{x})\mathbf{z}\ \phi(\|\mathbf{z}\|_\sharp)d\mathbf{z}d\mathbf{x} = K_\delta\mathcal{O}(\delta^2)\int_{\Omega\cup\Omega_I}\Delta v(x)d\mathbf{x}$$

and

$$\mathcal{O}(\delta^4)\int_{\Omega\cup\Omega_I}\int_{B_{\delta,\bullet}(\mathbf{0})}\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z}d\mathbf{x} = K_\delta\mathcal{O}(\delta^4)\mathcal{O}(\delta^{-2})|\Omega\cup\Omega_I|$$

as well as

$$\begin{aligned}\int_{\Omega\cup\Omega_I}\int_{B_{\delta,\bullet}(\mathbf{0})}\Big(\frac{1}{2}\mathbf{z}^T H_u(\mathbf{x})\mathbf{z}\Big)\Big(\frac{1}{2}\mathbf{z}^T H_v(\mathbf{x})\mathbf{z}\Big)\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z}d\mathbf{x}\\ = K_\delta\mathcal{O}(\delta^2)\frac{1}{4}\sum_{i,j,k,l}\int_{\Omega\cup\Omega_I}H_u^{ij}(\mathbf{x})H_v^{kl}(\mathbf{x})d\mathbf{x}.\end{aligned}$$

Finally, by part iii) of Lemma 7.1.2 we find

$$\begin{aligned}&\int_{\Omega\cup\Omega_I}\int_{B_{\delta,\bullet}(\mathbf{0})}\nabla u(\mathbf{x})^T\mathbf{z}\ \nabla v(\mathbf{x})^T\mathbf{z}\ \phi(\|\mathbf{z}\|_\sharp)d\mathbf{z}d\mathbf{x}\\ &= \int_{\Omega\cup\Omega_I}\nabla u(\mathbf{x})^T\Big(\int_{B_{\delta,\bullet}(\mathbf{0})}\mathbf{z}\mathbf{z}^T\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z}\Big)\nabla v(\mathbf{x})d\mathbf{x} = K_\delta\int_{\Omega\cup\Omega_I}\nabla u(\mathbf{x})^T\nabla v(\mathbf{x})d\mathbf{x}.\end{aligned}$$

All in all we therefore obtain

$$K_\delta^{-1}A_\delta(u,v) = \int_{\Omega\cup\Omega_I}\nabla u(\mathbf{x})^T\nabla v(\mathbf{x})d\mathbf{x} + \mathcal{O}(\delta^2) = \int_\Omega\nabla u(\mathbf{x})^T\nabla v(\mathbf{x})d\mathbf{x} + \mathcal{O}(\delta^2).$$

We do not achieve a higher convergence rate when using higher-order Taylor polynomials because then, e.g., first-order terms meet third-order terms, which only give a rate of $\mathcal{O}(\delta^2)$.

iii) As in i) we identify u, v by their zero extension to $\mathbb{R}^d$ and can consider

$$A_\delta(u,v) = \int_{\Omega\cup\Omega_I}\int_{B_{\delta,\bullet}(\mathbf{x})}(u(\mathbf{x})-u(\mathbf{z}))(v(\mathbf{x})-v(\mathbf{z}))\phi(\|\mathbf{x}-\mathbf{z}\|_\sharp)d\mathbf{z}d\mathbf{x}.$$

Because we have the triangulation $\mathcal{T}^h$ of $\Omega\cup\Omega_I$, we can further write

$$A_\delta(u,v) = \sum_{\mathcal{E}_j}\sum_{\mathcal{E}_{j'}}\int_{\mathcal{E}_j}\int_{\mathcal{E}_{j'}\cap B_{\delta,\bullet}(\mathbf{x})}(u(\mathbf{x})-u(\mathbf{z}))(v(\mathbf{x})-v(\mathbf{z}))\phi(\|\mathbf{x}-\mathbf{z}\|_\sharp)d\mathbf{z}d\mathbf{x}.$$

Without loss of generality let $\delta \ll h_{\mathcal{T}^h}$, such that all elements $\mathcal{E}_j$ only interact with their neighboring elements. Thus we have

$$A_\delta(u,v) = \sum_{\mathcal{E}_j}\int_{\mathcal{E}_j}\int_{(\mathcal{E}_j\cup\mathcal{E}_{j\,I})\cap B_{\delta,\bullet}(\mathbf{x})}(u(\mathbf{x})-u(\mathbf{z}))(v(\mathbf{x})-v(\mathbf{z}))\phi(\|\mathbf{x}-\mathbf{z}\|_\sharp)d\mathbf{z}d\mathbf{x},$$

where $\mathcal{E}_{j\,I}$ denotes the interaction domain corresponding to $\mathcal{E}_j$. Now let us consider the problem element-wise. By partitioning an element δ into $\mathcal{E}_{in}$ and $\mathcal{E}_{out}$, where $\mathcal{E}_{out} := \mathcal{E}\cap B_{\delta,\bullet}(\partial\mathcal{E})$ and $\mathcal{E}_{in} := \mathcal{E}\backslash\mathcal{E}_{out}$ (see Figure 7.1), we obtain

$$\int_{\mathcal{E}}\int_{(\mathcal{E}\cup\mathcal{E}_I)\cap B_{\delta,\bullet}(\mathbf{x})}(u(\mathbf{x})-u(\mathbf{z}))(v(\mathbf{x})-v(\mathbf{z}))\phi(\|\mathbf{x}-\mathbf{z}\|_\sharp)d\mathbf{z}d\mathbf{x}$$

$$
\begin{aligned}
&= \int_{\mathcal{E}_{in}} \int_{B_{\delta,\bullet}(\mathbf{x})} (u(\mathbf{x}) - u(\mathbf{z}))(v(\mathbf{x}) - v(\mathbf{z}))\phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
&+ \int_{\mathcal{E}_{out}} \int_{(\mathcal{E}\cup\mathcal{E}_I)\cap B_{\delta,\bullet}(\mathbf{x})} (u(\mathbf{x}) - u(\mathbf{z}))(v(\mathbf{x}) - v(\mathbf{z}))\phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
&= T_1 + T_2.
\end{aligned}
$$

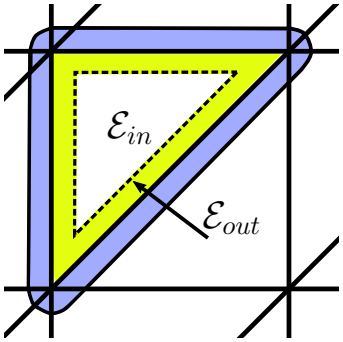


Figure 7.1: Illustration of the subsets $\mathcal{E}_{out} := \mathcal{E} \cap B_{\delta,\bullet}(\partial\mathcal{E})$ and $\mathcal{E}_{in} := \mathcal{E}\backslash\mathcal{E}_{out}$.

Let us first consider T_1. Since $\mathbf{x}, \mathbf{z} \in \mathcal{E}$ we can consider the Taylor series expansion $-(u(\mathbf{x}) - u(\mathbf{x}+\mathbf{z})) = \nabla u(\mathbf{x})^T\mathbf{z} + \frac{1}{2}\mathbf{z}^T H_u(\mathbf{x})\mathbf{z} + \mathcal{O}(\delta^2)$ leading to

$$
\begin{aligned}
T_1 = &\\
&\int_{\mathcal{E}_{in}} \int_{B_{\delta,\bullet}(\mathbf{0})} \left(\nabla u(\mathbf{x})^T\mathbf{z} + \frac{\mathbf{z}^T H_u(\mathbf{x})\mathbf{z}}{2} + \mathcal{O}(\delta^2) \right) \cdot \\
&\qquad \cdot \left(\nabla v(\mathbf{x})^T\mathbf{z} + \frac{\mathbf{z}^T H_v(\mathbf{x})\mathbf{z}}{2} + \mathcal{O}(\delta^2) \right) \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
= &\int_{\mathcal{E}_{in}} \int_{B_{\delta,\bullet}(\mathbf{0})} \nabla u(\mathbf{x})^T\mathbf{z} \nabla v(\mathbf{x})^T\mathbf{z} \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} \\
&+ \int_{\mathcal{E}_{in}} \int_{B_{\delta,\bullet}(\mathbf{0})} \nabla u(\mathbf{x})^T\mathbf{z} \left(\frac{1}{2}\mathbf{z}^T H_v(\mathbf{x})\mathbf{z} \right) \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} \\
&+ \mathcal{O}(\delta^2) \int_{\mathcal{E}_{in}} \int_{B_{\delta,\bullet}(\mathbf{0})} \nabla u(\mathbf{x})^T\mathbf{z} \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} \\
&+ \int_{\mathcal{E}_{in}} \int_{B_{\delta,\bullet}(\mathbf{0})} \nabla v(\mathbf{x})^T\mathbf{z} \left(\frac{1}{2}\mathbf{z}^T H_u(\mathbf{x})\mathbf{z} \right) \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} \\
&+ \int_{\Omega\cup\Omega_I} \int_{B_{\delta,\bullet}(\mathbf{0})} \left(\frac{1}{2}\mathbf{z}^T H_u(\mathbf{x})\mathbf{z} \right) \left(\frac{1}{2}\mathbf{z}^T H_v(\mathbf{x})\mathbf{z} \right) \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
&+ \mathcal{O}(\delta^2)\frac{1}{2} \int_{\Omega\cup\Omega_I} \int_{B_{\delta,\bullet}(\mathbf{0})} \mathbf{z}^T H_u(\mathbf{x})\mathbf{z}\ \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
&+ \mathcal{O}(\delta^2) \int_{\mathcal{E}_{in}} \int_{B_{\delta,\bullet}(\mathbf{0})} \nabla v(\mathbf{x})^T\mathbf{z} \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z}
\end{aligned}
$$

$$+ \mathcal{O}(\delta^2) \int_{\Omega \cup \Omega_I} \int_{B_{\delta,\bullet}(\mathbf{0})} \mathbf{z}^T H_v(\mathbf{x}) \mathbf{z} \ \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x}$$
$$+ \mathcal{O}(\delta^4) \int_{\Omega \cup \Omega_I} \int_{B_{\delta,\bullet}(\mathbf{0})} \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x}.$$

Because all first-order terms and products of first-order terms and third-order terms vanish identically under the integral by Lemma 7.1.2 i), we obtain

$$\begin{aligned} K_\delta^{-1} T_1 &= \int_{\mathcal{E}_{in}} \nabla u(\mathbf{x})^T \nabla v(\mathbf{x}) d\mathbf{z} + \mathcal{O}(\delta^2) \frac{1}{4} \sum_{i,j,k,l} \int_{\Omega \cup \Omega_I} H_u^{ij}(\mathbf{x}) H_v^{kl}(\mathbf{x}) d\mathbf{x} \\ &\quad + \mathcal{O}(\delta^2) \int_{\Omega \cup \Omega_I} \Delta u(x) d\mathbf{x} + \mathcal{O}(\delta^2) \int_{\Omega \cup \Omega_I} \Delta v(x) d\mathbf{x} + \mathcal{O}(\delta^4)\mathcal{O}(\delta^{-2}) |\Omega \cup \Omega_I| \\ &= \int_{\mathcal{E}_{in}} \nabla u(\mathbf{x})^T \nabla v(\mathbf{x}) d\mathbf{z} + \mathcal{O}(\delta^2). \end{aligned}$$

Since $|\mathcal{E} \backslash \mathcal{E}_{in}| = \mathcal{O}(\delta)$ (and the gradients may be constant) we thus find that

$$T_1 = K_\delta \Big(\int_K \nabla u(\mathbf{x})^T \nabla v(\mathbf{x}) d\mathbf{z} + \mathcal{O}(\delta) \Big).$$

Now let us consider the second term T_2. Here, we integrate over the boundary of the elements and we can no longer apply the same Taylor expansion argument to $u(\mathbf{x}) - u(\mathbf{z})$, because the convex combination of $\mathbf{x}$ and $\mathbf{z}$ may now cross the boundary of $\mathcal{E}$ and the gradient is allowed to have jumps there. We can get around this by developing the Taylor series around a point on the boundary. More precisely, let $\mathbf{x}_p$ denote the orthogonal projection of $\mathbf{x}$ onto an edge of $\mathcal{E}$ which is intersected by the convex combination of $\mathbf{x}$ and $\mathbf{z}$. We have that $\|\mathbf{x}_p - \mathbf{x}\| < \delta$, since $\mathbf{x}_p$ lies in the ball around $\mathbf{x}$. Further, by splitting the integral over $(\mathcal{E} \cup \mathcal{E}_I)$ into a sum of integrals over neighboring elements $\mathcal{E}'$, it is clear that $\mathbf{x}_p$ only depends on $\mathbf{x}$, which is crucial here. We now consider the expansions

$$\begin{aligned} u(\mathbf{x}) &= u(\mathbf{x}_p) + \nabla u(\mathbf{x}_p)^T (\mathbf{x} - \mathbf{x}_p) + \mathcal{O}(\delta) \\ u_I(\mathbf{z}) &= u(\mathbf{x}_p) + \nabla u_I(\mathbf{x}_p)^T (\mathbf{z} - \mathbf{x}_p) + \mathcal{O}(\delta). \end{aligned}$$

The subscript on u_I indicates that $\mathbf{z}$ may also lie in the interaction domain of $\mathcal{E}$, implying that $\nabla u_I(\mathbf{x}_p)$ and $\nabla u(\mathbf{x}_p)$ depend on the element from which one we approach the limit. Now, we subtract both lines and exploit that $\mathbf{z} - \mathbf{x}_p = \mathbf{z} - \mathbf{x} + \mathbf{x} - \mathbf{x}_p$, which gives

$$u(\mathbf{x}) - u_I(\mathbf{z}) = (\nabla u(\mathbf{x}_p) - \nabla u_I(\mathbf{x}_p))^T (\mathbf{x} - \mathbf{x}_p) - \nabla u_I(\mathbf{x}_p)^T (\mathbf{z} - \mathbf{x}) + \mathcal{O}(\delta).$$

We apply the same procedure to the function v. Let us define

$$d_u(\mathbf{x}_p) := \nabla u(\mathbf{x}_p) - \nabla u_I(\mathbf{x}_p), \quad g_u(\mathbf{x}_p) := \nabla u_I(\mathbf{x}_p)$$

and the same for v. Inserting this into the expression for T_2 yields

$$T_2 = \int_{\mathcal{E}_{out}} \int_{(\mathcal{E} \cup \mathcal{E}_I) \cap B_{\delta,\bullet}(\mathbf{x})} (u(\mathbf{x}) - u(\mathbf{z}))(v(\mathbf{x}) - v(\mathbf{z})) \phi(\|\mathbf{x} - \mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x}$$

$$
\begin{aligned}
&= \int_{\mathcal{E}_{out}} \int_{(\mathcal{E}\cup\mathcal{E}_I)\cap B_{\delta,\bullet}(\mathbf{x})} \Big(d_u(\mathbf{x}_p)^T(\mathbf{x}-\mathbf{x}_p) - g_u(\mathbf{x}_p)^T(\mathbf{z}-\mathbf{x}) + \mathcal{O}(\delta)\Big) \cdot \\
&\qquad \cdot \Big(d_v(\mathbf{x}_p)^T(\mathbf{x}-\mathbf{x}_p) - g_v(\mathbf{x}_p)^T(\mathbf{z}-\mathbf{x}) + \mathcal{O}(\delta)\Big)\, \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
&= \int_{\mathcal{E}_{out}} d_u(\mathbf{x}_p)^T(\mathbf{x}-\mathbf{x}_p) d_v(\mathbf{x}_p)^T(\mathbf{x}-\mathbf{x}_p) \int_{B_{\delta,\bullet}(\mathbf{x})} \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
&- \int_{\mathcal{E}_{out}} d_u(\mathbf{x}_p)^T(\mathbf{x}-\mathbf{x}_p) g_v(\mathbf{x}_p)^T \int_{B_{\delta,\bullet}(\mathbf{x})} (\mathbf{z}-\mathbf{x}) \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
&\quad + \mathcal{O}(\delta) \int_{\mathcal{E}_{out}} d_u(\mathbf{x}_p)^T(\mathbf{x}-\mathbf{x}_p) \int_{B_{\delta,\bullet}(\mathbf{x})} \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
&\quad - \int_{\mathcal{E}_{out}} d_v(\mathbf{x}_p)^T(\mathbf{x}-\mathbf{x}_p) g_u(\mathbf{x}_p)^T \int_{B_{\delta,\bullet}(\mathbf{x})} (\mathbf{z}-\mathbf{x}) \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
&\quad + \int_{\mathcal{E}_{out}} g_u(\mathbf{x}_p)^T \left(\int_{B_{\delta,\bullet}(\mathbf{x})} (\mathbf{z}-\mathbf{x})(\mathbf{z}-\mathbf{x})^T \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} \right) g_v(\mathbf{x}_p) d\mathbf{x} \\
&\quad - \mathcal{O}(\delta) \int_{\mathcal{E}_{out}} g_u(\mathbf{x}_p)^T \int_{B_{\delta,\bullet}(\mathbf{x})} (\mathbf{z}-\mathbf{x}) \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
&\quad + \mathcal{O}(\delta) \int_{\mathcal{E}_{out}} d_v(\mathbf{x}_p)^T(\mathbf{x}-\mathbf{x}_p) \int_{B_{\delta,\bullet}(\mathbf{x})} \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
&\quad - \mathcal{O}(\delta) \int_{\mathcal{E}_{out}} g_v(\mathbf{x}_p)^T \int_{B_{\delta,\bullet}(\mathbf{x})} (\mathbf{z}-\mathbf{x}) \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x} \\
&\quad + \mathcal{O}(\delta^2) \int_{\mathcal{E}_{out}} \int_{B_{\delta,\bullet}(\mathbf{x})} \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} d\mathbf{x}.
\end{aligned}
$$

By Lemma 7.1.2, we then have

$$\int_{B_{\delta,\bullet}(\mathbf{x})} (\mathbf{z}-\mathbf{x}) \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} = (0,\dots,0)$$

and

$$\int_{B_{\delta,\bullet}(\mathbf{x})} (\mathbf{z}-\mathbf{x})(\mathbf{z}-\mathbf{x})^T \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} = K_\delta Id$$

as well as

$$\int_{B_{\delta,\bullet}(\mathbf{x})} \phi(\|\mathbf{x}-\mathbf{z}\|_\sharp) d\mathbf{z} = K_\delta \mathcal{O}(\delta^{-2})$$

and

$$\int_{\mathcal{E}_{out}} d\mathbf{x} = \mathcal{O}(\delta),$$

so that we obtain

$$
\begin{aligned}
{K_\delta}^{-1} T_2 &= \mathcal{O}(\delta^{-2}) \int_{K_{out}} d_u(\mathbf{x}_p)^T(\mathbf{x}-\mathbf{x}_p) d_v(\mathbf{x}_p)^T(\mathbf{x}-\mathbf{x}_p) d\mathbf{x} \\
&\quad + \mathcal{O}(\delta)\mathcal{O}(\delta^{-2}) \int_{K_{out}} d_u(\mathbf{x}_p)^T(\mathbf{x}-\mathbf{x}_p) d\mathbf{x} + \int_{K_{out}} g_u(\mathbf{x}_p)^T g_v(\mathbf{x}_p) d\mathbf{x} \\
&\quad + \mathcal{O}(\delta)\mathcal{O}(\delta^{-2}) \int_{K_{out}} d_v(\mathbf{x}_p)^T(\mathbf{x}-\mathbf{x}_p) d\mathbf{x} + \mathcal{O}(\delta^2)\mathcal{O}(\delta^{-2})\mathcal{O}(\delta).
\end{aligned}
$$

Recalling that $\|\mathbf{x} - \mathbf{x}_p\|_2 = \mathcal{O}(\delta)$ and $|\mathcal{E}_{out}| = \mathcal{O}(\delta)$, we can estimate this by $|{K_\delta}^{-1} T_2| \leq \mathcal{O}(\delta)$. Finally, by summing up over all elements we obtain the estimate. □

Remark 7.2.2. *On a fixed grid, (iii) in Proposition 7.2.1 implies that the difference between any of the entries of the nonlocal and local stiffness matrices $K_\delta^{-1}\mathbf{A}_\delta$ and $\mathbf{A}_0$, respectively, is $\mathcal{O}(\delta)$.*

As a consequence of Proposition 7.2.1 we obtain the following result concerning convergence of the operator with respect to the $L^2(\Omega)$ norm.

Corollary 7.2.3 (L^2 ***convergence***)**.** *For all $u \in C_0^\infty(\Omega \cup \Omega_I)$ with support in Ω we find*

$$\|K_\delta^{-1}\mathcal{L}_\delta u - \Delta u\|_{L^2(\Omega)} \underset{\delta\to 0}{\longrightarrow} 0.$$

Proof. Let $u \in C_0^\infty(\Omega \cup \Omega_I)$ with support in Ω. Then by Proposition 7.2.1 i) we have shown the pointwise convergence of the function $-K_\delta^{-1}\mathcal{L}_\delta u$ to $-\Delta u$ as $\delta \to 0$. In order to apply the dominated convergence theorem we further need an integrable upper bound. For this purpose let us consider, as in the proof of Proposition 7.2.1, the Taylor expansion

$$-\mathcal{L}_\delta u(\mathbf{x}) = -K_\delta \Delta u(\mathbf{x}) + \mathcal{O}(\delta^4) \int_{B_{\delta,\bullet}(\mathbf{0})} \phi(\|\mathbf{y}\|_\sharp) d\mathbf{y}.$$

Because $u \in C_0^\infty$ with support in the bounded domain Ω, we find a constant $C(u) > 0$ independent of $\mathbf{x}$ such that $|-\Delta u(\mathbf{x})| \leq C(u)$. Due to the definition of the scaling constant K_δ, we can further find a constant $C' > 0$ such that

$$\int_{B_{\delta,\bullet}(\mathbf{0})} \phi(\|\mathbf{y}\|_\sharp) d\mathbf{y} \leq K_\delta C' \delta^{-2}.$$

Thus

$$|-K_\delta^{-1}\mathcal{L}_\delta u(\mathbf{x})| \leq C(u) + C''(u)\delta^4 C'\delta^{-2} = C(u) + C''(u)C'\delta^2 \leq C(u) + C''(u)C'$$

where the last inequality follows because $\delta \ll 1$. As the right-hand side is independent of $\mathbf{x}$ and δ, we find

$$\sup_{\delta\in(0,1]} \sup_{\mathbf{x}\in\Omega} |-K_\delta^{-1}\mathcal{L}_\delta u(\mathbf{x})| < \infty.$$

□

7.3 Convergence results for weak solutions

The results of Section 7.2 state the convergence of the nonlocal operator and bilinear form to their local counterparts. In this section, we show the L^2-convergence of the nonlocal weak solution $u_\delta(\mathbf{x})$ to the local weak solution $u_0(\mathbf{x})$ as $\delta \to 0$. For this purpose we need to provide additional results which enable us to invoke the same convergence proof as in [85, Theorem 2.5] which in this paper is established for the "standard" case

$$\gamma_\delta(\mathbf{x}, \mathbf{y}) = \phi(\|\mathbf{x} - \mathbf{y}\|_2)\chi_{B_{\delta,2}(\mathbf{x})}(\mathbf{y}). \tag{7.3}$$

Crucial for the convergence proof in [85] are two results, a uniform nonlocal Poincaré inequality and a compactness result. For convenience, we recall these results.

Lemma 7.3.1 (***Uniform Poincaré inequality*** [85, Lemma 3.2])**.** *There exists a constant $C_p > 0$ such that $\|u\|^2_{L^2(\Omega)} \leq C_p |||u|||_\delta$ for all $\delta \in (0,1]$ and $u \in L^2_c(\Omega \cup \Omega_I)$.*

Lemma 7.3.2 (***Compactness result*** [85, Lemma 3.3])**.** *Given a sequence $\{\delta_n\}_n$ such that $\delta_n \to 0$ as $n \to \infty$, we denote by $u_n \in V^n_c(\Omega \cup \Omega_I)$ the associated nonlocal solution. If $\sup_n |||u_n||| < \infty$, then there exists a subsequence $(u_{n_j})_j$ and a function $u^* \in H^1_0(\Omega)$ such that $u_{n_j} \to u^*$ in $L^2(\Omega)$.*

In order to make these results applicable to our more general case, i.e., to the kernel (7.1), we have to estimate the energy norm induced by such a kernel against the energy norm induced by the standard kernel (7.3). The following lemma is devoted to this aim. Note that an additional monotonicity assumption is invoked.

Lemma 7.3.3. *Let $\gamma_\delta(\mathbf{x},\mathbf{y})$ be as in (7.1) and assume that $\phi\colon (0,\infty) \to \mathbb{R}$ is monotonically decreasing. Then, there exists a kernel $\gamma'_\delta(\mathbf{x},\mathbf{y}) = \gamma'_{\delta,2,2}(\mathbf{x},\mathbf{y})$ with*

$$\int_{\mathbb{R}^d} z_1^2 \gamma'_\delta(\|\mathbf{z}\|_2) d\mathbf{z} = 1$$

and a constant $C > 0$ independent of δ such that

$$|||u|||' \leq C|||u|||,$$

where $|||\cdot|||'$ and $|||\cdot|||$ denote the energy norms induced by the kernels γ'_δ and γ_δ, respectively. For instance $|||\cdot||| = \sqrt{A(\cdot,\cdot)}$.

Proof. Because we consider homogeneous volume constraints, we can identify the associated nonlocal solutions with their zero extension to all of $\mathbb{R}^d$. From the assumptions, it follows that there exist constants $c_1 \geq 1$ and $c_2 < 1$ such that
• either the norms have already the desired property to exploit the monotonicity of ϕ or we have to increase $\|\cdot\|_2$, so that

$$\phi(c_1\|\mathbf{z}\|_2) \leq \phi(\|\mathbf{z}\|_\sharp),$$

and
• we want to consider a $\|\cdot\|_2$-ball which is contained in the $\|\cdot\|_\sharp$-ball, so that

$$\chi_{B_{c_2\delta,2}(\mathbf{x})}(\mathbf{y}) \leq \chi_{B_{\delta,\bullet}(\mathbf{x})}(\mathbf{y}) \quad \text{or equivalently} \quad c_2 B_{\delta,2}(\mathbf{0}) \subset B_{\delta,\bullet}(\mathbf{0}).$$

Then, we introduce the auxiliary kernel

$$\gamma'_\delta(\mathbf{x},\mathbf{y}) := \frac{1}{K'_\delta}\phi(c_1\|\mathbf{z}\|_2)\chi_{B_{c_2\delta,2}(\mathbf{x})}(\mathbf{y}),$$

where

$$\begin{aligned} K'_\delta &:= \int_{\mathbb{R}^d} z_1^2 \phi(c_1\|\mathbf{z}\|_2)\chi_{B_{c_2\delta,2}(\mathbf{0})}(\mathbf{y}) d\mathbf{z} = \int_{B_{c_2\delta,2}(\mathbf{0})} z_1^2 \phi(c_1\|\mathbf{z}\|_2) d\mathbf{z} \\ &\leq \int_{B_{\delta,\bullet}(\mathbf{0})} z_1^2 \phi(\|\mathbf{z}\|_\sharp) d\mathbf{z} = K_\delta. \end{aligned}$$

Then, by construction, $\gamma'_\delta(\mathbf{x},\mathbf{y}) = \gamma'_\delta(\|\mathbf{x}-\mathbf{y}\|_2)$, $\int_{\mathbb{R}^d} z_1^2\gamma'_\delta(\|\mathbf{z}\|_2)d\mathbf{z} = 1$, and $\gamma'_\delta(\mathbf{x},\mathbf{y}) \leq \frac{K_\delta}{K'_\delta}\gamma_\delta(\mathbf{x},\mathbf{y})$. We therefore find that

$$|||u|||' \leq \frac{K_\delta}{K'_\delta}|||u|||.$$

In order to verify the latter, let us consider $u\colon \mathbb{R}^d \to \mathbb{R}$ with $u \equiv 0$ on $\mathbb{R}^d\backslash\Omega$ and $|||u||| < \infty$. Then we find

$$\begin{aligned}
|||u|||'^2 &= \int_{\Omega\cup\Omega'_\delta}\int_{\Omega\cup\Omega'_I}(u(\mathbf{x})-u(\mathbf{y}))^2\gamma'_\delta(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}\\
&= \int_\Omega\int_\Omega(u(\mathbf{x})-u(\mathbf{y}))^2\gamma'_\delta(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x} + 2\int_\Omega u(\mathbf{x})^2\int_{\Omega'_I}\gamma'_\delta(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}\\
&\leq \frac{K_\delta}{K'_\delta}\int_\Omega\int_\Omega(u(\mathbf{x})-u(\mathbf{y}))^2\gamma_\delta(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x} + 2\frac{K_\delta}{K'_\delta}\int_\Omega u(\mathbf{x})^2\int_{\Omega'_I}\gamma_\delta(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}\\
&\leq \frac{K_\delta}{K'_\delta}\int_\Omega\int_\Omega(u(\mathbf{x})-u(\mathbf{y}))^2\gamma_\delta(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x} + 2\frac{K_\delta}{K'_\delta}\int_\Omega u(\mathbf{x})^2\int_{\Omega_I}\gamma_\delta(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}\\
&= \frac{K_\delta}{K'_\delta}|||u|||^2,
\end{aligned}$$

where the last inequality follows from $\Omega'_I \subset \Omega_I$ since $c_2B_{\delta,2}(\mathbf{0}) \subset B_{\delta,*}(\mathbf{0})$ due to the choice of c_2. To complete the proof, we show that

$$\frac{K_\delta}{K'_\delta} = \frac{\int_{B_{\delta,\bullet}(\mathbf{0})} z_1^2\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z}}{\int_{B_{c_2\delta,2}(\mathbf{0})} z_1^2\phi(c_1\|\mathbf{z}\|_2)d\mathbf{z}}$$

can be bounded above by some constant independent of δ. For this purpose we first note that we can choose c_1 and c_2 above in such a way that $c_1c_2 \leq 1$. This is due to the fact that c_1 can be bounded (due to norm equivalences) and depending on that bound we can then choose c_2 sufficiently small. Consequently, $\phi(c_1c_2\|\mathbf{z}\|_2) \geq \phi(\|\mathbf{z}\|_2)$ due to the monotonicity assumption on ϕ. Furthermore, due to the norm equivalences, we can choose a bounded $c_3 \geq 1$ independent of δ such that $B_{\delta,\bullet}(\mathbf{0}) \subset c_3B_{\delta,2}(\mathbf{0})$ and $\phi(c_3\|\mathbf{z}\|_\sharp) \leq \phi(\|\mathbf{z}\|_2)$. Combining all these considerations we find

$$\begin{aligned}
K'_\delta &= \int_{B_{c_2\delta,2}(\mathbf{0})} z_1^2\phi(c_1\|\mathbf{z}\|_2)d\mathbf{z} = c_2^{d+2}\int_{B_{\delta,2}(\mathbf{0})} z_1^2\phi(c_1c_2\|\mathbf{z}\|_2)d\mathbf{z}\\
&\geq c_2^{d+2}\int_{B_{\delta,2}(\mathbf{0})} z_1^2\phi(\|\mathbf{z}\|_2)d\mathbf{z}
\end{aligned}$$

and also

$$\begin{aligned}
K_\delta = \int_{B_{\delta,\bullet}(\mathbf{0})} z_1^2\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z} \leq \int_{c_3B_{\delta,2}(\mathbf{0})} z_1^2\phi(\|\mathbf{z}\|_\sharp)d\mathbf{z} &= c_3^{d+2}\int_{B_{\delta,2}(\mathbf{0})} z_1^2\phi(c_3\|\mathbf{z}\|_\sharp)d\mathbf{z}\\
&\leq c_3^{d+2}\int_{B_{\delta,2}(\mathbf{0})} z_1^2\phi(\|\mathbf{z}\|_2)d\mathbf{z}.
\end{aligned}$$

Thus, we find that

$$\frac{K_\delta}{K'_\delta} \leq \frac{c_3^{d+2}}{c_2^{d+2}} =: C(d, \bullet, \sharp, 2) = C$$

which defines an upper bound for this quotient that is independent of δ. □

With the help of this result, it is straightforward to see that Lemmata 7.3.1 and 7.3.2 also hold for the nonstandard case. Thus, these two results together with Proposition 7.2.1 and Corollary 7.2.3, provide the necessary foundations for the following theorem, whose proof proceeds in complete analogy to the one in [85, Theorem 2.5] for the "standard" case and therefore is omitted here.

Theorem 7.3.4 (***Convergence of weak solutions***)**.** *With the same assumptions as in Lemma 7.3.3, we obtain*

$$\|u_\delta - u_0\|_{L^2(\Omega)} \to 0 \quad \text{as } \delta \to 0.$$

7.4 Example scaling constants

The following lemma provides the correct scaling constant K_δ for several combinations of kernels and ball types which is necessary for concrete numerical computations.

Lemma 7.4.1. *Let the kernel $\gamma_\delta(\mathbf{x}, \mathbf{y})$ be as in (7.1), K_δ be as in (7.2) and let $\Phi(r)$ denote the first antiderivative of the kernel function $\phi(r)$. Then,*

i) for the the $\|\cdot\|_2$-ball and ϕ radial with respect to the $\|\cdot\|_2$-norm

$$K_\delta = \frac{b_d}{d}\int_0^\delta \phi(r) r^{d+1} dr \qquad \text{with} \quad \begin{array}{c|c|c|c|c|c} d & 1 & 2 & 3 & 4 & 5 \\ \hline b_d & 2 & \pi & \frac{4}{3}\pi & \frac{1}{2}\pi^2 & \frac{8}{15}\pi^2 \end{array}.$$

ii) for $\phi \equiv 1$ (see Figure 7.2 for an illustration in 1d) and

a) the $\|\cdot\|_1$-ball, $\quad K_\delta = \dfrac{2^{d+1}}{(d+2)!}\delta^{d+2}$

b) the $\|\cdot\|_2$-ball, $\quad K_\delta = \dfrac{b_d}{d+2}\delta^{d+2}$

c) the $\|\cdot\|_\infty$-ball, $\quad K_\delta = \dfrac{2^d}{3}\delta^{d+2}$

iii) a few examples for $d = 2$

a) ϕ radial with respect to $\|\cdot\|_1$-norm and $\|\cdot\|_1$-ball

$$K_\delta = \int_{B_{\delta,1}(\mathbf{0})} z_1^2 \phi(\|\mathbf{z}\|_1) d\mathbf{z} = 2 \cdot 2 \left(\frac{\delta^3}{3}\Phi(\delta) - \int_0^\delta s^2 \Phi(s) ds \right)$$

b) ϕ radial with respect to $\|\cdot\|_2$-norm and $\|\cdot\|_2$-ball

$$K_\delta = \int_{B_{\delta,2}(\mathbf{0})} z_1^2 \phi(\|\mathbf{z}\|_2) d\mathbf{z} = 3 \cdot \pi \left(\frac{\delta^3}{3} \Phi(\delta) - \int_0^\delta s^2 \Phi(s) ds \right)$$

c) ϕ radial with respect to $\|\cdot\|_\infty$-norm and $\|\cdot\|_\infty$-ball

$$K_\delta = \int_{B_{\delta,\infty}(\mathbf{0})} z_1^2 \phi(\|\mathbf{z}\|_\infty) d\mathbf{z} = 4 \cdot 4 \left(\frac{\delta^3}{3} \Phi(\delta) - \int_0^\delta s^2 \Phi(s) ds \right)$$

d) ϕ radial with respect to $\|\cdot\|_2$-norm and $\|\cdot\|_\infty$-ball

$$K_\delta = \pi \int_0^\delta \phi(r) r^3 dr + 4\Big(\int_0^{\pi/4} \cos(\theta)^2 \int_\delta^{\delta/\cos(\theta)} \phi(r) r^3 dr d\theta + \int_{\pi/4}^{\pi/2} \cos(\theta)^2 \int_\delta^{\delta/\sin(\theta)} \phi(r) r^3 dr d\theta \Big).$$

Note that the assumption $\phi \equiv 1$ is without loss of generality.

Proof. i) We apply the transformation $\Phi \colon [0,\pi]^{d-2} \times [0,2\pi] \times [0,\delta] \to B_{\delta,2}(\mathbf{0})$, where

$$\Phi(\theta, r) := r \begin{pmatrix} \cos(\theta_1) \\ \sin(\theta_1)\cos(\theta_2) \\ \vdots \\ \sin(\theta_1)\sin(\theta_2)\cdots\sin(\theta_{d-2})\cos(\theta_{d-1}) \\ \sin(\theta_1)\sin(\theta_2)\cdots\sin(\theta_{d-1}) \end{pmatrix}$$

with $|\det \Phi(\theta, r)| = r^{d-1} \prod_{i=1}^{d-2} \sin(\theta_i)^{d-i-1}$. Thus by Lemma 7.1.2 iii) we obtain

$$\begin{aligned} dK_\delta &= \int_{B_{\delta,2}(\mathbf{0})} \mathbf{z}^T \mathbf{z} \phi(\|\mathbf{z}\|_2) d\mathbf{z} \\ &= \int_0^\pi \dots \int_0^\pi \int_0^{2\pi} \int_0^\delta \Phi_d(\theta, r)^2 \phi(r) r^{d-1} \prod_{i=1}^{d-2} \sin(\theta_i)^{d-i-1} dr d\theta_{d-1} \dots d\theta_1 \\ &= \Big(\int_0^\delta \phi(r) r^{d+1} dr \Big) \Big(\int_0^\pi \dots \int_0^\pi \int_0^{2\pi} \Big(\prod_{i=1}^{d-1} \sin(\theta_i) \Big)^2 \prod_{i=1}^{d-2} \sin(\theta_i)^{d-i-1} d\theta_{d-1} \dots d\theta_1 \Big). \end{aligned}$$

Now it remains to compute the second factor. We find

$$\begin{aligned} b_d &:= \int_0^\pi \dots \int_0^\pi \int_0^{2\pi} \left(\prod_{i=1}^{d-1} \sin(\theta_i) \right)^2 \prod_{i=1}^{d-2} \sin(\theta_i)^{d-i-1} d\theta_{d-1} \dots d\theta_1 \\ &= \left(\int_0^{2\pi} \sin(\theta_{d-1})^2 d\theta_{d-1} \right) \cdot \prod_{i=1}^{d-2} \int_0^\pi \sin(\theta)^{d-i+1} d\theta = \pi \prod_{i=1}^{d-2} \int_0^\pi \sin(\theta)^{d-i+1} d\theta. \end{aligned}$$

Further, by applying formulas for integrating powers of trigonometric functions we obtain

$$\int_0^\pi \sin(x)^k dx = \begin{cases} \left(\frac{k-1}{k}\frac{k-3}{k-2}\cdots\frac{1}{2}\right)\pi & : k \text{ even} \\ \left(\frac{k-1}{k}\frac{k-3}{k-2}\cdots\frac{2}{3}\right)2 & : \text{else.} \end{cases}$$

For different k we find for example

k	2	3	4	5
$\int_0^\pi \sin(x)^k dx$	π	$\frac{4}{3}$	$\frac{3}{8}\pi^2$	$\frac{16}{15}$

Finally, by building the product up to $(d-2)$ we find the values for b_d presented in the lemma.

iia) We can show that

$$K_\delta = \int_{B_{\delta,1}(\mathbf{0})} z_1^2 d\mathbf{z} = 2^d \int_0^\delta \int_0^{s_1} \cdots \int_0^{s_{d-1}} s_d^2 ds_d \dots ds_1,$$

which can be computed with the help of the following formula (which is a consequence of Fubini's Theorem)

$$\int_0^\delta \int_0^{s_1} \cdots \int_0^{s_{d-1}} f(s_d) ds_d \dots ds_1 = \frac{1}{(d-1)!} \int_0^\delta (\delta - t)^{d-1} f(t) dt.$$

Here $f(t) = t^2$ and thus

$$K_\delta = \frac{2^d}{(d-1)!} \int_0^\delta (\delta - t)^{d-1} t^2 dt.$$

Applying integration by parts twice leads to the stated result.

iib) This is a special case of i).

iic) Here we have

$$K_\delta = \int_{B_{\delta,\infty}(\mathbf{0})} z_1^2 d\mathbf{z} = 2^d \int_0^\delta \cdots \int_0^\delta s_1^2 ds_d \dots ds_1 = 2^d \delta^{d-1} \frac{\delta^3}{3}.$$

iiia) Here

$$\begin{aligned} K_\delta &= \int_{B_{\delta,1}(\mathbf{0})} z_1^2 \phi(\|\mathbf{z}\|_1) d\mathbf{z} = 4 \int_0^\delta \int_0^{\delta - s} s^2 \phi(s+t) dt ds = 4 \int_0^\delta s^2 \int_s^\delta \phi(t) dt ds \\ &= 4 \int_0^\delta s^2 \left(\Phi(\delta) - \Phi(s)\right) ds = 4 \left(\frac{\delta^3}{3} \Phi(\delta) - \int_0^\delta s^2 \Phi(s) ds \right). \end{aligned}$$

iiib) This is a special case of i) and we can rewrite it in analogy to the other cases with the help of partial integration

$$K_\delta = \pi \int_0^\delta \phi(r) r^3 dr = \pi \Big(s^3 \Phi(s) |_0^\delta - \int_0^\delta 3 s^2 \Phi(s) ds \Big) = \pi \Big(\delta^3 \Phi(\delta) - 3 \int_0^\delta s^2 \Phi(s) ds \Big).$$

iiic) We apply the transformation $\Phi(\theta, r) := (s, t)$ such that $B_{\delta,\infty}(\mathbf{0})$ $= \{\Phi(s,t) : s,t \in (-\delta,\delta)\}$ and $|\det d\Phi(s,t)| = 1$ and only integrate over the first (positive) quadrant. Thus we obtain

$$\begin{aligned}
K_\delta &= \int_{B_{\delta,\infty}} z_1^2 \phi(\|\mathbf{z}\|_\infty) d\mathbf{z} = 4\int_0^\delta \int_0^\delta s^2 \phi(\max\{s,t\}) dt ds \\
&= 4\int_0^\delta \int_0^s s^2 \phi(s) dt ds + 4\int_0^\delta \int_s^\delta s^2 \phi(t) dt ds \\
&= 4\int_0^\delta s^3 \phi(s) ds + 4\int_0^\delta s^2 (\Phi(\delta) - \Phi(s)) ds \\
&= 4\left(s^3 \Phi(s)|_0^\delta - \int_0^\delta 3s^2 \Phi(s) ds \right) + 4\frac{\delta^3}{3}\Phi(\delta) - 4\int_0^\delta s^2 \Phi(s) ds \\
&= 4\delta^3 \Phi(\delta) - 12\int_0^\delta s^2 \Phi(s) ds + 4\frac{\delta^3}{3}\Phi(\delta) - 4\int_0^\delta s^2 \Phi(s) ds \\
&= 16\left(\frac{\delta^3}{3}\Phi(\delta) - \int_0^\delta s^2 \Phi(s) ds \right).
\end{aligned}$$

iiid) First we partition the domain of integration which yields

$$K_\delta = \int_{B_{\delta,\infty}(\mathbf{0})} z_1^2 \phi(\|\mathbf{z}\|_2) d\mathbf{z} = \int_{B_{\delta,2}(\mathbf{0})} y_1^2 \phi(\|\mathbf{z}\|_2) d\mathbf{z} + \int_{B_{\delta,\infty}(\mathbf{0}) \setminus B_{\delta,2}(\mathbf{0})} z_1^2 \phi(\|\mathbf{z}\|_2) d\mathbf{z}.$$

We apply the same procedure as in the proof of Lemma 6.1.5 with the difference that we have to partition the domain into 4 instead of 8 equal pieces now due to the presence of the term z_1^2. We obtain

$$\begin{aligned}
&\int_{B_{\delta,\infty}(\mathbf{0}) \setminus B_{\delta,2}(\mathbf{0})} z_1^2 \phi(\|\mathbf{z}\|_2) d\mathbf{z} \\
=&4\left(\int_0^{\pi/4} \int_\delta^{u_1(\theta)} \cos(\theta)^2 \phi(r) r^3 dr d\theta + \int_{\pi/4}^{\pi/2} \int_\delta^{u_2(\theta)} \cos(\theta)^2 \phi(r) r^3 dr d\theta \right)
\end{aligned}$$

where $u_1(\theta) = \delta/\cos(\theta)$ and in the same way we can derive $u_2(\theta) = \delta/\sin(\theta)$, again by intersecting "rays" with the boundary of $B_{\delta,\infty}(\mathbf{0})$. More precisely,

$$u_2(\theta) := \|r(\theta) \cap \{(0,\delta) + \lambda(\delta, 0) : \lambda \in (0,1)\}\|_2$$

where

$$r(\theta) = \{\lambda(\cos(\theta), \sin(\theta)) : \lambda \geq \delta\}.$$

□

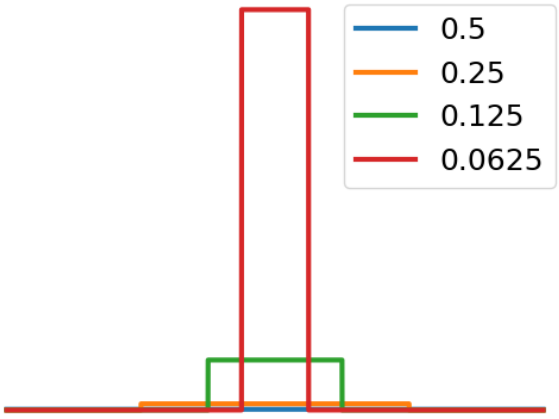


Figure 7.2: A one-dimensional example of a truncated constant kernel $\gamma(|x|) = {K_\delta}^{-1}\chi_{|x|<\delta}$, which is scaled according to ii) in Lemma 7.4.1 for $\delta \in \{0.5, 0.25, 0.125, 0.0625\}$. As $\delta \to 0$ this kernel approaches the Dirac delta function.

7.5 Numerical examples

We provide two numerical examples illustrating the theoretical results given in Proposition 7.2.1. In all our tests we discretize the weak form of the nonlocal equation by the finite element method using continuous piecewise linear basis functions.

Example 1. Let $\Omega = (0, 3)$, $\gamma = \frac{3}{2\delta^3}\chi_{(x-\delta,x+\delta)}(y)$, and $\delta \in \{0.1 \cdot 2^{-i} \ : \ i = 0, \dots, 9\}$. For fixed functions u, v, we investigate the convergence behavior of $|A_\delta(u, v) - a_0(u, v)|$ as $\delta \to 0$.

Row (a) in Table 7.1 corresponds to case (i) of Proposition 7.2.1 since u and v are C^∞ functions. As predicted, we observe a convergence rate of 2.001; this rate is obtained by a linear fit of the data. Row (b) corresponds to case (iii) of the same lemma since u and v are piecewise smooth functions. In general, we should expect only a first-order convergence rate. However, an examination of the proof of Proposition 7.2.1 given in Section 7.2 shows that the overall rate of 1 relies on the regularity of the gradients. In example (b) either u or v is zero outside of $[1, 2]$, i.e., T_2 (from the proof of Proposition 7.2.1 iii)) convergences to 0 with a rate higher than 1. Thus, the convergence depends on T_1, given by

$$K_\delta^{-1} T_1 = \int_{1+\delta}^{2-\delta} \nabla u(\mathbf{x})^T \nabla v(\mathbf{x}) d\mathbf{x} + \mathcal{O}(\delta^2).$$

Because $\nabla u(\mathbf{x})^T \nabla v(\mathbf{x})$ is not constant for u and v in (b), in fact $u'(x) \sim \sqrt{x}$, we find that

$$K_\delta^{-1} T_1 = \int_{1+\delta}^{2-\delta} u'(x)v'(x)dx + \mathcal{O}(\delta^2) \to \int_1^2 u'(x)v'(x)dx$$

with a rate higher than 1, namely 1.5. In fact, in our experiment, we observe a rate of 1.498. Row (c) also corresponds to case (iii), but in this case we consider functions

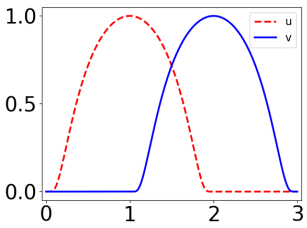
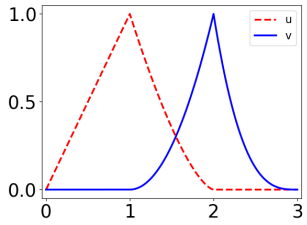
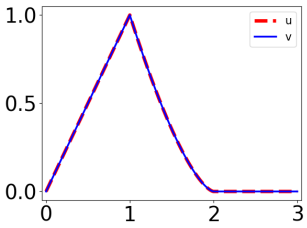
(a)		$u(\mathbf{x}) = \begin{cases} e^{\frac{-1}{1-(x-1)^2}} & x \in [0,2] \\ 0 & else \end{cases}$ $v(\mathbf{x}) = \begin{cases} e^{\frac{-1}{1-(x-2)^2}} & x \in [1,3] \\ 0 & else \end{cases}$
(b)		$u(x) = \begin{cases} x & x \in [0,1] \\ (2-x)^{\frac{3}{2}} & x \in (1,2] \\ 0 & else \end{cases}$ $v(x) = \begin{cases} (x-1)^2 & x \in [1,2] \\ (3-x)^3 & x \in (2,3] \\ 0 & else \end{cases}$
(c)		$u(x) = \begin{cases} x & x \in [0,1] \\ (2-x)^{\frac{3}{2}} & x \in (1,2] \\ 0 & else \end{cases}$ $v(x) = u(x)$

Table 7.1: Plots of $u(\mathbf{x})$ (red) and $v(\mathbf{x})$ (blue).

whose gradients are constant in $[0,1]$. Therefore the best we can obtain is a rate of 1; in fact, we observe a rate of 1.012.

Example 2. We consider the nonlocal problem (2.1) on the domain $\Omega = (0,1)^2$ with a kernel $\gamma(\mathbf{x},\mathbf{y}) = \frac{1}{C_{\delta,\bullet}}\chi_{B_{\delta,\bullet}(\mathbf{x})}(\mathbf{y})$ for $\|\cdot\|_1$-, $\|\cdot\|_2$-, and $\|\cdot\|_\infty$-balls, a constant source term $f \equiv 1$, and homogeneous Dirichlet data $g \equiv 0$. We discretize the problem with P_1 elements on a uniform grid of size $h = 1/8$ for several interaction radii $\delta \in \{0.1 \cdot 2^{-i} : i = 0, \dots, 6\}$ and compare the resulting stiffness matrices and weak solutions to their local counterparts. The results are presented in Table 7.2. As shown in Lemma 7.2.1 we observe a rate of 1 as $\delta \to 0$.

Ball	$\|\mathbf{A}_\delta - \mathbf{A}_0\|_F$	$\|u_\delta - u_0\|_{H_0^1(\Omega)}$
ℓ^∞	0.926	1.007
ℓ^2	0.940	1.062
ℓ^1	0.967	0.999

Table 7.2: Convergence rates as $\delta \to 0$ derived as linear fit of the measurements.

Chapter 8

Shape optimization for identifying interfaces in nonlocal models

Shape optimization problems which are constrained by partial differential equations are of interest in many fields of application [69, 51, 76, 77] and particularly for inverse problems where the parameter to be estimated, e.g., the diffusivity in a heat equation model, is assumed to be defined piecewise on certain subdomains. Given a rough picture of the configuration, shape optimization techniques can be successfully applied to identify the detailed shape of these subdomains [73, 71, 72, 92].

In this chapter we transfer the problem of parameter identification into a nonlocal regime. Here, the parameter of interest is given by the kernel which specifies the nonlocal convection-diffusion model. We assume that the kernel is defined piecewise with respect to certain disjoint subdomains $\Omega_i \subset \Omega$ with smooth boundaries. Since the kernel is a two-point function accounting for interactions between two possibly disjoint points, it has to interrelate all subdomains. Thus, such a kernel is naturally composed of certain partial kernels γ_{ij} each accounting for nonlocal interactions between one of the possible combinations of two subdomains $\Omega_i \times \Omega_j$. We refer to such a kernel as *mixed kernel* in the remainder. A nonlocal convection-diffusion model defined through a mixed kernel depends on the interfaces between the respective subdomains. Under the assumption that we know the rough setting but are lacking in details, we can apply the techniques developed in the aforementioned papers to identify those interfaces.

For this purpose we formulate a shape optimization problem which is constrained by an interface-dependent nonlocal convection-diffusion model. Here, we do not aim at investigating conceptual improvements of existing shape optimization algorithms. On the contrary, we want to study the applicability of established methods for problems of this type. Thus this chapter can be regarded as a feasibility study where we set a focus on the numerical implementation.

The realization of this plan basically requires two ingredients both of which are worked out in this chapter. First, we need to define a reasonable interface-dependent nonlocal model and provide a finite element code which discretizes a variational formulation thereof. Second, we need to derive the shape derivative of the corresponding

nonlocal bilinear form which is then implemented into an overall shape optimization algorithm.

This leads to the following organization of the present chapter. In Section 8.1 we formulate the shape optimization problem including an interface-dependent nonlocal model. Once established, we briefly recall basic concepts from the shape optimization universe in Section 8.2. Then, an individual section, namely Section 8.3, is devoted to the task of computing the shape derivative of the nonlocal bilinear form since this represents a novelty in the current literature. Along the way we have to further specialize the kernel and therefore remark some properties of the resulting nonlocal model at the end of this section. Finally we present numerical illustrations in Section 8.4 corroborating theoretical findings and close this chapter with a summarizing section including an outline for future investigations in this field.

8.1 Problem formulation

Let $\Omega \subset \mathbb{R}^d$ denote a bounded domain. We assume that this domain is partitioned into a simply connected interior subdomain $\Omega_1 \subset \Omega$ with boundary $\Gamma := \partial\Omega_1$ and a domain $\Omega_2 := \Omega \backslash \overline{\Omega}_1$. Thus we have $\Omega = \Omega(\Gamma) = \Omega_1 \dot{\cup} \Gamma \dot{\cup} \Omega_2$, where $\dot{\cup}$ denotes the disjoint union. Figure 8.1 illustrates this situation. In the following, the boundary Γ of the interior domain Ω_1 is called the *interface* and is assumed to be an element of an appropriate shape space; see also Section 8.2 for a related discussion. Now suppose we are given certain measurements $\bar{u} \colon \Omega \to \mathbb{R}$ on the domain Ω, which we assume to follow a nonlocal model of the form (2.1) with a kernel γ, a forcing term f and Dirichlet data $g = 0$. We assume that the system parameters of the model, or more specific the kernel, account for anisotropic diffusion and convective effects induced by the interface Γ. Thus we think of the interface as being variable and consider a kernel $\gamma = \gamma_\Gamma \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ and also a forcing term $f = f_\Gamma \colon \Omega \to \mathbb{R}$ which both depend in a way on the interface. Then, given the data $\bar{u}$ we aim at identifying the interface Γ for which the nonlocal solution $u(\Gamma)$ corresponding to the kernel γ_Γ and the right-hand side f_Γ is the "best approximation" to the measurements.

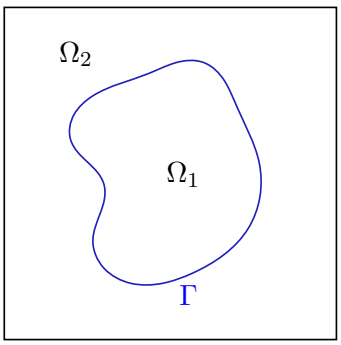


Figure 8.1: An example interface configuration.

Mathematically spoken, we formulate an optimal control problem with a tracking-type objective functional where the interface Γ, modeled as a shape, represents the control variable. We now assume $\Omega := (0,1)^2$ and $\bar{u} \in L^2(\Omega)$ and introduce the following nonlocally constrained shape optimization problem

$$\begin{aligned} &\min_{\Gamma} J(u,\Gamma) \\ &\text{s.t. } A_\Gamma(u,v) = \ell_\Gamma(v) \quad \text{for all } v \in V_c(\Omega \cup \Omega_I). \end{aligned}$$

The objective functional is given by

$$J(u,\Gamma) := j(u,\Gamma) + j_{reg}(\Gamma) := \frac{1}{2}\int_\Omega (u-\bar{u})^2 \, d\mathbf{x} + \nu \int_\Gamma 1 ds.$$

The first term $j(u,\Gamma)$ is a standard L^2 tracking-type functional "projecting" the data on the set of reachable solutions, whereas the second term $j_{reg}(\Gamma)$ is known as the perimeter regularization, which is commonly used in the related literature to overcome possible ill-posedness of optimization problems [9]. The constraint equation is the *homogeneous* weak formulation (3.7) of the nonlocal convection-diffusion problem with defining forms

$$A_\Gamma(u,v) := (-\mathcal{L}_\Gamma u, p)_{L^2(\Omega)} \quad \text{and} \quad \ell_\Gamma(v) := (f_\Gamma, v)_{L^2(\Omega)}, \tag{8.1}$$

where $\mathcal{L}_\Gamma$ denotes the interface-dependent nonlocal convection-diffusion operator

$$-\mathcal{L}_\Gamma u(\mathbf{x}) := \int_{\mathbb{R}^d} \left(u(\mathbf{x})\gamma_\Gamma(\mathbf{x},\mathbf{y}) - u(\mathbf{y})\gamma_\Gamma(\mathbf{y},\mathbf{x})\right) d\mathbf{y}.$$

We postpone the definition of an appropriate energy space $V_c(\Omega \cup \Omega_I)$ until the end of this section.

Remark 8.1.1. *For simplicity, we consider homogeneous Dirichlet data $g = 0$ in this chapter. In this case we find $\mathcal{L} = \mathcal{L}_\Omega$ according to the splitting defined in* (3.1)*. Thus, by adding some appropriate zero terms, the nonlocal bilinear form as defined in* (3.5) *coincides with the definition given in* (8.1)*, which reads as*

$$A_\Gamma(u,v) = \int_\Omega (-\mathcal{L}_\Gamma u)\, v d\mathbf{x} = \int_\Omega v \int_{\mathbb{R}^d} \left(u\gamma_\Gamma - u'\gamma'_\Gamma\right) d\mathbf{y} d\mathbf{x}. \tag{8.2}$$

Since throughout this chapter we need to manipulate the nonlocal bilinear form a couple of times, for the ease of illustration it is advantageous to use representation (8.2) *despite potential zero evaluations of the integrand.*

The operator $\mathcal{L}_\Gamma$ depends on the interface Γ through the kernel $\gamma_\Gamma \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. According to (2.23), by defining

$$\boldsymbol{\alpha}(\mathbf{x},\mathbf{y}) := \frac{(\mathbf{x}-\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|_2}, \qquad \boldsymbol{\Theta}_\Gamma(\mathbf{x},\mathbf{y}) := \gamma_\Gamma^s \, \mathbf{Id}, \qquad \boldsymbol{\mu}_\Gamma(\mathbf{x},\mathbf{y}) := -\gamma_\Gamma^a \boldsymbol{\alpha},$$

we find that

$$\mathcal{L}_\Gamma u := \mathcal{D}(\boldsymbol{\Theta}_\Gamma \mathcal{G} u) - \mathcal{D}(\boldsymbol{\mu}_\Gamma u). \tag{8.3}$$

We consider a *mixed* kernel accounting for anisotropic diffusivity $\boldsymbol{\Theta}_\Gamma$ and convectivity $\boldsymbol{\mu}_\Gamma$ induced by the interface in the following way

$$\gamma_\Gamma(\mathbf{x},\mathbf{y}) = \begin{cases} \gamma_{11}(\mathbf{x},\mathbf{y}): & (\mathbf{x},\mathbf{y}) \in \Omega_1 \times \Omega_1 \\ \gamma_{22}(\mathbf{x},\mathbf{y}): & (\mathbf{x},\mathbf{y}) \in (\Omega_2 \cup \Omega_I) \times \Omega_2 \\ \gamma_{12}(\mathbf{x},\mathbf{y}): & (\mathbf{x},\mathbf{y}) \in \Omega_1 \times (\Omega_2 \cup \Omega_I) \\ \gamma_{21}(\mathbf{x},\mathbf{y}): & (\mathbf{x},\mathbf{y}) \in (\Omega_2 \cup \Omega_I) \times \Omega_1. \end{cases} \tag{8.4}$$

We can write this more concisely as

$$\gamma_\Gamma(\mathbf{x},\mathbf{y}) = \sum_{i,j=1,2} \gamma_{ij}(\mathbf{x},\mathbf{y}) \chi_{\Omega_i \times \Omega_j}(\mathbf{x},\mathbf{y}),$$

where we suppress an explicit designation of the interaction domain Ω_I by redefining

$$\Omega_2 := (\Omega \cup \Omega_I) \backslash \overline{\Omega}_1.$$

By recalling the definition $\gamma_\Gamma^s = \frac{1}{2}(\gamma_\Gamma + \gamma_\Gamma')$ we find for the symmetric part of this kernel

$$\gamma_\Gamma^s(\mathbf{x},\mathbf{y}) = \sum_{i,j=1,2} \frac{(\gamma_{ij} + \gamma'_{ji})}{2} \chi_{\Omega_i \times \Omega_j}(\mathbf{x},\mathbf{y}).$$

Thus, if $\gamma_{ii} = \gamma'_{ii}$, $\gamma_{12} = \gamma'_{21}$ and $\gamma'_{12} = \gamma_{21}$, then we have $\gamma = \gamma^s$ implying that the mixed kernel is symmetric in such cases. Further we assume that each of the partial kernels γ_{ij} is truncated by the same family of interaction sets satisfying (S1) and (S2) from Section 3.2, so that we can consider

$$\gamma_\Gamma(\mathbf{x},\mathbf{y}) = \phi_\Gamma(\mathbf{x},\mathbf{y}) \chi_{S(\mathbf{x})}(\mathbf{y}),$$

for some appropriate interface-dependent kernel function ϕ_Γ. The associated nonlocal bilinear form (8.2) is then given by

$$\begin{aligned} A_\Gamma(u,v) &= \int_\Omega v(\mathbf{x}) \int_{\mathbb{R}^d} (u(\mathbf{x})\gamma_\Gamma(\mathbf{x},\mathbf{y}) - u(\mathbf{y})\gamma_\Gamma(\mathbf{y},\mathbf{x})) d\mathbf{y} d\mathbf{x} \\ &= \sum_{i,j=1,2} \int_{\Omega_i} v(\mathbf{x}) \int_{\Omega_j \cap S(\mathbf{x})} (u(\mathbf{x})\phi_{ij}(\mathbf{x},\mathbf{y}) - u(\mathbf{y})\phi_{ji}(\mathbf{y},\mathbf{x}))\, d\mathbf{y} d\mathbf{x}, \end{aligned}$$

for $u, v \in V_c(\Omega \cup \Omega_I)$. Since there are mixed kernels γ_Γ, for which the nonlocal bilinear form A_Γ is not necessarily coercive (we later provide an example in (8.28)), we cannot define $|||\cdot|||$ as in (3.6).

Along the optimization process we also need to consider the adjoint operator $\mathcal{L}_\Gamma^*$ of the nonlocal convection-diffusion operator $\mathcal{L}_\Gamma$ with respect to the $L^2(\Omega)$ inner product. We find

$$\mathcal{L}_\Gamma^* v(\mathbf{x}) = \int_{\mathbb{R}^d} (v(\mathbf{y}) - v(\mathbf{x})) \gamma_\Gamma(\mathbf{x},\mathbf{y}) d\mathbf{y},$$

which, with the above definitions of the parameters $\boldsymbol{\Theta}_\Gamma$ and $\boldsymbol{\mu}_\Gamma$, can be written as

$$\mathcal{L}_\Gamma^* v(\mathbf{x}) = \mathcal{D}\boldsymbol{\Theta}_\Gamma \mathcal{G} v - \int_{\mathbb{R}^d} \mathcal{G} v^T \boldsymbol{\mu}_\Gamma \, d\mathbf{y}.$$

This duality is an immediate consequence of the relation $\mathcal{D}^* = -\mathcal{G}$ when applied to (8.3). Note that if the kernel is symmetric, i.e., if $\gamma_\Gamma^a = 0$, then the operator is self-adjoint, i.e., $\mathcal{L}_\Gamma^* = \mathcal{L}_\Gamma$. The associated bilinear form is given by

$$\begin{aligned} A_\Gamma^*(v,u) = (-\mathcal{L}_\Gamma^* v, u)_{L^2(\Omega)} &= \int_\Omega u(\mathbf{x}) \int_{\mathbb{R}^d} (v(\mathbf{x}) - v(\mathbf{y}))\, \gamma_\Gamma(\mathbf{x},\mathbf{y}) d\mathbf{y} d\mathbf{x} \\ &= \sum_{i,j=1,2} \int_{\Omega_i} u(\mathbf{x}) \int_{\Omega_j \cap S(\mathbf{x})} (v(\mathbf{x}) - v(\mathbf{y}))\, \phi_{ij}(\mathbf{x},\mathbf{y}) d\mathbf{y} d\mathbf{x}. \end{aligned}$$

By the definition of the adjoint operator we have

$$A_\Gamma^*(v,u) = (-\mathcal{L}_\Gamma^* v, u)_{L^2(\Omega)} = (v, -\mathcal{L}_\Gamma u)_{L^2(\Omega)} = A_\Gamma(u,v) = A_\Gamma'(v,u). \tag{8.5}$$

Remark 8.1.2. *Relation* (8.5) *also implies that the nonlocal stiffness matrix* $\mathbf{A}^*$ *corresponding to the adjoint bilinear form* A_Γ^* *is simply the transpose of the nonlocal stiffness matrix* $\mathbf{A}$ *corresponding to the primal bilinear form* A_Γ*, since*

$$a_{kj}^* = \mathbf{A}^*(\varphi_k, \varphi_j) = \mathbf{A}(\varphi_j, \varphi_k) = a_{jk}.$$

Note that we choose the same finite-dimensional finite element spaces for ansatz functions φ_k *and test functions* φ_j*. Consequently we only need a routine which assembles one of the stiffness matrices. Since for a symmetric kernel the operator takes the form of the adjoint operator, we implement a routine assembling the stiffness matrix corresponding to the adjoint bilinear form.*

For the forcing term we assume a dependency on the interface in the following way

$$f_\Gamma(\mathbf{x}) := \begin{cases} f_1(\mathbf{x}): & \mathbf{x} \in \Omega_1 \\ f_2(\mathbf{x}): & else, \end{cases} \tag{8.6}$$

where we assume that $f_i \in H^1(\Omega_i)$. The right-hand side of the constraint equation is then given by

$$\ell_\Gamma(v) = \int_\Omega f_\Gamma v \, d\mathbf{x} = \int_{\Omega_1} f_1 v \, d\mathbf{x} + \int_{\Omega \setminus \Omega_1} f_2 v \, d\mathbf{x}. \tag{8.7}$$

Our ultimate goal is to solve the shape optimization problem with derivative-based optimization methods. For this purpose it is necessary to require a certain regularity level for the weak solutions of the variational problems corresponding to the primal as well as the adjoint nonlocal bilinear form. In Chapter 3 however, we have pointed out, that for certain kernels we cannot expect a smoothing of the data at all. In fact, for the class of integrable kernels considered in Subsection 3.3.1 there is generally no gain

of regularity; see the discussion at the end of this subsection. That is why we regularize our problem by perturbing the nonlocal operator with a "small" Laplacian which is independent of the interface (isotropic diffusivity). The new bilinear form governing the constraint equation then reads as

$$\begin{aligned}\tilde{A}_\Gamma(u,v) &:= A_\Gamma(u,v) + c_{per}A^{loc}(u,v)\\ &:= (-\mathcal{L}_\Gamma u, v)_{L^2(\Omega)} + c_{per}(-\Delta u, v)_{L^2(\Omega)}\\ &= \sum_{i,j=1,2}\int_{\Omega_i} v \int_{\Omega_j\cap S(\mathbf{x})} \left(u\phi_{ij} - u'\phi'_{ji}\right) d\mathbf{y}d\mathbf{x} + c_{per}\int_\Omega \nabla u^T \nabla v \; d\mathbf{x}.\end{aligned} \tag{8.8}$$

We choose a small *perturbation parameter* $c_{per} > 0$, which in practice does not significantly affect the nonlocal model, but theoretically guarantees sufficient regularity of state and adjoint variables. It is therefore reasonable to define the energy space $V_c(\Omega\cup\Omega_I)$ to consist of weakly differentiable functions with support in Ω. Specifically, we define

$$(V_c(\Omega\cup\Omega_I), |||\cdot|||) := \left(H_c^1(\Omega\cup\Omega_I), |\cdot|_{H^1(\Omega)}\right),$$

where

$$H_c^1(\Omega\cup\Omega_I) := \left\{u\in H^1(\Omega\cup\Omega_I)\colon u = 0 \;\text{ on } \Omega_I\right\}$$

and

$$|u|_{H^1(\Omega)} := \sqrt{A^{loc}(u,u)}.$$

By considering zero extensions we find that $H_c^1(\Omega\cup\Omega_I)$ is equivalent to the standard space $H_0^1(\Omega)$. Throughout the remainder of this Chapter we now consider the *perturbed or regularized shape optimization problem*

$$\begin{aligned}&\min_\Gamma\; J(u,\Gamma)\\ &\text{s.t. }\; \tilde{A}_\Gamma(u,v) = \ell_\Gamma(v) \quad \text{for all } v\in V_c(\Omega\cup\Omega_I).\end{aligned} \tag{8.9}$$

In Figure 8.2 the nonlocal counterpart to what is depicted in Figure 8.1 is illustrated.

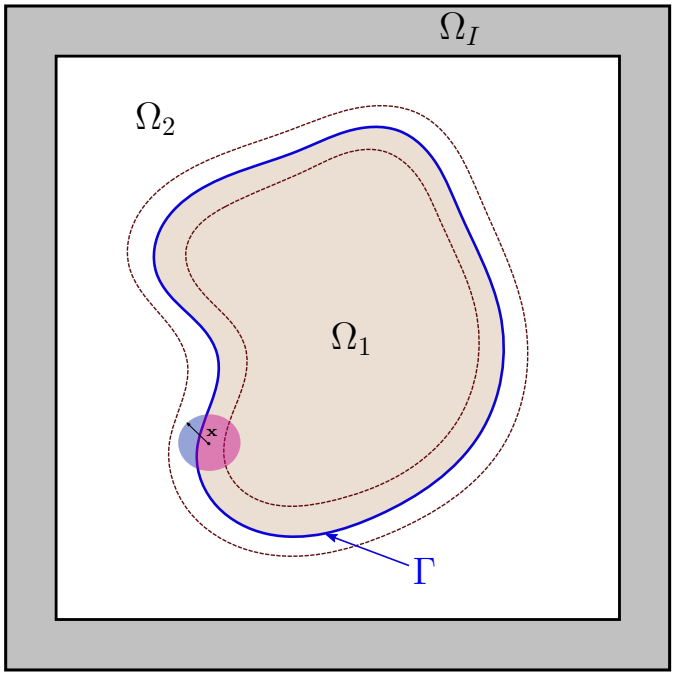


Figure 8.2: An example for the interface configuration in the nonlocal setting. Compared to Figure 8.1, we additionally have the interaction domain Ω_I. Also, the dashed lines parallel to the interface Γ, forming some sort of tube, can be considered a *nonlocal interface*

$$\Gamma_I := \{\mathbf{y} \in \Omega \colon \mathbf{y} \in S(\mathbf{x}) \text{ for some } \mathbf{x} \in \Gamma\}\,.$$

Within this tube we find that nonlocal interactions occurring within the interaction set, in this figure an Euclidean ball centered at $\mathbf{x}$, are weighted differently according to the mixed kernel; this is indicated by the different colors used to fill the ball. More precisely, let us consider the depicted $\mathbf{x} \in \Omega_1$. If $\mathbf{y}$ is located in Ω_2, then the partial kernel function ϕ_{12} is activated (see the blueish colored area of the ball). Whereas if $\mathbf{y} \in \Omega_1$ then the kernel function ϕ_{11} quantifies nonlocal interactions (see the reddish portion of the ball). Thus these partially differently weighted interactions close to the interface induce convective effects.

8.2 Basic concepts in shape optimization

For solving the constrained shape optimization problem (8.9) we want to use the same shape optimization algorithms as they are developed in [73, 71, 70] for problem classes that are comparable in structure. Thus, in this section we briefly introduce the basic concepts and ideas of the therein applied shape formalism, which are sufficient enough to understand and realize the numerical implementation of these methods. For a rigorous introduction to shape spaces, shape derivatives and shape calculus in general, we refer to the monographs [31, 82, 92].

8.2.1 Notations and definitions

Based on our perception of the interface, we now refer to the image of a simple closed and smooth curve as a *shape*, i.e., the spaces of interest are subsets of

$$\mathcal{A} := \left\{ \Gamma := \varphi(S^1) \colon \varphi \in C^\infty(S^1, \Omega) \text{ injective; } \varphi' \neq 0 \right\}. \tag{8.10}$$

By the Jordan curve theorem [50] such a shape $\Gamma \in \mathcal{A}$ divides the plane into two (simply) connected components with common boundary Γ. One of them is the bounded interior, which in our situation can then be identified with Ω_1.
Functionals $J \colon \mathcal{A} \to \mathbb{R}$ which assign a real number to a shape are called *shape functionals*. Since this chapter deals with minimizing such shape functionals, i.e., with so-called shape optimization problems, we need to introduce the notion of an appropriate *shape derivative*. To this end we consider a family of mappings $F_t \colon \overline{\Omega} \to \mathbb{R}^d$ with $F_0 = id$, where $t \in [0, T]$ and $T > 0$, which transform a shape Γ into a family of *perturbed shapes* $\{\Gamma_t\}_{t\in[0,T]}$, where $\Gamma_t := F_t(\Gamma)$ with $\Gamma_0 = \Gamma$. Here the family of mappings $\{F_t\}_{t\in[0,T]}$ is described by the *perturbation of identity*, which for a smooth vector field $\mathbf{V} \in C_0^k(\Omega, \mathbb{R}^d)$, $k \in \mathbb{N}$, is defined by

$$F_t(\mathbf{x}) := \mathbf{x} + t\mathbf{V}(\mathbf{x}).$$

We note that for sufficiently small $t \in [0, T]$ the mapping $\mathbf{x} \mapsto t\mathbf{V}(\mathbf{x})$ defines a contraction on Ω implying that F_t is injective, and thus $\Gamma_t \in \mathcal{A}$. Then the *Eulerian or directional derivative* of a shape functional J at a shape Γ in direction of a vector field $\mathbf{V} \in C_0^k(\Omega, \mathbb{R}^d)$, $k \in \mathbb{N}$, is defined by

$$DJ(\Gamma)[\mathbf{V}] := \left.\frac{d}{dt}\right|_{t=0^+} J(F_t(\Gamma)) = \lim_{t \searrow 0} \frac{(J(F_t(\Gamma)) - J(\Gamma))}{t}. \tag{8.11}$$

If $DJ(\Gamma)[\mathbf{V}]$ exists for all $\mathbf{V} \in C_0^k(\Omega, \mathbb{R}^d)$ and $\mathbf{V} \mapsto DJ(\Gamma)[\mathbf{V}]$ is in the dual space $\left(C_0^k(\Omega, \mathbb{R}^d)\right)^*$, then $DJ(\Gamma)[\mathbf{V}]$ is called the *shape derivative* of J [92, Definition 4.6].

For demonstration purposes and later use we compute the shape derivative of the right-hand side, i.e., of the shape functional

$$\Gamma \mapsto \ell_\Gamma(v) = \int_{\Omega_1} f_1 v d\mathbf{x} + \int_{\Omega\setminus\Omega_1} f_2 v d\mathbf{x},$$

with ℓ_Γ defined in (8.7) and a fixed test function $v \in V_c(\Omega \cup \Omega_I)$. Therefore, let $\mathbf{V} \in C_0^k(\Omega, \mathbb{R}^d)$, then due to $\Omega(F_t(\Gamma)) = F_t(\Omega_1) \dot\cup F_t(\Gamma) \dot\cup F_t(\Omega\setminus\Omega_1)$ we find for the perturbed functional

$$\ell_{F_t(\Gamma)} = \int_{F_t(\Omega_1)} f_1 v d\mathbf{x} + \int_{F_t(\Omega\setminus\Omega_1)} f_2 v d\mathbf{x}. \tag{8.12}$$

Now let us representatively consider the first summand in (8.12). According to the definition of the directional derivative (8.11) we now have to compute variations in t. To

effect this computation, we first shift the time dependency from the domain of integration to the integrand by applying the transformation formula. This yields

$$\int_{F_t(\Omega_1)} f_1 v d\mathbf{x} = \int_{\Omega_1} f_1(F_t(\mathbf{x}))v(F_t(\mathbf{x}))\,|\det dF_t(\mathbf{x})|\,d\mathbf{x}.$$

Now, we invoke the theorem about differentiating parameter integrals, which allows us to interchange the order of differentiation and integration. We obtain

$$\begin{aligned}\frac{d}{dt}\Big|_{t=0^+} \int_{\Omega_1} f_1(F_t(\mathbf{x}))v(F_t(\mathbf{x}))\,|\det dF_t(\mathbf{x})|\,d\mathbf{x}\\ = \int_{\Omega_1} \frac{d}{dt}\Big|_{t=0^+} (f_1(F_t(\mathbf{x}))v(F_t(\mathbf{x}))\,|\det dF_t(\mathbf{x})|)\,d\mathbf{x}.\end{aligned}$$

It can be shown that [53, Section 4.6]

$$\frac{d}{dt}\Big|_{t=0^+} |\det dF_t(\mathbf{x})| = \operatorname{div}\mathbf{V}(\mathbf{x}).$$

Also, let us define the derivative of a function $v\colon \Omega \to \mathbb{R}$ in direction of $\mathbf{V}$ by

$$D_m v(\mathbf{x}) := \dot{v}(\mathbf{x}) := \frac{d}{dt}\Big|_{t=0^+} v(F_t(\mathbf{x})) = \nabla v(\mathbf{x})^T \mathbf{V}(\mathbf{x}).$$

For functions v, which do not explicitly depend on the shape, $D_m v$ is equivalent to what is called the *material derivative*. For their general definition and more details we refer to the literature, e.g., [61]. Due to the following product rule

$$D_m(vw) = D_m(v)w + vD_m(w),$$

we obtain

$$\int_{\Omega_1} \frac{d}{dt}\Big|_{t=0^+} (f_1(F_t(\mathbf{x}))v(F_t(\mathbf{x}))\,|\det dF_t(\mathbf{x})|)\,d\mathbf{x} = \int_{\Omega_1} \left(\dot{f}_1 v + f_1 \dot{v}\right) + f_1 v \operatorname{div}\mathbf{V}\; d\mathbf{x}.$$

All in all, by applying this procedure to the second summand we arrive at

$$\begin{aligned}D(\ell_\Gamma(v))[\mathbf{V}] &= \int_{\Omega_1} (\dot{f}_1 v + f_1 \dot{v}) + f_1 v\ \operatorname{div}\mathbf{V}\ d\mathbf{x} + \int_{\Omega\setminus\Omega_1} (\dot{f}_2 v + f_2 \dot{v}) + f_2 v\ \operatorname{div}\mathbf{V}\ d\mathbf{x}\\ &= \int_\Omega \dot{f}_\Gamma v d\mathbf{x} + \int_\Omega f_\Gamma v\ \operatorname{div}\mathbf{V}\ d\mathbf{x} + \ell_\Gamma(\dot{v}),\end{aligned} \tag{8.13}$$

where $\dot{f}_\Gamma$ has to be understood piecewise according to the definition of f_Γ (8.6). Below, we proceed in the same manner to derive the shape derivative of the nonlocal bilinear form.

8.2.2 Optimization approach: Formal Lagrangian

Let us assume that for each admissible shape Γ there exists a unique solution $u(\Gamma)$ of the constraint equation, i.e., $u(\Gamma)$ satisfies $\tilde{A}_\Gamma(u(\Gamma), v) = \ell_\Gamma(v)$ for all $v \in V_c(\Omega \cup \Omega_I)$. Then we can consider the *reduced problem*

$$\min_\Gamma J^{red}(\Gamma) := J(u(\Gamma), \Gamma).$$

In order to employ derivative based minimization algorithms we need to derive the shape derivative of the reduced objective functional J^{red}. By formally applying the chain rule we obtain

$$DJ^{red}(\Gamma)[\mathbf{V}] = D_u J(u(\Gamma), \Gamma) D_\Gamma u(\Gamma)[\mathbf{V}] + D_\Gamma J(u(\Gamma), \Gamma)[\mathbf{V}],$$

where D_u and D_Γ denote the partial derivatives of the objective J with respect to the state variable u and the control Γ, respectively. In applications we typically do not have an explicit formula for the control-to-state mapping $u(\Gamma)$, so that we cannot analytically quantify the sensitivity of the unique solution $u(\Gamma)$ with respect to the interface Γ. Thus, a formula for the shape derivative $D_\Gamma u(\Gamma)[\mathbf{V}]$ is unattainable. A common approach to still access this derivative is to consider the associated adjoint equation; see, e.g., [92]. More precisely, by introducing an adjoint variable v (also called *Lagrange multiplier*), we define the so-called *Lagrange functional*, or *Lagrangian*, as

$$L(u, \Gamma, v) := J(u, \Gamma) + \tilde{A}_\Gamma(u, v) - \ell_\Gamma(v)$$

and aim to find a saddle point (u, Γ, v), such that, for all variations $(du, \mathbf{V}, dv)$,

$$\begin{aligned}
0 &= D_v L(u, \Gamma, v)[dv] = \tilde{A}_\Gamma(u, dv) - \ell_\Gamma(dv), && \text{(state)} \\
0 &= D_u L(u, \Gamma, v)[du] = \tilde{A}^*_\Gamma(v, du) + (u - \bar{u}, du)_{L^2(\Omega)}, && \text{(adjoint)} \\
0 &= D_\Gamma L(u, \Gamma, v)[\mathbf{V}] = D_\Gamma J(u, \Gamma)[\mathbf{V}] + D_\Gamma \tilde{A}_\Gamma(u, v)[\mathbf{V}] - D_\Gamma \ell_\Gamma(u, v)[\mathbf{V}]. && \text{(design)}
\end{aligned}$$

By inserting the definition (8.8) of the perturbed bilinear form $\tilde{A}_\Gamma = A_\Gamma + c_{per} A^{loc}$ and exploiting that A^{loc} is symmetric, i.e., $A^{loc} = (A^{loc})^*$, and does not depend on the interface, so that $D_\Gamma A^{loc} = 0$, we find

$$\begin{aligned}
0 &= D_v L(u, \Gamma, v)[dv] = A_\Gamma(u, dv) + c_{per} A^{loc}(u, dv) - \ell_\Gamma(dv), && \text{(state)} \\
0 &= D_u L(u, \Gamma, v)[du] = A^*_\Gamma(v, du) + c_{per} A^{loc}(v, du) + (u - \bar{u}, du)_{L^2(\Omega)}, && \text{(adjoint)} \\
0 &= D_\Gamma L(u, \Gamma, v)[\mathbf{V}] = D_\Gamma J(u, \Gamma)[\mathbf{V}] + D_\Gamma A_\Gamma(u, v)[\mathbf{V}] - D_\Gamma \ell_\Gamma(u, v)[\mathbf{V}]. && \text{(design)}
\end{aligned} \tag{8.14}$$

The partial shape derivative of the Lagrangian $D_\Gamma L(u, \Gamma, v)[\mathbf{V}]$ evaluated in a saddle point $(u(\Gamma), \Gamma, v(\Gamma))$ then serves as shape derivative $DJ^{red}(\Gamma)[\mathbf{V}]$ of the reduced objective functional. This has to be proven in each specific setting. In the related literature this is typically done by an application of a theorem of Correa and Seeger [92, Theorem 4.18], which states the differentiability of a min-max function, in our case the Lagrangian. Here, we assume that the prerequisites of this theorem hold.

The first two equations of the system (8.14), namely the state and adjoint equation, are standard and can be found, e.g., in [30, Section 3.3]. In our situation they imply

$$\begin{aligned} A_\Gamma(u, dv) - \ell_\Gamma(dv) &= -c_{per} A^{loc}(u, dv), && \text{(state)} \\ A^*_\Gamma(v, du) - \ell^*(du) &= -c_{per} A^{loc}(v, du), && \text{(adjoint)} \end{aligned}$$

where under abuse of notation we define

$$\ell^*(q) := -\int_\Omega (u - \bar{u}) q \; d\mathbf{x}.$$

The third equation in (8.14), i.e., the design equation, involves the shape derivative and therefore needs careful examination. In fact, the crucial task is the computation of the occurring shape derivatives which then enable us to implement suitable shape optimization algorithms. The shape derivatives of the objective functional and the right-hand side are standard and we recall below the formulas from the pertinent literature; see also (8.13). However the shape derivative of the system model, in particular the shape derivative of the nonlocal bilinear form is nonstandard and cannot yet be found in literature. That is why we devote Section 8.3 to this task.

We end this subsection by inserting the already available shape derivatives of the objective functional and the right-hand side into the design equation of the saddle point system (8.14). Since we only consider transformation vector fields $\mathbf{V} \in C^k_0(\Omega, \mathbb{R}^d)$ which are zero on the boundary, we find that the tracking-type functional $j(u, \Gamma) = \frac{1}{2}\int_\Omega (u - \bar{u})^2 \; d\mathbf{x}$ does not react on such variations, i.e., $j(u, \Gamma) = j(u, F_t(\Gamma))$ for all $t > 0$ and consequently

$$D_\Gamma j(u, \Gamma)[\mathbf{V}] = 0 \quad \text{for all} \quad \mathbf{V} \in C^k_0(\Omega, \mathbb{R}^d).$$

The shape derivative of the regularization term is an immediate consequence of [92, Theorem 4.13] and is given by

$$D j_{reg}(u, \Gamma)[\mathbf{V}] = \nu \int_\Gamma \operatorname{div}_\Gamma \mathbf{V} \; ds = \nu \int_\Gamma \operatorname{div} \mathbf{V} - \mathbf{n}^T \nabla \mathbf{V} \mathbf{n} \; ds,$$

where $\mathbf{n}$ denotes the outer normal of Ω_1. The shape derivate of the right-hand side ℓ_Γ has been computed in (8.13). Finally, for an interface Γ, let $u = u(\Gamma)$ and $v = v(\Gamma)$ solve the state and adjoint equation given in (8.14), respectively, then by inserting the shape derivative formulas derived so far into the design equation of the saddle point system (8.14), we obtain

$$\begin{aligned} &DJ^{red}(\Gamma)[\mathbf{V}] \\ &= D_\Gamma L(u, \Gamma, v)[\mathbf{V}] \\ &= D_\Gamma J(u, \Gamma)[\mathbf{V}] + D_\Gamma A_\Gamma(u, v)[\mathbf{V}] - D_\Gamma \ell_\Gamma(u, v)[\mathbf{V}] \\ &= \nu \int_\Gamma \operatorname{div}_\Gamma \mathbf{V} \; ds + D_\Gamma A_\Gamma(u, v) - \left(\int_\Omega \dot{f}_\Gamma v \; d\mathbf{x} + \int_\Omega f_\Gamma v \; \operatorname{div} \mathbf{V} \; d\mathbf{x} + \ell_\Gamma(\dot{v}) \right). \end{aligned} \tag{8.15}$$

If $\nu = 0$ and f_Γ is piecewise constant with respect to the interface, this expression further reduces to

$$DJ^{red}(\Gamma)[\mathbf{V}] = D_\Gamma A_\Gamma(u, v) - \left(\ell_\Gamma(\dot{v}) + \int_\Omega f_\Gamma v \; \operatorname{div} \mathbf{V} \; d\mathbf{x} \right).$$

It remains to derive an explicit formula for the shape derivative of the nonlocal bilinear form A_Γ; see Section 8.3.

8.2.3 Optimization algorithm

Let us assume for a moment that we have an explicit formula for the shape derivative of the reduced objective functional (8.15). We now briefly recall the techniques developed in [73] and describe how to exploit this derivative for implementing gradient based optimization methods or even Quasi-Newton methods, such as L-BFGS, to solve the constrained shape optimization problem (8.9).

Talking about gradients requires the notion of an inner product, or more generally a Riemannian metric. Unfortunately, shape spaces typically do not admit the structure of a linear space. However, in particular situations it is possible to define appropriate quotient spaces, which can be equipped with a Riemannian structure. For instance consider the set $\mathcal{A}$ introduced in (8.10). Since we are only interested in the image of the defining embedding, a re-parametrization thereof does not lead to a different shape. Consequently, two curves that are equal modulo (diffeomorphic) re-parametrizations define the same shape. This conception naturally leads to the quotient space $\mathrm{Emb}(S^1, \mathbb{R}^d)/\operatorname{Diff}(S^1, S^1)$, which can be considered an infinite-dimensional Riemannian manifold [58]. This example already intimates the difficulty of translating abstract shape derivatives into discrete optimization methods; see, e.g., the thesis [93] on this topic. A detailed discussion of these issues is not the intention of this chapter and we now outline Algorithm 2.

The basic idea can be intuitively explained in the following way. Starting with an initial guess Γ_0, we aim to iterate in a steepest-descent fashion over interfaces Γ_k until we reach a "stationary point" of the reduced objective functional J^{red}. The interface Γ_k is encoded in the finite element mesh and transformations thereof are realized by adding vector fields $\mathbf{U}\colon \Omega \to \mathbb{R}^d$ (which can be interpreted as tangent vectors at a fixed interface) to the finite element nodes which we denote by Ω_k. Thus, the essential part is to update the finite element mesh after each iteration by adding an appropriate transformation vector field. For this purpose, we use the solution $\mathbf{U}(\Gamma)\colon \Omega(\Gamma) \to \mathbb{R}^d$ of the so-called *deformation equation*

$$a_\Gamma(\mathbf{U}(\Gamma), \mathbf{V}) = DJ^{red}(\Gamma)[\mathbf{V}] \quad \text{for all } \mathbf{V} \in H_0^1(\Omega(\Gamma), \mathbb{R}^2). \tag{8.16}$$

The right-hand side of this equation is given by the shape derivative of the reduced objective functional (8.15) and the left-hand side denotes an inner product on the vector field space $H_0^1(\Omega, \mathbb{R}^2)$. In the view of the manifold interpretation, we can consider a_Γ as inner product on the tangent space at Γ, so that $\mathbf{U}(\Gamma)$ is interpretable as the gradient of the shape functional J^{red} at Γ. The solution $\mathbf{U}(\Gamma)\colon \Omega \to \mathbb{R}^2$ of (8.16) is then added to the coordinates Ω_k of the finite element nodes.
A common choice for a_Γ is the bilinear form associated to the linear elasticity equation given by

$$a_\Gamma(\mathbf{U}, \mathbf{V}) = \int_{\Omega(\Gamma)} \sigma(\mathbf{U}) : \epsilon(\mathbf{V})\, dx,$$

for $\mathbf{U}, \mathbf{V} \in H_0^1(\Omega, \mathbb{R}^2)$, where

$$\sigma(\mathbf{U}) := \lambda \mathrm{tr}(\epsilon(\mathbf{U}))I + 2\mu\epsilon(\mathbf{U}) \tag{8.17}$$

and

$$\epsilon(\mathbf{U}) := \frac{1}{2}(\nabla \mathbf{U} + \nabla \mathbf{U}^T)$$

are the strain and stress tensors, respectively. Deformation vector fields $\mathbf{V}$ which do not change the interface do not have an impact on the reduced objective functional, so that

$$DJ^{red}(\Gamma)[\mathbf{V}] = 0 \quad \text{for all } \mathbf{V} \text{ with } \mathrm{supp}(\mathbf{V}) \cap \Gamma = \emptyset.$$

Therefore, the right-hand side $DJ^{red}(\Gamma)[\mathbf{V}]$ is only assembled for test vector fields whose support intersects with the interface Γ and set to zero for all other basis vector fields. This prevents wrong mesh deformations resulting from discretization errors as outlined and illustrated in [71]. Furthermore, λ and μ in (8.17) denote the Lamé parameters which do not need to have a physical meaning here. It is more important to understand their effect on the mesh deformation. They enable us to control the stiffness of the material and thus can be interpreted as some sort of step size. In [70], it is observed that locally varying Lamé parameters have a stabilizing effect on the mesh. A good strategy is to choose $\lambda = 0$ and μ as solution of the following Poisson equation

$$\begin{aligned} -\Delta\mu &= 0 && \text{in } \Omega \\ \mu &= \mu_{\max} && \text{on } \Gamma \\ \mu &= \mu_{\min} && \text{on } \partial\Omega. \end{aligned} \tag{8.18}$$

Therefore $\mu_{\min}, \mu_{\max} \in \mathbb{R}$ influence the step size of the optimization algorithm. A small step is achieved by the choice of a large $\mu_{\max}$. Note that a_Γ then depends on the interface Γ through the parameter $\mu = \mu(\Gamma)\colon \Omega(\Gamma) \to \mathbb{R}$.

Algorithm 2: Shape optimization algorithm

1 Initialize: $\gamma_\Gamma, f_\Gamma, \Gamma_0, \bar{u}$, $k = 1$
2 **while** $\|DJ^{red}(\Gamma_k)\| > tol$ **do**
3 Interpolate $\bar{u}$ onto the current finite element mesh Ω_k
4 Assemble $\tilde{A}_\Gamma$ and solve state and adjoint equation (8.15)
5 $\rightarrow u(\Gamma_k)$, $v(\Gamma_k)$
6 Compute the mesh deformation
7 Assemble shape derivative
8 $DJ^{red}(\Gamma_k)[\mathbf{V}] = D_\Gamma L(u(\Gamma_k), \Gamma_k, v(\Gamma_k))[\mathbf{V}]$ (8.15)
9 Set $DJ^{red}(\Gamma_k)[\mathbf{V}] = 0$ for all $\mathbf{V}$ with $\operatorname{supp}(\mathbf{V}) \cap \Gamma_k = \emptyset$
10 Compute locally varying Lamé parameter by solving (8.18)
11 Assemble linear elasticity a_{Γ_k} and solve the deformation equation (8.16)
12 $\rightarrow \mathbf{U}_k$
13 Perform L-BFGS update if curvature condition is satisfied, otherwise choose gradient
14 $\rightarrow \tilde{\mathbf{U}}_k$
15 Backtracking line search (with parameters $\alpha = 1, \tau, c \in (0,1)$)
16 **while** $J^{red}(\Gamma_k - \alpha\tilde{\mathbf{U}}_k) \geq cJ^{red}(\Gamma_k)$ **do**
17 $\alpha = \tau\alpha$
18 **end while**
19 $\rightarrow \alpha_k$
20 Update mesh
21 $\Omega_{k+1} = \Omega_k - \alpha_k\tilde{\mathbf{U}}_k(\Omega_k) = \left\{\mathbf{x} - \alpha_k\tilde{\mathbf{U}}_k(\mathbf{x}) \colon \mathbf{x} \in \Omega_k\right\}$
22 $k = k + 1$
23 **end while**

How to perform the limited memory L-BFGS update in Line 13 of Algorithm 2 within the shape formalism is investigated in [72, Section 4]. Here, we only mention that the therein examined vector transport is approximated with the identity operator, so that we finally treat the gradients $\mathbf{U}_k \colon \Omega_k \to \mathbb{R}^d$ as vectors in $\mathbb{R}^{d|\Omega_k|}$ and implement the standard L-BFGS update [70, Section 5].

8.3 Nonlocal shape optimization

In Section 8.2 we have depicted the optimization methodology that we follow in this chapter to numerically solve the constrained shape optimization problem (8.9). As pointed out, the missing piece to implement the respective algorithmic realization presented in Subsection 8.2.3 is the shape derivative of the nonlocal bilinear form. We now compute this derivative for a certain class of mixed kernels and insert this formula into the saddle point system (8.14) in order to determine the shape derivative $DJ^{red}(\Gamma)[\mathbf{V}]$ of the reduced objective functional, which is used in Line 8 of the optimization Algorithm 2.

8.3.1 Shape derivative of the nonlocal bilinear form

For a fixed pair of state and adjoint variables (u, v) we have to consider the shape functional

$$\Gamma \mapsto A_\Gamma(u,v) = \sum_{i,j=1,2} \int_{\Omega_i} v(\mathbf{x}) \int_{\Omega_j \cap S(\mathbf{x})} \left(u(\mathbf{x})\phi_{ij}(\mathbf{x},\mathbf{y}) - u(\mathbf{y})\phi_{ji}(\mathbf{y},\mathbf{x})\right) d\mathbf{y} d\mathbf{x}.$$

For a given vector field $\mathbf{V} \in C_0^k(\Omega, \mathbb{R}^d)$ we now want to compute its shape derivative, which is defined by

$$D_\Gamma A_\Gamma(u,v)[\mathbf{V}] = \frac{d}{dt}\Big|_{t=0^+} A_{F_t(\Gamma)}(u,v).$$

Due to $F_t(\Omega \cup \Omega_I) = F_t(\Omega_1) \cup F_t(\Gamma) \cup F_t(\Omega_2)$ we find for the perturbed bilinear form

$$A_{F_t(\Gamma)}(u,v) = \sum_{i,j=1,2} \int_{F_t(\Omega_i)} \int_{F_t(\Omega_j)\cap S(\mathbf{x})} \psi_{ij}(\mathbf{x},\mathbf{y}) d\mathbf{y} d\mathbf{x},$$

where we define

$$\psi_{ij}(\mathbf{x},\mathbf{y}) := \psi_{ij,(u,v)}(\mathbf{x},\mathbf{y}) := v(\mathbf{x})\left(u(\mathbf{x})\phi_{ij}(\mathbf{x},\mathbf{y}) - u(\mathbf{y})\phi_{ji}(\mathbf{y},\mathbf{x})\right).$$

We aim to proceed in the same double-stage fashion as we have done to calculate the shape derivative of the right-hand side ℓ_Γ in Subsection 8.2.1. Thus we first apply transformation formula in order to shift the dependency on the parameter t from the domain of integration to the integrand and then interchange the order of differentiation and integration. Since we deal with a double integral, let us first apply transformation formula to the outer integral, which yields

$$A_{F_t(\Gamma)}(u,v) = \sum_{i,j=1,2} \int_{\Omega_i} \int_{F_t(\Omega_j)\cap S(F_t(\mathbf{x}))} \psi_{ij}(F_t(\mathbf{x}),\mathbf{y}) \left|\det dF_t(\mathbf{x})\right| d\mathbf{y} d\mathbf{x}. \tag{8.19}$$

Now we are confronted with the following difficulty. We cannot straightforwardly apply transformation formula a second time to the inner integral due to the complicated integration domain $F_t(\Omega_j) \cap S(F_t(\mathbf{x}))$. In order to proceed in this way we would either need to represent this domain as $F_t(\Omega_j) \cap S(F_t(\mathbf{x})) = F_t(M(\mathbf{x}))$ for some appropriate set $M(\mathbf{x})$ or, after changing the order of differentiation and integration only for the outer integral, we would need to compute variations in t for the function

$$t \mapsto \int_{F_t(\Omega_j)\cap S(F_t(\mathbf{x}))} \psi_{ij}(F_t(\mathbf{x}),\mathbf{y}) \left|\det dF_t(\mathbf{x})\right| d\mathbf{y}.$$

In a general setting it is not straightforward to derive analytically tractable formulas for one of the two tasks. Therefore, in the following we restrict to a specific class of mixed kernels which avoid this difficulty. Before doing so, we first need to uncover a relation between the shape derivative of the adjoint bilinear form and the shape derivative of the

original bilinear form, which proves useful to identify such kernels. Recall that we have $A_\Gamma^* = A_\Gamma'$, see (8.5), so that we find for the corresponding perturbed bilinear forms

$$A_{F_t(\Gamma)}(u,v) = A^*_{F_t(\Gamma)}(v,u),$$

where

$$A^*_\Gamma(u,v) = \sum_{i,j=1,2} \int_{\Omega_i} \int_{\Omega_j \cap S(\mathbf{x})} v(\mathbf{x})\,(u(\mathbf{x}) - u(\mathbf{y}))\,\phi_{ij}(\mathbf{x},\mathbf{y}) d\mathbf{y} d\mathbf{x}.$$

Therefore, if we can find an expression for the shape derivative of $A^*_\Gamma(u,v)$ we simultaneously have one for $A_\Gamma(u,v)$ ensured by the relation

$$\frac{d}{dt}\Big|_{t=0^+} A_{F_t(\Gamma)}(u,v) = \frac{d}{dt}\Big|_{t=0^+} A^*_{F_t(\Gamma)}(v,u). \tag{8.20}$$

The crucial advantage of considering the adjoint bilinear form with regard to our goal relies on the fact that it only involves the partial kernels ϕ_{ij} with unswapped arguments.

We are now in a position to identify tractable mixed kernels. In fact, for a kernel which depends on the interface only through its first argument, so that $\phi_{ij} = \phi_{ii}$, we find

$$A^*_\Gamma(u,v) = \sum_{i=1,2}\sum_{j=1,2} \int_{\Omega_i} \int_{\Omega_j \cap S(\mathbf{x})} \psi^*_{ij}(\mathbf{x},\mathbf{y}) d\mathbf{y} d\mathbf{x} = \sum_{i=1,2} \int_{\Omega_i} \int_{S(\mathbf{x})} \psi^*_{ii}(\mathbf{x},\mathbf{y}) d\mathbf{y} d\mathbf{x}, \tag{8.21}$$

where we abuse notation and define

$$\psi^*_{ij}(\mathbf{x},\mathbf{y}) := v(\mathbf{x})\,(u(\mathbf{x}) - u(\mathbf{y}))\,\phi_{ij}(\mathbf{x},\mathbf{y}).$$

Consequently, perturbations of the interface only affect the outer integral and thereby avoiding the central difficulty arising in the general setting (8.19) pointed out in the beginning of this subsection. More precisely, for points $\mathbf{x}$ within the interaction domain Γ_I of the interface we do not have to cope with partially weighted balls $S(\mathbf{x})$. In Figure 8.3 we find a zoom-in onto to the nonlocal interface illustrating the difference between the general setting and this specialized setting.
Now let us have a look at the perturbed adjoint bilinear form (see also (8.21)). Applying transformation formula to the outer integral yields

$$A^*_{F_t(\Gamma)}(u,v) = \sum_{i=1,2} \int_{\Omega_i} \int_{S(F_t(\mathbf{x}))} \psi^*_{ii}(F_t(\mathbf{x}),\mathbf{y})\,|\det dF_t(\mathbf{x})|\,d\mathbf{y} d\mathbf{x}.$$

We now assume that the family of interaction sets is translation invariant, so that $S(F_t(\mathbf{x})) = F_t(\mathbf{x}) + S(\mathbf{0})$. Then we find

$$A^*_{F_t(\Gamma)}(u,v) = \sum_{i=1,2} \int_{\Omega_i} \int_{S(\mathbf{0})} \psi^*_{ii}(F_t(\mathbf{x}),F_t(\mathbf{x})+\mathbf{y})\,|\det dF_t(\mathbf{x})|\,d\mathbf{y} d\mathbf{x}.$$

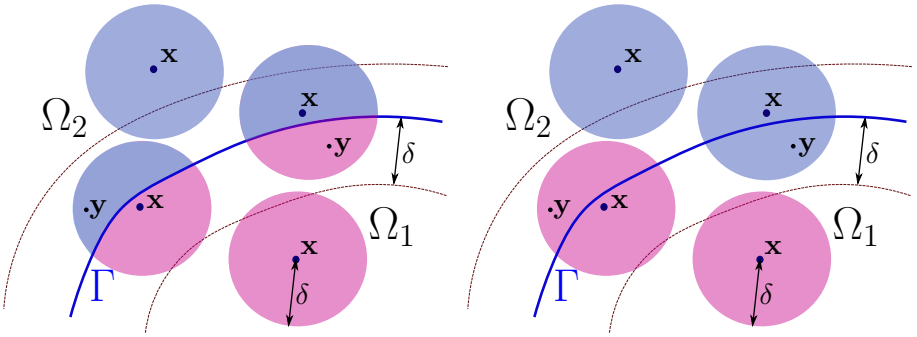


Figure 8.3: Zoom-in to the interface Γ (blue line) for a general mixed kernel (left) and for a kernel with interface-dependency only in $\mathbf{x}$. Here, both are truncated by an Euclidean ball of radius $\delta > 0$. The fundamental difference in the right situation relies on the fact, that for points within the nonlocal interface (i.e., for points within the dashed lines) nonlocal interactions are now weighted by the same partial kernel as indicated by the colors used to fill the balls.

Finally, computing the time derivative by interchanging the order of differentiation and integration and assuming that the partial kernels ϕ_{ii} are differentiable, finally leads to

$$\begin{aligned} D_\Gamma A^*_\Gamma(u,v)[\mathbf{V}] &= \frac{d}{dt}\Big|_{t=0^+} A^*_{F_t(\Gamma)}(u,v) \\ &= \frac{d}{dt}\Big|_{t=0^+} \sum_{i=1,2} \int_{\Omega_i} \int_{S(\mathbf{0})} \psi^*_{ii}(F_t(\mathbf{x}), F_t(\mathbf{x})+\mathbf{y}) \left|\det dF_t(\mathbf{x})\right| d\mathbf{y}d\mathbf{x} \\ &= \sum_{i=1,2} \int_{\Omega_i} \int_{S(\mathbf{0})} \frac{d}{dt}\Big|_{t=0^+} \left(\psi^*_{ii}(F_t(\mathbf{x}), F_t(\mathbf{x})+\mathbf{y}) \left|\det dF_t(\mathbf{x})\right|\right) d\mathbf{y}d\mathbf{x} \\ &= \sum_{i=1,2} \int_{\Omega_i} \int_{S(\mathbf{x})} \left(\nabla_{\mathbf{x}}\psi^*_{ii}(\mathbf{x},\mathbf{y}) + \nabla_{\mathbf{y}}\psi^*_{ii}(\mathbf{x},\mathbf{y})\right)^T \mathbf{V}(\mathbf{x}) + \psi_{ii}(\mathbf{x},\mathbf{y}) \operatorname{div}\mathbf{V}(\mathbf{x}) d\mathbf{y}d\mathbf{x}. \end{aligned} \tag{8.22}$$

We want to remark that the transformation formula is applicable since the perturbations F_t are diffeomorphic and the integrands are integrable. The theorem about the differentiation of parameter integrals is applicable since the integrands are piecewise H^1 on the subdomains, because we consider the perturbed nonlocal model. We proceed computing the gradients of the integrand. Recalling that $\psi^*_{ii}(\mathbf{x},\mathbf{y}) = v(\mathbf{x})\left(u(\mathbf{x}) - u(\mathbf{y})\right)\phi_{ii}(\mathbf{x},\mathbf{y})$ we find

$$\begin{aligned} &\nabla_{\mathbf{x}}\psi^*_{ii}(\mathbf{x},\mathbf{y}) \\ &\quad = \nabla v(\mathbf{x})(u(\mathbf{x}) - u(\mathbf{y}))\phi_{ii}(\mathbf{x},\mathbf{y}) + v(\mathbf{x})\nabla u(\mathbf{x})\phi_{ii}(\mathbf{x},\mathbf{y}) + v(\mathbf{x})(u(\mathbf{x}) - u(\mathbf{y}))\nabla_{\mathbf{x}}\phi_{ii}(\mathbf{x},\mathbf{y}) \end{aligned}$$

and

$$\nabla_{\mathbf{y}}\psi^*_{ii}(\mathbf{x},\mathbf{y}) = -v(\mathbf{x})\nabla u(\mathbf{y})\phi_{ii}(\mathbf{x},\mathbf{y}) + v(\mathbf{x})(u(\mathbf{x}) - u(\mathbf{y}))\nabla_{\mathbf{y}}\phi_{ii}(\mathbf{x},\mathbf{y}).$$

Thus by combining these two formulas we arrive at

$$\begin{aligned}\nabla_{\mathbf{x}}\psi_{ii}^*(\mathbf{x},\mathbf{y}) + \nabla_{\mathbf{y}}\psi_{ii}^*(\mathbf{x},\mathbf{y}) =& \nabla v(\mathbf{x})(u(\mathbf{x}) - u(\mathbf{y}))\phi_{ii}(\mathbf{x},\mathbf{y}) \\ &+ v(\mathbf{x})(\nabla u(\mathbf{x}) - \nabla u(\mathbf{y}))\phi_{ii}(\mathbf{x},\mathbf{y}) \\ &+ v(\mathbf{x})(u(\mathbf{x}) - u(\mathbf{y}))\left(\nabla_{\mathbf{x}}\phi_{ii}(\mathbf{x},\mathbf{y}) + \nabla_{\mathbf{y}}\phi_{ii}(\mathbf{x},\mathbf{y})\right).\end{aligned}$$

By inserting this expression into (8.22) we find

$$\begin{aligned}&D_\Gamma A_\Gamma^*(u,v)[\mathbf{V}] \\ &= \sum_{i=1,2}\int_{\Omega_i}\int_{S(\mathbf{x})} \left(\nabla_{\mathbf{x}}\psi_{ii}^* + \nabla_{\mathbf{y}}\psi_{ii}^*\right)^T \mathbf{V} + \psi_{ii}\operatorname{div}\mathbf{V}\; d\mathbf{y}d\mathbf{x} \\ &= \sum_{i=1,2}\Bigg(\int_{\Omega_i}\int_{S(\mathbf{x})} \nabla v^T\mathbf{V}(u-u')\phi_{ii}\; d\mathbf{y}d\mathbf{x} \\ &\qquad + \int_{\Omega_i}\int_{S(\mathbf{x})} v(\nabla u - \nabla u')^T\mathbf{V}\phi_{ii}\; d\mathbf{y}d\mathbf{x} \\ &\qquad + \int_{\Omega_i}\int_{S(\mathbf{x})} v(u-u')\left(\nabla_{\mathbf{x}}\phi_{ii} + \nabla_{\mathbf{y}}\phi_{ii}\right)^T\mathbf{V}\; d\mathbf{y}d\mathbf{x} \\ &\qquad + \int_{\Omega_i}\int_{S(\mathbf{x})} v\left(u-u'\right)\phi_{ii}\operatorname{div}\mathbf{V}\; d\mathbf{y}d\mathbf{x}\Bigg) \\ &= A_\Gamma^*(u,\dot v) + A_\Gamma^*(\dot u, v) \\ &\quad + \sum_{i=1,2}\int_{\Omega_i}\int_{S(\mathbf{x})} v(u-u')\left(\nabla_{\mathbf{x}}\phi_{ii} + \nabla_{\mathbf{y}}\phi_{ii}\right)^T\mathbf{V}\; d\mathbf{y}d\mathbf{x} \\ &\quad + \sum_{i=1,2}\int_{\Omega_i}\int_{S(\mathbf{x})} v\left(u-u'\right)\phi_{ii}\operatorname{div}\mathbf{V}\; d\mathbf{y}d\mathbf{x}.\end{aligned}$$

Then by (8.20) we eventually arrive at

$$\begin{aligned}D_\Gamma A_\Gamma(u,v)[\mathbf{V}] =\; & A_\Gamma(\dot u, v) + A_\Gamma(u,\dot v) \\ &+ \sum_{i=1,2}\int_{\Omega_i}\int_{S(\mathbf{x})} u(v-v')\left(\nabla_{\mathbf{x}}\phi_{ii} + \nabla_{\mathbf{y}}\phi_{ii}\right)^T\mathbf{V}\; d\mathbf{y}d\mathbf{x} \\ &+ \sum_{i=1,2}\int_{\Omega_i}\int_{S(\mathbf{x})} u\left(v-v'\right)\phi_{ii}\operatorname{div}\mathbf{V}\; d\mathbf{y}d\mathbf{x}.\end{aligned} \tag{8.23}$$

Further simplifications arise in the case of piecewise smooth and radial kernel functions $\phi_{ii}(\mathbf{x},\mathbf{y}) = \phi_{ii}(\|\mathbf{x}-\mathbf{y}\|_2)$. Here, we find

$$\nabla_{\mathbf{x}}\phi_{ii}(\mathbf{x},\mathbf{y}) = \phi_{ii}'(\|\mathbf{x}-\mathbf{y}\|_2)\frac{1}{2}\frac{1}{\|\mathbf{x}-\mathbf{y}\|_2}2(\mathbf{x}-\mathbf{y})\cdot 1 = \phi_{ii}'(\|\mathbf{x}-\mathbf{y}\|_2)\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|_2}$$

and similarly

$$\nabla_{\mathbf{y}}\phi_{ii}(\mathbf{x},\mathbf{y}) = -\phi_{ii}'(\|\mathbf{x}-\mathbf{y}\|_2)\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|_2}.$$

Thus, we have $\nabla_{\mathbf{x}}\phi_{ii}(\mathbf{x},\mathbf{y}) + \nabla_{\mathbf{y}}\phi_{ii}(\mathbf{x},\mathbf{y}) = 0$ and therefore

$$D_\Gamma A_\Gamma(u,v)[\mathbf{V}] = A_\Gamma(\dot{u},v) + A_\Gamma(u,\dot{v}) + \sum_{i=1,2}\int_{\Omega_i}\int_{S(\mathbf{x})} u\,(v-v')\,\phi_{ii}\,\mathrm{div}\,\mathbf{V}\;d\mathbf{y}d\mathbf{x}.$$

For reasons of clarity let us summarize all the assumptions on the kernel that we have formulated on the way to arrive at the shape derivative (8.23):

- **Interface dependency only in $\mathbf{x}$.** We assume that the mixed kernel depends on the interface solely through its first argument, i.e., we consider kernels of the form

$$\gamma_\Gamma(\mathbf{x},\mathbf{y}) = \begin{cases} \gamma_{11}(\mathbf{x},\mathbf{y}): & \mathbf{x}\in\overline{\Omega}_1 \\ \gamma_{22}(\mathbf{x},\mathbf{y}): & \mathbf{x}\in\Omega_2 = (\Omega\cup\Omega_I)\backslash\overline{\Omega}_1. \end{cases} \tag{8.24}$$

- **Truncation by translation invariant interaction sets.** We require the partial kernels to be truncated by the same family of interaction sets $\{S(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$, which we assume to be translation invariant, i.e., $S(\mathbf{x}) = \mathbf{x} + S(\mathbf{0})$.

- **Differentiability of partial kernels.** We assume that the partial kernels ϕ_{ii}, $i = 1, 2$, are weakly differentiable.

8.3.2 Shape derivative of the reduced objective functional

Finally, it remains to insert expression (8.23) into equation (8.15) in order to derive an expression for the shape derivative of the reduced objective functional. We obtain

$$\begin{aligned} DJ^{red}(\Gamma)[\mathbf{V}] = \nu\int_\Gamma \mathrm{div}_\Gamma\,\mathbf{V}\;ds - \left(\int_\Omega \dot{f}_\Gamma v\;d\mathbf{x} + \int_\Omega f_\Gamma v\;\mathrm{div}\,\mathbf{V}\;d\mathbf{x} + \ell_\Gamma(\dot{v})\right) \\ + A_\Gamma(u,\dot{v}) + A_\Gamma(\dot{u},v) + \sum_{i=1,2}\int_{\Omega_i}\int_{S(\mathbf{x})} u\,(v-v')\,\phi_{ii}\,\mathrm{div}\,\mathbf{V}\;d\mathbf{y}d\mathbf{x} \\ + \sum_{i=1,2}\int_{\Omega_i}\int_{S(\mathbf{x})} u(v-v')\,(\nabla_{\mathbf{x}}\phi_{ii} + \nabla_{\mathbf{y}}\phi_{ii})^T\,\mathbf{V}\;d\mathbf{y}d\mathbf{x}. \end{aligned}$$

By exploiting both, the state and adjoint equation (8.15), as well as $A_\Gamma(\dot{u},v) = A_\Gamma^*(v,\dot{u})$, we obtain

$$\begin{aligned} DJ^{red}(\Gamma)[\mathbf{V}] = \nu\int_\Gamma \mathrm{div}_\Gamma\,\mathbf{V}\;ds - \int_\Omega \dot{f}_\Gamma v\;d\mathbf{x} - \int_\Omega f_\Gamma v\;\mathrm{div}\,\mathbf{V}\;d\mathbf{x} \\ - c_{per}A^{loc}(u,\dot{v}) + \ell^*(\dot{u}) - c_{per}A^{loc}(v,\dot{u}) \\ + \sum_{i=1,2}\int_{\Omega_i}\int_{S(\mathbf{x})} u\,(v-v')\,\phi_{ii}\,\mathrm{div}\,\mathbf{V}\;d\mathbf{y}d\mathbf{x} \\ + \sum_{i=1,2}\int_{\Omega_i}\int_{S(\mathbf{x})} u(v-v')\,(\nabla_{\mathbf{x}}\phi_{ii} + \nabla_{\mathbf{y}}\phi_{ii})^T\,\mathbf{V}\;d\mathbf{y}d\mathbf{x}. \end{aligned} \tag{8.25}$$

Now we exploit the following formula for the shape derivative of the local bilinear form (see, e.g., [92, Theorem 4.21])

$$\begin{aligned} &D_\Gamma A^{loc}(u,v) \\ &\quad = A^{loc}(u,\dot{v}) + A^{loc}(\dot{u},v) + \int_\Omega \nabla u^T \nabla v \ \operatorname{div} \mathbf{V} d\mathbf{x} - \int_\Omega \nabla u^T \left(\nabla \mathbf{V} + \nabla \mathbf{V}^T\right) \nabla v d\mathbf{x}. \end{aligned}$$

Since A^{loc} is independent of the interface, so that $D_\Gamma A^{loc}(u,v) = 0$, we obtain

$$-\left(A^{loc}(u,\dot{v}) + A^{loc}(\dot{u},v)\right) = \int_\Omega \nabla u^T \nabla v \ \operatorname{div} \mathbf{V} d\mathbf{x} - \int_\Omega \nabla u^T \left(\nabla \mathbf{V} + \nabla \mathbf{V}^T\right) \nabla v d\mathbf{x}.$$

By inserting this equation into (8.25) we eventually arrive at

$$\begin{aligned} DJ^{red}(\Gamma)[\mathbf{V}] = \nu \int_\Gamma \operatorname{div}_\Gamma \mathbf{V} \ ds - \int_\Omega \dot{f}_\Gamma v \ d\mathbf{x} - \int_\Omega f_\Gamma v \ \operatorname{div} \mathbf{V} \ d\mathbf{x} + \ell^*(\dot{u}) \\ + c_{per} \left(\int_\Omega \nabla u^T \nabla v \ \operatorname{div} \mathbf{V} d\mathbf{x} - \int_\Omega \nabla u^T \left(\nabla \mathbf{V} + \nabla \mathbf{V}^T\right) \nabla v d\mathbf{x} \right) \\ + \sum_{i=1,2} \int_{\Omega_i} \int_{S(\mathbf{x})} u \left(v - v'\right) \phi_{ii} \ \operatorname{div} \mathbf{V} \ d\mathbf{y} d\mathbf{x} \\ + \sum_{i=1,2} \int_{\Omega_i} \int_{S(\mathbf{x})} u(v - v') \left(\nabla_{\mathbf{x}} \phi_{ii} + \nabla_{\mathbf{y}} \phi_{ii}\right)^T \mathbf{V} \ d\mathbf{y} d\mathbf{x}. \end{aligned} \tag{8.26}$$

If f_Γ is piecewise constant, i.e., $f_i = const$, $\nu = 0$, and if the partial kernel functions ϕ_{ii} are radial, this further simplifies to

$$\begin{aligned} DJ^{red}(\Gamma)[\mathbf{V}] = &- \int_\Omega f_\Gamma v \ \operatorname{div} \mathbf{V} \ d\mathbf{x} - \int_\Omega (u - \bar{u}) \ \nabla u^T \mathbf{V} \ d\mathbf{x} \\ &+ c_{per} \left(\int_\Omega \nabla u^T \nabla v \ \operatorname{div} \mathbf{V} \ d\mathbf{x} - \int_\Omega \nabla u^T \left(\nabla \mathbf{V} + \nabla \mathbf{V}^T\right) \nabla v d\mathbf{x} \right) \\ &+ \sum_{i=1,2} \int_{\Omega_i} \int_{S(\mathbf{x})} u \left(v - v'\right) \phi_{ii} \ \operatorname{div} \mathbf{V} d\mathbf{y} d\mathbf{x}. \end{aligned}$$

We observe that, if $\nu = 0$ and $u = \bar{u}$, then $\ell^*(\cdot) = 0$, which implies $v = (-\mathcal{L}_\Gamma^*)^{-1}(u - \bar{u}) = 0$, and thus $DJ^{red}(\Gamma)[\mathbf{V}] = 0$ in a saddle point where $u = \bar{u}$. In other words, if there is a shape Γ so that $\bar{u} = u(\Gamma)$, then this shape is a stationary point of the reduced objective functional.

8.3.3 Some thoughts on the unperturbed nonlocal model

In this subsection we want to highlight some aspects of the unperturbed nonlocal interface problem with the specialized kernel from (8.24), i.e.,

$$\gamma_\Gamma(\mathbf{x},\mathbf{y}) = \gamma_{11}(\mathbf{x},\mathbf{y})\chi_{\Omega_1}(\mathbf{x}) + \gamma_{22}(\mathbf{x},\mathbf{y})\chi_{\Omega_2}(\mathbf{x})$$

$$= (\phi_{11}(\mathbf{x},\mathbf{y})\chi_{\Omega_1}(\mathbf{x}) + \phi_{22}(\mathbf{x},\mathbf{y})\chi_{\Omega_2}(\mathbf{x}))\,\chi_{S(\mathbf{x})}(\mathbf{y}).$$

Let us insert this kernel into the bilinear form A_Γ and partition the domain of integration $\Omega \times (\Omega \cup \Omega_I)$ according to the subdomains Ω_1 and $\Omega_2 = (\Omega \cup \Omega_I)\backslash\overline{\Omega_1}$. We obtain

$$\begin{aligned} A_\Gamma(u,v) = &\int_{\Omega_1}\int_{\Omega_1} v(u\gamma_{11} - u'\gamma'_{11})d\mathbf{y}d\mathbf{x} + \int_{\Omega_1}\int_{\Omega_2} v(u\gamma_{11} - u'\gamma'_{22})d\mathbf{y}d\mathbf{x} \\ &+ \int_{\Omega_2}\int_{\Omega_1} v(u\gamma_{22} - u'\gamma'_{11})d\mathbf{y}d\mathbf{x} + \int_{\Omega_2}\int_{\Omega_2} v(u\gamma_{22} - u'\gamma'_{22})d\mathbf{y}d\mathbf{x}. \end{aligned}$$

By an application of Fubini's theorem we further find

$$\begin{aligned} &A_\Gamma(u,v) \\ &= \left(\sum_{i=1,2}\int_{\Omega_i}\int_{\Omega_i} v(u\gamma_{ii} - u'\gamma'_{ii})d\mathbf{y}d\mathbf{x}\right) + \int_{\Omega_1}\int_{\Omega_2}(v-v')(u\gamma_{11} - u'\gamma'_{22})d\mathbf{y}d\mathbf{x} \\ &= \left(\sum_{i=1,2}\int_{\Omega_i}\int_{\Omega_i\cap S(\mathbf{x})} v(u\phi_{ii} - u'\phi'_{ii})d\mathbf{y}d\mathbf{x}\right) + \int_{\Omega_1}\int_{\Omega_2\cap S(\mathbf{x})}(v-v')(u\phi_{11} - u'\phi'_{22})d\mathbf{y}d\mathbf{x}. \end{aligned} \tag{8.27}$$

The bilinear form

$$(u,v) \mapsto \int_{\Omega_i}\int_{\Omega_i\cap S(\mathbf{x})} v(u\gamma_{ii} - u'\gamma'_{ii})d\mathbf{y}d\mathbf{x},$$

can be associated with a nonlocal model posed on the *nonlocal interior* $\overset{\circ}{\Omega}_i := \Omega_i\backslash(\Omega_i \cap \Gamma_I)$ of Ω_i with interaction domain $(\Omega_i \cap \Gamma_I)$ and kernel γ_{ii}. The last summand

$$\begin{aligned} &\int_{\Omega_1}\int_{\Omega_2\cap S(\mathbf{x})}(v-v')(u\phi_{11} - u'\phi'_{22})d\mathbf{y}d\mathbf{x} \\ &\qquad = \int_{\Omega_1\cap\Gamma_I}\int_{\Omega_2\cap\Gamma_I\cap S(\mathbf{x})}(v-v')(u\phi_{11} - u'\phi'_{22})d\mathbf{y}d\mathbf{x} \end{aligned}$$

can then be interpreted as the coupling between these two nonlocal models.

Counterexample for the coercivity of A_Γ

Let us consider the simple piecewise constant kernel

$$\gamma_\Gamma(\mathbf{x},\mathbf{y}) := \begin{cases} c_1 & : \mathbf{x} \in \Omega_1 \\ c_2 & : \mathbf{x} \in \Omega_2 \end{cases}$$

and a piecewise constant function $u\colon \Omega \cup \Omega_I \to \mathbb{R}$ defined by

$$u(\mathbf{x}) := \begin{cases} u_1 & : \mathbf{x} \in \Omega_1 \\ u_2 & : \mathbf{x} \in \Omega_2. \end{cases}$$

Then by (8.27) we find that

$$\begin{aligned} A_\Gamma(u,u) &= \sum_{i=1,2} \int_{\Omega_i} \int_{\Omega_i \cap S(\mathbf{x})} u(u\phi_{ii} - u'\phi'_{ii}) d\mathbf{y} d\mathbf{x} \\ &\qquad + \int_{\Omega_1} \int_{\Omega_2 \cap S(\mathbf{x})} (u-u')(u\phi_{11} - u'\phi'_{22}) d\mathbf{y} d\mathbf{x} \\ &= (u_1 - u_2)(u_1 c_1 - u_2 c_2) \int_{\Omega_1} \int_{\Omega_2 \cap S(\mathbf{x})} d\mathbf{y} d\mathbf{x}. \end{aligned}$$

Hence, if we choose u_i and c_i in such a way that

$$1 < \frac{u_1}{u_2} < \frac{c_2}{c_1}$$

we find

$$(u_1 - u_2)(u_1 c_1 - u_2 c_2) < 0 \tag{8.28}$$

implying that $A_\Gamma(u,u) < 0$ so that the nonlocal bilinear form for this kernel is not coercive.

Remark 8.3.1. *As a consequence of* (8.28) *we cannot rely on the Lax-Milgram theorem for establishing a well-posedness result for the nonlocal interface problem without further assumptions on the kernel.*

Towards a well-posedness proof via Fredholm alternative

Let us assume that the partial kernels are integrable in such a way that $\int_{\mathbb{R}^d} \gamma(\mathbf{x},\mathbf{y}) d\mathbf{y}$ exists for all $\mathbf{x} \in \Omega$, then we can rewrite the nonlocal convection-diffusion operator as

$$-\mathcal{L}_\Gamma u(\mathbf{x}) = \int_{\mathbb{R}^d} (u\gamma_\Gamma - u'\gamma'_\Gamma) d\mathbf{y} = u \int_{\mathbb{R}^d} \gamma_\Gamma d\mathbf{y} - \int_{\mathbb{R}^d} u'\gamma'_\Gamma d\mathbf{y} =: h(\mathbf{x})u(\mathbf{x}) - \mathcal{K}u(\mathbf{x}),$$

where we define the function $h\colon \Omega \to \mathbb{R}$ by

$$h(\mathbf{x}) = \int_{\mathbb{R}^d} \gamma_\Gamma(\mathbf{x},\mathbf{y}) d\mathbf{y} = \begin{cases} h_1(\mathbf{x}) := \int_{S(\mathbf{x})} \phi_{11}(\mathbf{x},\mathbf{y}) d\mathbf{y} & : \mathbf{x} \in \Omega_1 \\ h_2(\mathbf{x}) := \int_{S(\mathbf{x})} \phi_{22}(\mathbf{x},\mathbf{y}) d\mathbf{y} & : else \end{cases}$$

and the operator $\mathcal{K}\colon L^2(\Omega \cup \Omega_I) \to L^2(\Omega \cup \Omega_I)$ by

$$\mathcal{K}u(\mathbf{x}) := \int_{\mathbb{R}^d} u'\gamma'_\Gamma d\mathbf{y} = \int_{\Omega_1} u(\mathbf{y})\gamma_{11}(\mathbf{y},\mathbf{x}) d\mathbf{y} + \int_{\Omega_2} u(\mathbf{y})\gamma_{22}(\mathbf{y},\mathbf{x}) d\mathbf{y}.$$

In order to further specify the operator $\mathcal{K}$ let us assume $\gamma_{ii} \in L^2(\Omega \cup \Omega_I \times \Omega \cup \Omega_I)$ so that

$$\int_{\Omega\cup\Omega_I} \int_{\Omega\cup\Omega_I} \gamma_\Gamma^2 d\mathbf{y} d\mathbf{x} = \int_{\Omega_1} \int_{\Omega\cup\Omega_I} \gamma_{11}^2 d\mathbf{y} d\mathbf{x} + \int_{\Omega_2} \int_{\Omega\cup\Omega_I} \gamma_{22}^2 d\mathbf{y} d\mathbf{x} < \infty$$

which implies $\gamma_\Gamma \in L^2(\Omega \cup \Omega_I \times \Omega \cup \Omega_I)$. In this case, $\mathcal{K}\colon L^2(\Omega \cup \Omega_I) \to L^2(\Omega \cup \Omega_I)$ is well defined and compact; a so-called Hilbert-Schmidt operator [94, Section VI.6]. Consequently, we can interpret the nonlocal convection-diffusion operator $\mathcal{L}_\Gamma$ as a compact perturbation of a scalar multiple of identity which in turn allows us to apply Fredholm theory. In particular, the Fredholm alternative teaches us that the surjectivity of such operators is already implied by their injectivity. In other words, if the homogeneous equation $\mathcal{L}_\Gamma u = 0$ has the unique solution $u = 0$, then $\mathcal{L}_\Gamma u = f$ has a unique solution for all $f \in L^2(\Omega)$. Thus, by applying Fredholm alternative to our specific setting, we find the following sufficient condition for the well-posedness of our interface problem:

$$h(\mathbf{x})u(\mathbf{x}) = \left(\int_{\Omega_1} u(\mathbf{y})\gamma_{11}(\mathbf{y},\mathbf{x})d\mathbf{y} + \int_{\Omega_2} u(\mathbf{y})\gamma_{22}(\mathbf{y},\mathbf{x})d\mathbf{y}\right) \quad \text{for all } \mathbf{x} \in \Omega$$
$$\Rightarrow \quad u = 0.$$

The derivation of sufficient conditions on the kernel for this statement to hold is left to future work.

Regularity of solutions

In order to gain some deeper insights into the nonlocal interface problem at hand we want to highlight two properties that potential solutions necessarily satisfy. For this purpose we make the following assumptions.

- We assume that the interaction sets are translation invariant.
- We assume that the partial kernel functions are radial with respect to some norm $\|\cdot\|$ in $\mathbb{R}^d$ so that $\phi_{ii}(\mathbf{x},\mathbf{y}) = \phi_{ii}(\|\mathbf{x}-\mathbf{y}\|) > 0$. We further require that $\phi_{ii}(\|\cdot\|)$ is continuous on the closure of $\overline{S(\mathbf{0})}$ and thereby uniformly continuous.
- We assume that the solution is square integrable, i.e., $u \in L^2(\Omega)$.
- For the the forcing term f, we assume that f_i is continuous on Ω_i with continuous extension to $\overline{\Omega}_i$.

In this case we find that h is piecewise constant because

$$h_i = \int_{S(\mathbf{x})} \phi_{ii}(\|\mathbf{x}-\mathbf{y}\|)d\mathbf{y} = \int_{S(\mathbf{0})} \phi_{ii}(\|\mathbf{z}\|)d\mathbf{z}.$$

Since the kernel functions are assumed to be positive we have $h_i > 0$ and we find for $\mathbf{x} \in \Omega_i$ that the homogeneous solutions takes the form

$$u_i(\mathbf{x}) := u|_{\Omega_i}(\mathbf{x}) = \frac{1}{h_i}\mathcal{K}u(\mathbf{x}). \tag{8.29}$$

We now show that $\mathbf{x} \mapsto \mathcal{K}u(\mathbf{x})$ is continuous on Ω, so that u_i, for $i = 1, 2$ is continuous on the respective subdomain Ω_i. Therefore let $\epsilon > 0$ and $\mathbf{x}, \mathbf{y} \in \Omega$, then

$$|\mathcal{K}u(\mathbf{x}) - \mathcal{K}u(\mathbf{y})| = \left|\int_{\Omega_1} u(\mathbf{z})\gamma_{11}(\|\mathbf{x}-\mathbf{z}\|)d\mathbf{z} + \int_{\Omega_2} u(\mathbf{z})\gamma_{22}(\|\mathbf{x}-\mathbf{z}\|)d\mathbf{z}\right.$$

$$
\begin{aligned}
&\left| - \int_{\Omega_1} u(\mathbf{z})\gamma_{11}(\|\mathbf{y}-\mathbf{z}\|)d\mathbf{z} + \int_{\Omega_2} u(\mathbf{z})\gamma_{22}(\|\mathbf{y}-\mathbf{z}\|)d\mathbf{z} \right| \\
&\leq \sum_{i=1,2} \int_{\Omega_i \cap S(\mathbf{x})} |u(\mathbf{z})|\, |\phi_{ii}(\|\mathbf{x}-\mathbf{z}\|) - \phi_{ii}(\|\mathbf{y}-\mathbf{z}\|)|\, d\mathbf{z} \\
&\leq \sum_{i=1,2} \int_{S(\mathbf{x})} |u(\mathbf{z})|\, |\phi_{ii}(\|\mathbf{x}-\mathbf{z}\|) - \phi_{ii}(\|\mathbf{y}-\mathbf{z}\|)|\, d\mathbf{z}.
\end{aligned}
$$

For $\|\mathbf{y}-\mathbf{x}\| < \delta$ we find by the triangle inequality that $|\|\mathbf{x}-\mathbf{z}\| - \|\mathbf{y}-\mathbf{z}\|| < \delta$ for $\mathbf{z} \in S(\mathbf{x})$. Due to the uniform continuity of $\phi_{ii}(\cdot)$ we find for some $\epsilon_i > 0$ a corresponding $\delta_i > 0$ such that, for $\|\mathbf{x}-\mathbf{y}\| < \delta_i$, and for all $\mathbf{z} \in S(\mathbf{x})$,

$$
|\phi_{ii}(\|\mathbf{x}-\mathbf{z}\|) - \phi_{ii}(\|\mathbf{y}-\mathbf{z}\|)| < \epsilon_i
$$

and therefore

$$
\begin{aligned}
&\sum_{i=1,2} \int_{S(\mathbf{x})} |u(\mathbf{z})|\, |\phi_{ii}(\|\mathbf{x}-\mathbf{z}\|) - \phi_{ii}(\|\mathbf{y}-\mathbf{z}\|)|\, d\mathbf{z} \\
&< \sum_{i=1,2} \epsilon_i \int_{S(\mathbf{x})} |u(\mathbf{z})| d\mathbf{z} \leq \sum_{i=1,2} \epsilon_i \sqrt{|\Omega|} \|u\|_{L^2(\Omega)} \\
&\leq \max\{\epsilon_1, \epsilon_2\}\, 2\|u\|_{L^2(\Omega)} \sqrt{|\Omega|}.
\end{aligned}
$$

Thus we can choose ϵ_i so that

$$
|\mathcal{K}u(\mathbf{x}) - \mathcal{K}u(\mathbf{y})| \leq \max\{\epsilon_1, \epsilon_2\}\, 2\|u\|_{L^2(\Omega)} \sqrt{|\Omega_1|} < \epsilon
$$

for $\|\mathbf{x}-\mathbf{y}\| < \delta := \min\{\delta_1, \delta_2\}$, which states the desired continuity property. We also want to point out that for $\mathbf{x} \in \overset{\circ}{\Omega}_i = \Omega_i \backslash (\Omega_i \cap \Gamma_I)$, we find

$$
\mathcal{K}u(\mathbf{x}) = \int_{S(\mathbf{x})} u(\mathbf{y})\phi_{ii}(\|\mathbf{x}-\mathbf{y}\|)d\mathbf{y} = \int_{\mathbb{R}^d} u(\mathbf{y})\phi_{ii}(\|\mathbf{y}-\mathbf{x}\|)\chi_{S(\mathbf{0})}(\mathbf{y}-\mathbf{x})d\mathbf{y},
$$

so that $\mathcal{K}$ restricted to the nonlocal interior $\overset{\circ}{\Omega}_i$ of Ω_i is a convolution operator in our setting. A consequence of the continuity of $\mathcal{K}u$ is the piecewise continuity of the homogeneous solution u; more precisely by (8.29) we find that u_i is continuously extendable on $\overline{\Omega}_i$. Due to the assumption on f, we also find that the inhomogeneous solution u^f is piecewise continuous due to the relation

$$
u_i^f = \frac{1}{h_i}\left(f_i - \mathcal{K}u\right).
$$

We now derive the second property, namely, if the kernels γ_{ii} differ "significantly", then we find that the potential solution u must have a jump discontinuity along the interface. To this end, let $\bar{\mathbf{x}} \in \Gamma$. Further let us consider two sequences $\{\mathbf{x}_n^i\}_{n\in\mathbb{N}} \subset \Omega_i$ converging

to $\bar{\mathbf{x}}$, where each is approaching $\bar{\mathbf{x}}$ from either Ω_1 or Ω_2. Due to the continuity of u_i^f we find

$$u^f(\mathbf{x}_n^i) \longrightarrow \bar{u}_i^f = \frac{1}{h_i}(f_i - \mathcal{K}u(\bar{\mathbf{x}})) \quad (n \to \infty)$$

and therefore

$$\left|\bar{u}_1^f - \bar{u}_2^f\right| = \left|\left(\frac{f_1(\bar{\mathbf{x}})}{h_1} - \frac{f_2(\bar{\mathbf{x}})}{h_2}\right) + \left(\frac{1}{h_1} - \frac{1}{h_2}\right)\mathcal{K}u(\bar{\mathbf{x}})\right|.$$

Thus, unless the f_i are chosen such that

$$f_1(\bar{\mathbf{x}}) = \frac{h_1}{h_2} f_2(\bar{\mathbf{x}}) + \mathcal{K}u(\bar{\mathbf{x}})\left(\frac{h_1}{h_2} - 1\right)$$

we find that

$$\left|\bar{u}_1^f - \bar{u}_2^f\right| > 0$$

implying that the limits from either sides are not the same and u has a jump discontinuity across the interface.

Example in 1d

For illustration purposes we want to present a numerical example in 1d, which confirms the theoretical assertions above. In the one-dimensional setting the interface reduces to a single point, but as a cross section, the problem still captures all the features of the higher dimensional counterpart and therefore serves as a representative study case. We now consider $\Omega = (0,1)$ as well as norm induced interaction sets with interaction horizon $\delta > 0$ and an interface point $a \in (0,1)$ which splits $\Omega \cup \Omega_I = (-\delta, 1+\delta)$ into two subdomains $\Omega_1 := (-\delta, a)$ and $\Omega_2 := (a, 1+\delta)$. Let us choose the kernel

$$\gamma_a(x,y) = \left(\phi_1(x,y)\chi_{\Omega_1}(x) + \phi_2(x,y)\chi_{\Omega_2}(x)\right)\chi_{|x-y|<\delta}$$

where

$$\phi_1(x,y) = 0.1 c_\delta$$

and

$$\phi_2(x,y) = c_\delta\left(1 - \left(\frac{|x-y|}{\delta}\right)^2\right)$$

with $c_\delta := \frac{3}{2\delta^3}$ and an interaction horizon $\delta = 0.05$. We further choose $a = 0.5$, $f_i = 1$ and homogeneous Dirichlet data $g = 0$. Then all of the above assumptions are satisfied. Namely, the partial kernels γ_i are positive, radial and uniformly continuous on the translation invariant interaction sets. The partial forcing terms are also continuous on the respective subdomains. We solve the constraint equation from (8.9) with continuous piecewise linear basis functions on a uniform grid of size $h = 0.00025$. The results are presented and commented in Figure 8.4.

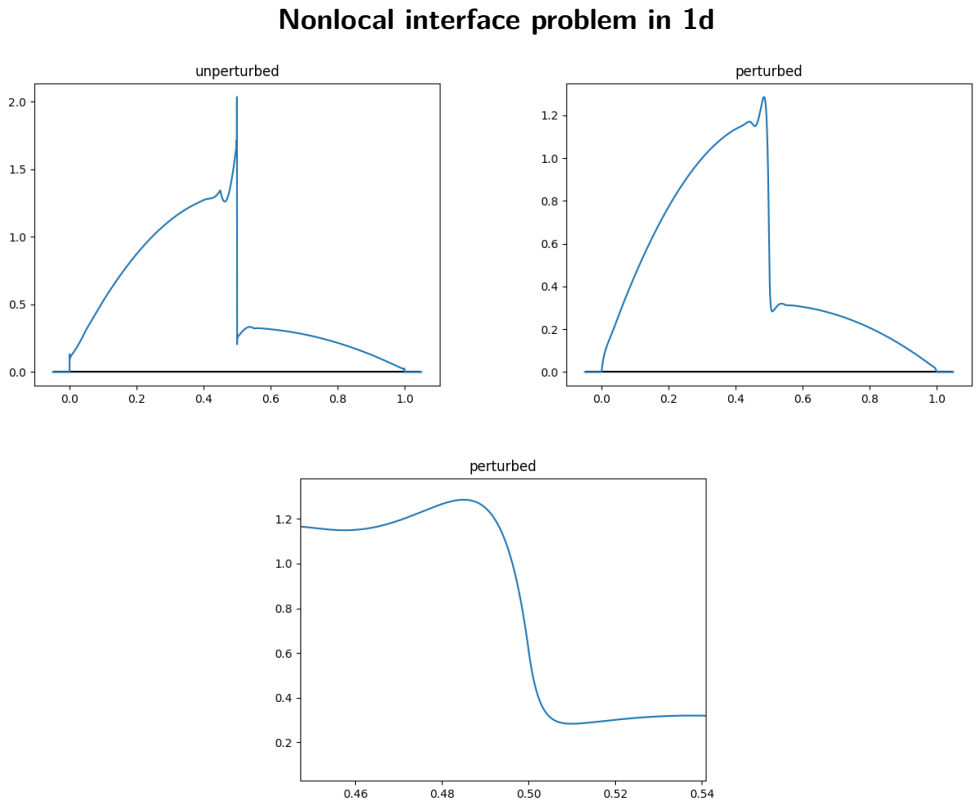


Figure 8.4: The plots show the finite element solution u^h of the nonlocal interface problem posed on $\Omega = (0, 1)$. In the top row on the left-hand side we have the unperturbed version ($c_{per} = 0$), i.e., a pure nonlocal model and on the right-hand side the perturbed version with parameter $c_{per} = 0.001$, i.e., a "small" Laplacian $-c_{per}\Delta$ is added to the nonlocal operator. The plot below is a zoom onto the nonlocal interface for the perturbed case. We first observe the predicted jump discontinuity at the interface point $a = 0.5$. Since we employ continuous basis functions this entails a slightly oscillating behavior for the unperturbed finite element solution. Although only mildly perturbed, this singular behavior is smoothened for the regularized finite element solution. Furthermore, let us interpret both solutions in the light of the introductory Section 2.1. First of all note that we have a nonlocal diffusion model on both sides of the interface, whereas close to the interface we observe convective effects. Since we employ the same source term $f_i = 1$ for $i = 1, 2$, particles are produced at the same rate on both sides. However by the choice of the partial kernel functions ϕ_{ii} we observe a slower diffusion of particles on Ω_1, i.e., on left-hand side of the interface, than on Ω_2, i.e., on the right-hand side of the interface. Therefore, the mass density u of particles on subdomain $\Omega_1 = (-0.1, 0.5)$ is larger than the one on $\Omega_2 = (0.5, 1.1)$. Particles within the nonlocal interface $(a - \delta, a + \delta) = (0.45, 0.55)$ experience an imbalanced motivation to move, thereby provoking the convective effects.

8.4 Numerical experiments

In this section, we want to put the above derived formula (8.26) for the shape derivative of the reduced objective functional into numerical practice.

In all of the following numerical examples we choose the kernel

$$\gamma_\Gamma(\mathbf{x},\mathbf{y}) = \left(\phi_1(\mathbf{x},\mathbf{y})\chi_{\Omega_1}(\mathbf{x}) + \phi_2(\mathbf{x},\mathbf{y})\chi_{\Omega_2}(\mathbf{x})\right)\chi_{B_{\delta,\infty}(\mathbf{x})}(\mathbf{y}),$$

where

$$\phi_1(\mathbf{x},\mathbf{y}) = \frac{1}{1000}c_\delta$$

and

$$\phi_2(\mathbf{x},\mathbf{y}) = 100c_\delta\left(1 - \left(\frac{\|\mathbf{y}-\mathbf{x}\|_\infty}{\delta}\right)^2\right)$$

with scaling constant $c_\delta := \frac{3}{4\delta^4}$ and an interaction horizon $\delta = 0.1$. We note that both partial kernels are radial and they are truncated by $\|\cdot\|_\infty$-balls so that $\Omega \cup \Omega_I = [-\delta, 1+\delta]^2$. We choose homogeneous Dirichlet data $g = 0$ and as right-hand side a piecewise constant function

$$f_\Gamma(\mathbf{x}) = 100\chi_{\Omega_1}(\mathbf{x}) + \chi_{(\Omega\setminus\Omega_1)}(\mathbf{x}),$$

i.e., $f_1 = 100$ and $f_2 = 1$. We do not use a perimeter regularization, i.e., we choose $\nu = 0$, so that we can make use of the formula (8.26) for the shape derivative of the reduced objective functional. We employ continuous piecewise linear basis functions on triangular grids for the discretization of the nonlocal constraint equation. For a detailed discussion on the assembly of the nonlocal stiffness matrix we refer to Chapter 4. Here we only want to amplify how to implement a mixed kernel. During the mesh generation each triangle is labeled according to its subdomain affiliation. Thus, whenever we integrate over a pair of two triangles, we can read out the labels (i,j) and choose the corresponding atomic kernel γ_{ij}.

The data $\bar{u}$ is generated as solution $u(\overline{\Gamma})$ of the constraint equation associated to a target shape $\overline{\Gamma}$. Thus the data is represented in a finite element basis and for the interpolation task in Line 3 of Algorithm 2 we solely need to translate between (non-matching) finite element grids (we use `scipy.interpolate.griddata` for this). In all examples below the target shape $\overline{\Gamma}$ is chosen to be a circle of radius 0.25 centered at $(0.5, 0.5)$.

We now present three different examples which differ in the choice of the initial guess Γ_0 and perturbation parameter c_{per}. They are presented and described in the Figures 8.5, 8.6 and 8.7. In each plot of the aforementioned Figures the red line represents the target interface $\overline{\Gamma}$. The black line represents the initial guess and the blue ones the shape iterates. For the second and third example we also used a re-meshing technique. More precisely, as outlined in Chapter 4, the assembly of the nonlocal stiffness matrix is very expensive from a computational perspective. Therefore, since we have to assemble the nonlocal system not only for solving the state and adjoint equation in Line 4 of

Algorithm 2, but also potentially several times for performing the backtracking line search in Line 16, Algorithm 2 is clearly a costly endeavor for the nonlocally constrained shape optimization problem at hand. In order to ease the computational effort we therefore first compute a fixed number of iterations on a coarse grid, then re-fine the mesh and use the interpolated iterate from the coarse grid onto the fine grid as a "warm-start" for the computations on the fine grid. We stop the final iteration when a sufficient decrease of the norm of the shape gradient is achieved.

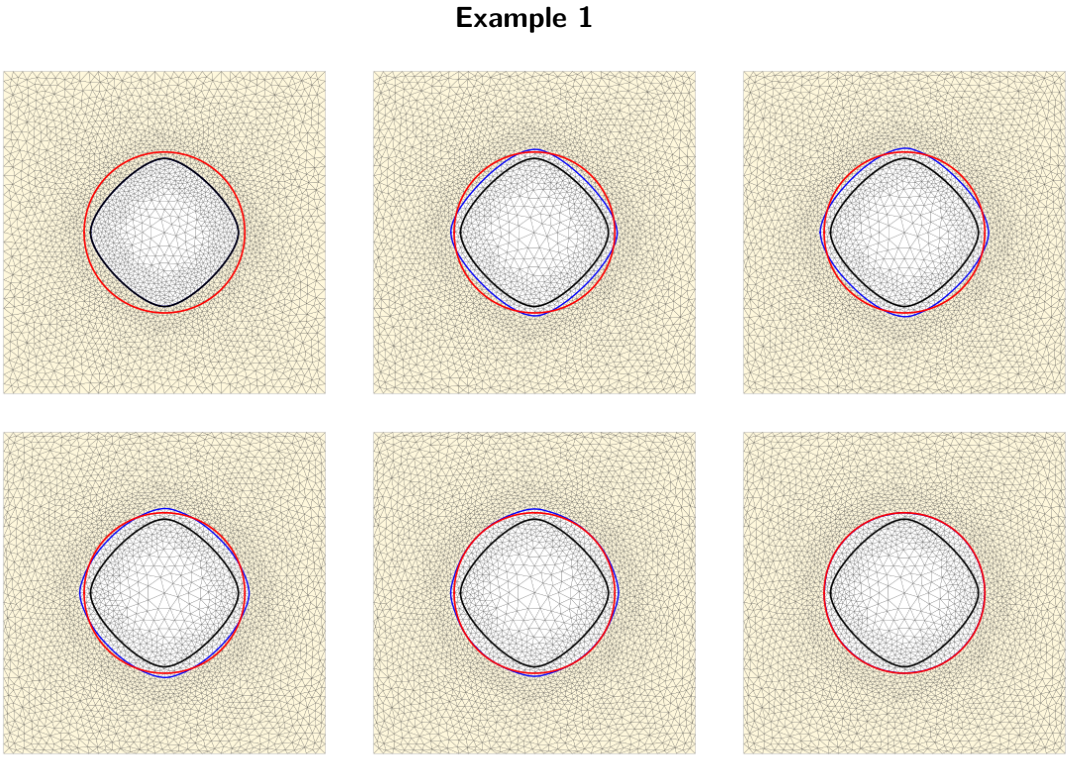


Figure 8.5: We have chosen a mild perturbation parameter $c_{per} = 0.0001$. The optimization algorithm terminated after 35 iterations. In the top row from left to right the reader finds the initial interface (black) as well as the iterates 1 and 2 (blue). Followed by the iterates 3 and 4 and the final shape in the bottom row. The finite element mesh for $\Omega \cup \Omega_I$ (note that Ω_I is not depicted in the image) consists of 7008 triangles with maximum diameter $h = 0.05$ and 3032 interior nodes in Ω, so that the stiffness matrix is an element of $\mathbb{R}^{3032\times 3032}$. As desired, we find that the blue iterates finally converge from the black initial shape to the red target shape.

We have implemented Algorithm 2 into a fully self-contained Python program including the assembly of the nonlocal stiffness matrix. The finite element meshes are generated with the free software Gmsh. As already alluded to, the overall optimization

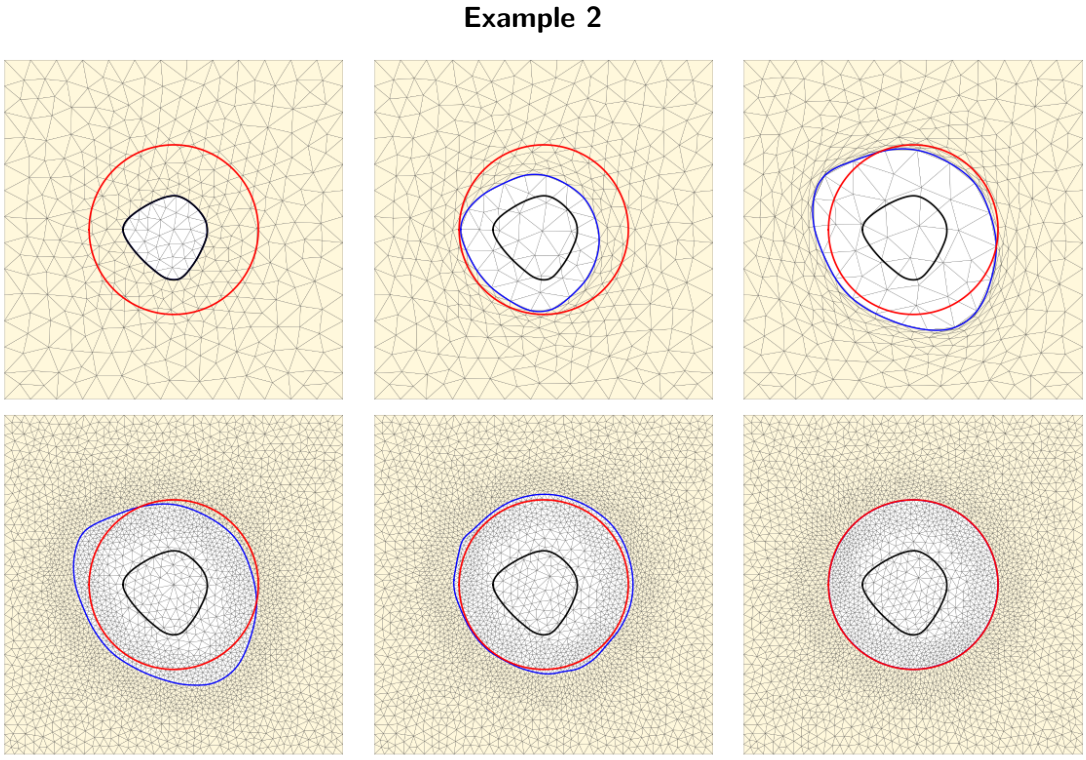


Figure 8.6: We have chosen $c_{per} = 0.001$. We run 25 iterations on the coarse grid and show in the upper row from left to right the initial configuration (black) as well as iterate 10 and the final iterate 25 (blue). The coarse grid consists of 870 triangles for $\Omega \cup \Omega_I$ (note that Ω_I is not depicted in the image) with a maximum diameter $h = 0.14$ and 408 interior nodes in Ω. We then perform the optimization on the fine grid, which ended after 32 iterations. In the bottom line the reader finds the initial interpolated interface (blue), iterate 4 and the solution after iteration 32. The fine grid consists of 6826 triangles with maximum diameter $h = 0.04$ and 3186 interior nodes, so that the stiffness matrix is in $\mathbb{R}^{3186\times 3186}$. We find that the blue iterates finally converge to the red target shape.

program is computationally very expensive due to the fact that we have to assemble the nonlocal stiffness matrix several times; see Line 4 and Line 16. Therefore, in order to keep computation times acceptable we assemble the nonlocal stiffness matrix by using approximate interaction sets. More precisely, we choose the approximation Example 2 from the numerical experiment described in Subsection 6.3.1, i.e., we integrate over whole triangles whose barycenter lie within the interaction set $S(\mathbf{x})$, here the $\|\cdot\|_\infty$-ball. The overall optimization program still runs a couple of hours. It is important to mention that computation times and the performance of Algorithm 2 in general are very sensi-

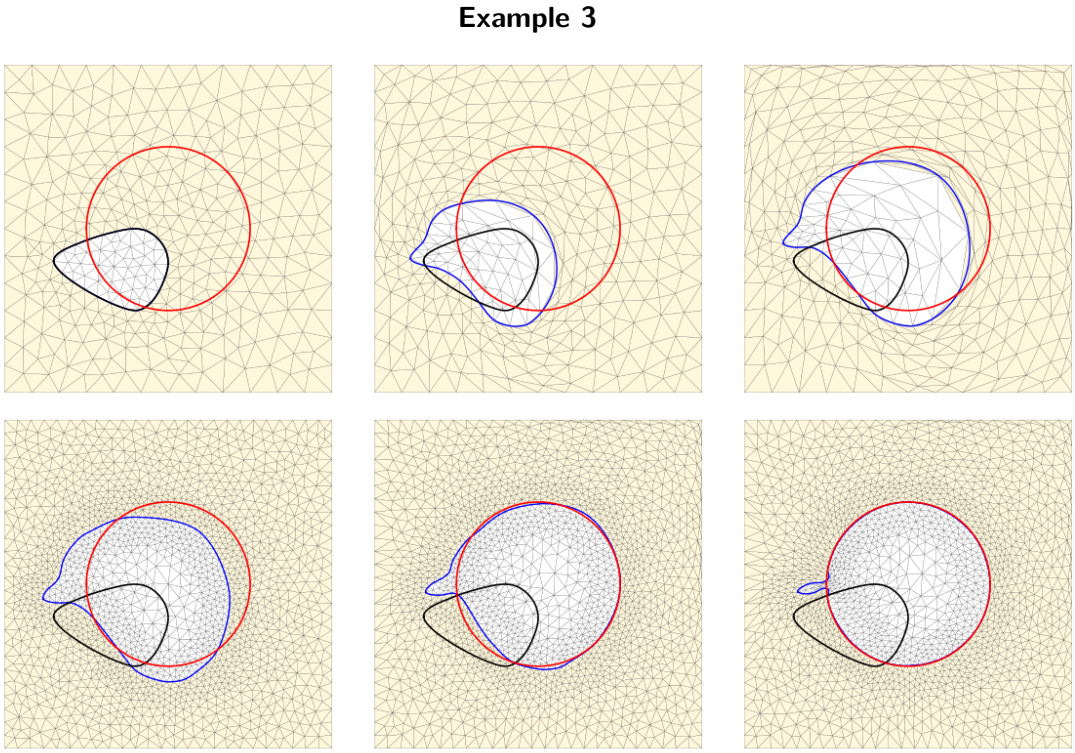


Figure 8.7: For this more complex experiment we had to choose a larger regularization parameter $c_{per} = 0.1$. We perform 50 iterations on the coarse grid and show in the upper row from left to right the initial configuration (black) as well as iterate 18 and the final iterate 50 (blue). The coarse grid consists of 768 triangles for $\Omega \cup \Omega_I$ with a maximum diameter $h = 0.21$ and 359 interior nodes in Ω. We then perform the optimization on the fine grid, which ended after 55 iterations. In the bottom line the reader finds the initial interpolated interface (blue), iterate 10 and the solution after iteration 55. The fine grid consists of 3924 triangles with maximum diameter $h = 0.08$ and 1789 interior nodes, so that the stiffness matrix is an element of $\mathbb{R}^{1789\times 1789}$. In this example, we do not exactly fit the target shape. In fact, on the lower left-hand side of the last shape iterate, a thin "bump" is formed. Different choices for the perimeter regularization ν and the initial Lamé parameter μ_{max} have led to similar corrupted simulations. We conjecture that these observations are due to pure discretization artefacts; see the related remarks in the concluding Section 8.5.

tive to the choice of parameters and may strongly vary, which is why reporting exact computation times is not very meaningful at this stage. Particularly delicate choices are those of the system parameters including the kernel (diffusion and convection) and the forcing term, which both determine the identifiability of the model. But also the choice of Lamé parameters to control the step size, specifically μ_{max} (we set $\mu_{min} = 0$ in all experiments, since we want the boundary of Ω to be fixed).

In order to make the algorithm more robust against an unfavorable choice of Lamé parameters, we have additionally implemented a dynamical adaption of μ_{max}. In the first phase of the optimization we typically have a larger distance between the state variable $u(\Gamma_k)$ and the target data $\bar{u}$. This leads to larger gradients and thus to mesh deformations of larger magnitude. An ideal choice of Lamé parameters would lead to an adequate stiffness of the mesh, which does not necessitate the backtracking line search in Line 16 and thereby would save costly assemblies of the nonlocal stiffness matrix. On the other hand, in a later phase of the algorithm, when the state variables $u(\Gamma_k)$ are closer to the data $\bar{u}$, mesh deformations decrease in magnitude and a stiff mesh resulting from a to largely chosen μ_{max} would lead to a stagnation of the algorithm. With these deliberations in mind, we perform a heuristic adaption of μ_{max} subsequent to the backtracking line search in the following way. We count the number of rounds of the while loop in Line 16, i.e., how often the step-size is downscaled through $\alpha = \tau\alpha$. A large number of rounds is an indication for potentially too large mesh deformations and we upscale μ_{max}; and vice-versa. Furthermore, we expand this procedure by a second measure to avoid unnecessary line search steps and thereby assemblies of the nonlocal stiffness matrix. As already pointed out, mesh deformations may be large in the early phase of the algorithm. Especially in the case of system parameters with high interface-sensitivity in combination with an inconveniently small μ_{max} which reinforces this behavior. Thus, such mesh deformations $\tilde{\mathbf{U}}_k$ of high magnitude lead to destroyed meshes if only mildly downscaled and an evaluation of the reduced objective functional $J^{red}(\Gamma_k - \alpha\tilde{\mathbf{U}}_k)$, which requires the assembly of the nonlocal stiffness matrix, becomes a pointless computation. In order to avoid such computations we first perform a line search depending on two simple mesh quality criteria. More precisely, we downscale the step size, i.e., $\alpha = \tau\alpha$, until the resulting interface $\Gamma_k - \alpha\tilde{\mathbf{U}}_k$ does not intersect itself (`self_intersect = False`) *and* all finite element nodes of the resulting mesh $\Omega_k - \alpha\tilde{\mathbf{U}}_k$ are a subset of Ω (`out_of_omega = False`). Here `self_intersect` and `out_of_omega` denote the boolean output of the two routines, which test for these mesh quality criteria. All in all, the backtracking line search in Line 16 of Algorithm 2 is substituted by the modified line search outlined in Algorithm 3. Throughout our experiments we have chosen the following set of parameters for Algorithm 3:

$$\mu_{max} = 20,\ n_{up} = 1,\ n_{down} = 4,\ c_{up} = 1.2 \text{ and } c_{down} = 0.8.$$

Also observe that we exploit the line search counter i in Algorithm 3 as a heuristic measure for the necessity of a potential restart of the overall optimization. In fact, it may happen from time to time that the L-BFGS updates "get stuck" and fail in determining a descent direction. This may result in a multiple downscaling to satisfy the line search criterion. Thus, if this is the case, we delete the L-BFGS memory (we store at most $m = 15$ vectors in our experiments) and restart the overall optimization with the current iterate as new initial guess. Although often motivated by the aim of saving memory storage, restart procedures are a common option in optimization software [67, 25, 80]. In addition to that, we can use the number of restarts as additional break criterion. In our computations we initiate a restart if the step size had to be downscaled

more than seven times, i.e., $n_{restart} = 8$ in Algorithm 3 and allow for at most 3 such restarts before we terminate the optimization; even if the gradients have not reached a sufficient decrease.

Algorithm 3: Modified line search with mesh quality check, dynamical adaption of μ_{max} and restart criterion

Additional parameters: $n_{up}, n_{down}, n_{restart} \in \mathbb{N},\ c_{up} > 1, c_{down} < 1$
$i = 0$
Linesearch depending on mesh quality
    while (`self_intersect or out_of_omega`) **do**
        $\alpha = \tau\alpha$
        $i = i + 1$
    end while
Linesearch depending on objective functional
    while $J^{red}(\Gamma_k - \alpha\tilde{\mathbf{U}}_k) \geq cJ^{red}(\Gamma_k)$ **do**
        $\alpha = \tau\alpha$
        $i = i + 1$
    end while
Adaption of μ_{max}
    if $i \geq n_{up}$ **then**
        $\mu_{max} = c_{up}\mu_{max}$
    end if
    if $i \leq n_{down}$ **then**
        $\mu_{max} = c_{down}\mu_{max}$
    end if
Restart criterion
    if $i \geq n_{restart}$ **then**
        $\rightarrow$ restart overall optimization

8.5 Concluding remarks and future work

We have conducted a study for shape optimization problems which are constrained by nonlocal system models. We have proven through numerical experiments the applicability of established shape optimization techniques for which the shape derivative of the nonlocal bilinear form represents the crucial ingredient. During this feasibility study we have uncovered a couple of interesting challenges, which are purely attributable to the involvement of nonlocality, such as the consideration of truncated kernels or the lack of regularity of nonlocal weak solutions. All in all, this chapter is only a first step along the exploration of the interesting field of nonlocally constrained shape optimization problems and many open questions surfaced during our studies. Therefore, we end this final chapter by listing a selection of open problems, which we consider being important for

further investigations and which we aim to address in future work.

Rigorous analysis of the saddle point system. It remains to establish a rigorous justification for the consideration of the reduced objective functional as is done in Subsection 8.2.3. As already alluded to, a common approach is to apply the theorem of Correa and Seeger and a proof for its applicability in our setting needs to be delivered subsequently.

Rigorous analysis of the (unperturbed) nonlocal interface problem. We conjecture that the unperturbed nonlocal interface problem is well-posed and we aim to present a rigorous proof for the existence of a (weak) solution. Furthermore, the pure nonlocal model with such an interface-dependent mixed kernel can be interpreted as the coupling of two nonlocal models and it is important to gain a deeper understanding of the underlying physics. In this regard the consideration of different interaction sets for the different atomic kernels is also of interest since it would contribute to an increase of generality.

Improving shape related implementation parts. The third numerical example presented in Figure 8.7 leaves us with the question whether the remaining unmatching "bump-like" part of the final iterate is a pure discretization artifact or ascribable to the nonlocal model. Under the assumption that the unperturbed nonlocal model admits a unique solution we have been able to show in Subsection 8.3.3 that this solution has a jump discontinuity across the interface. Thus, by only mildly perturbing the nonlocal operator with the Laplacian we still observe a steep gradient there. Thus the discretization with continuous basis functions may lead to a slightly oscillating behavior of the approximate finite element solution along the interface, which is not fully smoothened if c_{per} is chosen too small with respect to the kernel choice. Since the crucial contributions to the shape derivative are determined at the interface, these oscillations may negatively affect the performance of the optimization and may lead to degenerated solutions in more complicated cases such as the one depicted in this example. In addition to that, the problem may also be traced back to an inconvenient distribution chosen for the Lamé parameter μ and thereby leading to an unfavorable stiffness of the mesh. All in all, we are confident to prevent such misbehaviors by implementing discontinuous finite element basis functions and by investigating better choices for the distribution of the Lamé parameter μ; for the latter see also [70, Section 5].

Local limit. It is a well known fact that the nonlocal objects, such as the nonlocal operator, bilinear form and especially the corresponding weak solution, converge to appropriately chosen local counterparts as the interaction horizon vanishes $\delta \to 0$; see Chapter 7. Since the foregoing shape formalism is thoroughly investigated for constraint equations which are classical partial differential equations, it is interesting to investigate the local limit for the unperturbed nonlocal interface problem and the associated shape derivative. This is not only interesting in its own right but would also serve as a further diagnostic tool for the reliability of the derived (nonlocal) shape derivative.

Shape derivative for general mixed kernels. As we have proceeded to compute the shape derivative of the nonlocal bilinear we have restricted ourselves to a certain class of mixed kernels due to partially weighted interaction sets. For a wider application

spectrum it is of great importance to provide a method which enables us to numerically solve shape optimization problems as in the present case which involve general mixed kernels. This might require the application of shape optimization techniques other then those used in this chapter; see the next point.

Investigation of different shape optimization techniques. It might be worth considering other numerical techniques to approach the shape optimization problem at hand. One example would be the phase field approach; see [14, 13] and the references therein. Also, considering sharp interfaces and local gradients in a nonlocal setting is conceptually not coherent with the nonlocal "philosophy" and a different conception of nonlocal interface problems may lead to different optimization tools.

Nonlocal-to-local coupling. In mechanical applications a lot of effort has been put in the study of nonlocal-to-local coupling. The aim is to benefit from the computational ease that (local) partial differential models come along with, but retain accuracy in regions, where singular behaviors, such as cracks, are expected; see, e.g., [43] and the references therein. However, this requires the a priori knowledge about where to optimally apply the nonlocal model. Thus, it is of great importance to develop techniques which enable the identification of these regions in a rather automatic manner. One first step would be to formulate a shape optimization problem to identify the interface between a local and a nonlocal model.

Choice of inner product. In order to determine the mesh deformation vector field, i.e., the gradient, we solve equation (8.16), i.e., the linear elasticity equation. An optimal choice for the inner product would be the one determined by the Hessian of the reduced objective functional. Since we are considering nonlocally constrained optimization problems, it stands to reason to consider an inner product that is determined by a nonlocal operator. This would add to the computational costs but could lead to a faster and potentially grid-independent convergence behavior.

Chapter 9

Conclusion and outlook

In this thesis we have treated the nonlocal Dirichlet problem with volume constraints given in (2.1) from a modeling and analytical perspective, we have extensively discussed its discretization via the finite element method and finally formulated and numerically solved a shape optimization problem to identify an interface-dependent kernel. We have thereby set a focus on truncated kernels of the form

$$\gamma(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x}, \mathbf{y})\chi_{S(\mathbf{x})}(\mathbf{y}) \tag{9.1}$$

for interaction sets $S(\mathbf{x})$ which can be sandwiched between two Euclidean balls and for which the indicator function $\chi_{S(\mathbf{x})}(\mathbf{y})$ is symmetric. Although the case $S(\mathbf{x}) = \mathbb{R}^d$ is of interest, e.g., in the fractional calculus, we mainly worked with finitely truncated kernels. For the most part in the related literature only the case $S(\mathbf{x}) = B_{\delta,2}(\mathbf{x})$ is treated. Thus the consideration of more general interaction sets naturally gives rise to the question, whether established analytical results are transferable to kernels of type (9.1) involving such interaction sets. In this regard me we have made the following contributions to the **analysis and modeling** aspect:

- *Well-posedness.* We have proved the well-posedness for a variational formulation of problem (2.1) for two prominent kernel classes, namely integrable kernels (Subsection 3.3.1) and a certain type of singular kernels which are closely related to fractional diffusion (Subsection 3.3.2).

- *Comparison framework.* We have considered two nonlocal models, which solely differ in the choice of the kernel. We have derived estimates for the difference in bilinear forms (Corollary 6.1.1) and the difference in weak solutions (Proposition 6.1.4). A focus has been set on the comparison of kernels which solely differ in their interaction sets. In particular we have derived precise integrals to effect the comparison of Euclidean balls and $\|\cdot\|_\infty$-balls for various radii choices (Lemma 6.1.5).

While specializing to norm induced interaction balls, so that we can consider the size of the interaction sets as varying parameter, we have documented the following asymptotic results:

- *Infinite horizon.* We have shown that solutions corresponding to nontruncated kernels converge to solutions corresponding to related truncated kernels as the interaction horizon increases. For illustration purposes we have considered the fractional Poisson problem (Lemma 6.2.3). In order to make this statement relevant for numerical practice we have computed the related integrals (Lemma 6.2.2).

- *Vanishing horizon.* We have shown that nonlocal operators and weak solutions converge to their local counterparts as the interaction horizon vanishes (Proposition 7.2.1 and Theorem 7.3.4). We thereby allowed for kernels which are truncated by any norm induced interaction sets and which are characterized by kernel functions ϕ that are radial with respect to any norm. The crucial estimate to translate between the general and the standard case has been derived in Lemma 7.3.3. For numerical purposes we have derived precise integral respresentation for associated scaling constants (Lemma 7.4.1).

The latter results have been achieved by translating the general kernel into a standard kernel by deriving appropriate estimates and computing related integrals.

These analytical results then have served as a necessary foundation for subsequent numerical investigations of the computational benefits that nonstandard interaction sets come along with, in particular polyhedral and approximate interaction sets. We have served the **discretization** aspect in the following way.

- *Detailed description.* We have explained in detail the discretization of problem (2.1) via the finite element method including related challenges and the benefits of polyhedral interaction sets. We have documented useful representations for the entries of the nonlocal stiffness matrix based on the support of the basis functions (see (4.13) and (4.18)) and we have derived a fully discretized version thereof (4.28).

- *Approximate interaction sets.* We have introduced the notion of approximate interaction sets in Section 4.4 with a further investigation in Section 6.3 concerning their effect on the finite element error (Corollary 6.3.2).

- *Structure exploiting method.* Under appropriate assumptions on the kernel and the domain, we have introduced a finite element discretization, which yields multilevel Toeplitz stiffness matrices (Theorem 5.1.2). We have explained the computational advantages, which have enabled us to solve three-dimensional problems (Subsection 5.3.1).

In the final Chapter 8 we have formulated a shape **optimization** problem which is constrained by a nonlocal equation of type (2.1).

- *Problem formulation.* We have introduced a nonlocal convection-diffusion model, which is determined by an interface-dependent kernel (see (8.4)). In order to identify the interface from given measurements we formulate an optimization problem which is constrained by the nonlocal model at hand and where the control, i.e., the interface, is modeled as a shape.

- *Shape derivative.* We have derived the shape derivate of the nonlocal bilinear form (see (8.23)), which was necessary to deduce an expression for the shape derivative of the associated reduced objective functional (see (8.15)).

- *Numerical realization.* We have implemented a shape optimization algorithm and presented three numerical study cases.

All in all, nonlocal models with a finite range of interactions enrich our modeling spectrum and present useful alternative as well as supplemental approaches to classical differential equations. However, their numerical discretization for application-oriented problems is challenging. In this regard, it should be our ultimate goal to develop efficient discretization techniques, which enable us to solve nonlocal equations on complex three-dimensional domains. Therefore, we see great potential in approximate interaction sets which only necessitate the integration over whole elements. In future work, we therefore aim at extending the analysis of their effect on the quadrature error and thus on the finite element error. It is desirable to find an approximation approach, which does not result in a significant deterioration of the finite element error.

Clearly, solving the constraint equation is a necessary prerequisite for optimization problems which, e.g., enable the identification of related parameters in the model of interest. Thus, a computationally tractable assembly routine to generate the discretized system is of great importance. Besides this, the particular shape optimization problem discussed in Chapter 8 has led itself to many challenging and interesting problems as outlined in the concluding Section 8.5. In the near future, besides the necessary development of a rigorous analytical foundation for the introduced problem at hand, we also want to put an emphasized focus on applying the gathered insights to nonlocal-to-local coupling problems as they provide a promising alternative approach for easing the discretization of nonlocal models.

List of Figures

Bibliography

[1] G. Acosta, F.M. Bersetche, and J.P. Borthagaray. A short FE implementation for a 2D homogeneous Dirichlet problem of a fractional Laplacian. *Comput. Math. Appl.*, 74(4):784–816, 2017.

[2] G. Acosta and J.P. Borthagaray. A fractional Laplace equation: Regularity of solutions and finite element approximations. *SIAM Journal on Numerical Analysis*, 55(2):472–495, 2017.

[3] R. Adams. *Sobolev Spaces*. Academic Press, New York-London, 1975.

[4] M. Ainsworth and C. Glusa. Aspects of an adaptive finite element method for the fractional Laplacian: A priori and a posteriori error estimates, efficient implementation and multigrid solver. *Computer Methods in Applied Mechanics and Engineering*, 327:4 – 35, 2017.

[5] M. Ainsworth and C. Glusa. Towards an efficient finite element method for the integral fractional Laplacian on polygonal domains. *Contemporary Computational Mathematics - A Celebration of the 80th Birthday of Ian Sloan*, pages 17–57, 2018.

[6] B. Aksoylu and T. Mengesha. Results on nonlocal boundary value problems. *Numer. Funct. Anal. Optim.*, 31:1301–1317, 2010.

[7] B. Aksoylu and M.L. Parks. Variational theory and domain decomposition for nonlocal problems. *Appl. Math. Comput.*, 217:6498–6515, 2011.

[8] B. Alali, K. Liu, and M. Gunzburger. A generalized nonlocal vector calculus. *Zeitschrift für angewandte Mathematik und Physik*, 66(5):2807–2828, 2015.

[9] H.B. Ameur, M. Burger, and B. Hackl. Level set methods for geometric inverse problems in linear elasticity. *Inverse Problems*, 20(3):673–696, 2004.

[10] F. Andreu, J.M. Mazon, J.D. Rossi, and J. Toledo. *Nonlocal Diffusion Problems*, volume 165. Math. Surveys Monogr. AMS, Providence, RI, 2010.

[11] M.T. Barlow, R.F. Bass, Z.-Q. Chen, and M. Krassmann. Non-local Dirichlet forms and symmetric jump processes. *Transactions of the American Mathematical Society*, 361(4):1963–1999, 2009.

[12] R.F. Bass, M. Kassmann, and T. Kumagai. Symmetric jump processes: Localization, heat kernels and convergence. *Ann. Inst. H. Poincaré Probab. Statist.*, 46(1):59–71, 2010.

[13] L. Blank, H. Garcke, M. Hassan Farshbaf-Shaker, and V. Styles. Relating phase field and sharp interface approaches to structural topology optimization. *ESAIM: Control, Optimisation and Calculus of Variations*, 20(4):1025–1058, 2014.

[14] L. Blank, H. Garcke, C. Hecht, and C. Rupprecht. Sharp interface limit for a phase field model in structural optimization. *SIAM Journal on Control and Optimization*, 54(3):1558–1584, 2016.

[15] J. Fernández Bonder, A. Ritorto, and A. Salort. A class of shape optimization problems for some nonlocal operators. *Advances in Calculus of Variations*, 11(4):373–386, 2017.

[16] J. Fernández Bonder and J.F. Spedaletti. Some nonlocal optimal design problems. *Journal of Mathematical Analysis and Applications*, 459(2):906–931, 2018.

[17] D. Brockmann. Anomalous diffusion and the structure of human transportation networks. *The European Physical Journal Special Topics*, 157(1):173–189, 2008.

[18] D. Brockmann and F. Theis. Money circulation, trackable items, and the emergence of universal human mobility patterns. *IEEE Pervasive Computing*, 7(4):28–35, 2008.

[19] A. Burchard, R. Choksi, and I. Topaloglu. Nonlocal shape optimization via interactions of attractive and repulsive potentials. 2018. arXiv:1512.07282.

[20] O. Burkovska and M. Gunzburger. Regularity and approximation analyses of nonlocal variational equality and inequality problems. 2018. arXiv:1804.10282.

[21] T.F. Chan. An optimal circulant preconditioner for Toeplitz systems. *SIAM journal on scientific and statistical computing*, 9(4):766–771, 1988.

[22] E. Chasseigne, M. Chaves, and J.D. Rossi. Asymptotic behavior for nonlocal diffusion equations. *Journal de Mathématiques Pures et Appliquées*, 86(3):271–291, 2006.

[23] J. Chen, T.L.H. Li, and M. Anitescu. A parallel linear solver for multilevel Toeplitz systems with possibly several right-hand sides. *Parallel Computing*, 40(8):408–424, 2014.

[24] M. Chen and W. Deng. Convergence analysis of a multigrid method for a nonlocal model. *SIAM Journal on Matrix Analysis and Applications*, 38(3):869–890, 2017.

[25] C. Chih-Chao. Periodically restarted quasi-Newton updates in constant arc-length method. *Computers and Structures*, 41:963–972, 1991.

[26] R. Cools and A. Haegemans. An embedded pair of cubature formulae of degree 5 and 7 for the triangle. *BIT Numerical Mathematics*, 28(2):357–359, 1988.

[27] C. Cortazar, M. Elgueta, and J.D. Rossi. Nonlocal diffusion problems that approximate the heat equation with Dirichlet boundary conditions. *Israel Journal of Mathematics*, 170(1):53–60, 2009.

[28] N. Cusimano, A. Bueno-Orovio, I. Turner, and K. Burrage. On the order of the fractional Laplacian in determining the spatio-temporal evolution of a space-fractional model of cardiac electrophysiology. *PLOS ONE*, 10(12):1–16, 12 2015.

[29] A.-L. Dalibard and D. Gérard-Varet. On shape optimization problems involving the fractional Laplacian. *ESAIM: Control, Optimisation and Calculus of Variations*, 19(4):976–1013, 2013.

[30] J.P. de los Reyes. *Numerical PDE-Constrained Optimization*. SpringerBriefs in Optimization. Springer Cham Heidelberg New York Dordrecht London, 2015.

[31] M.C. Delfour and J.-P. Zolésio. *Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization*. Advances in Design and Control. SIAM Philadelphia, 2011.

[32] M. D'Elia, C. Glusa, and E. Otárola. A priori error estimates for the optimal control of the integral fractional Laplacian. 2018. arXiv:1810.04262.

[33] M. D'Elia and M. Gunzburger. The fractional Laplacian operator on bounded domains as a special case of the nonlocal diffusion operator. *Computers and Mathematics with Applications*, 66(7):1245–1260, 2013.

[34] M. D'Elia and M. Gunzburger. Optimal distributed control of nonlocal steady diffusion problems. *SIAM Journal on Control and Optimization*, 52(1):243–273, 2014.

[35] M. D'Elia and M. Gunzburger. Identification of the diffusion parameter in nonlocal steady diffusion problems. *Appl. Math. Optim.*, 73(2):227–249, 2016.

[36] Q. Du, M. Gunzburger, R.B. Lehoucq, and K. Zhou. Analysis and Approximation of Nonlocal Diffusion Problems with Volume Constraints. *SIAM Review*, 54(4):667–696, 2012.

[37] Q. Du, M. Gunzburger, R.B. Lehoucq, and K. Zhou. A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws. *Mathematical Models and Methods in Applied Sciences*, 23(03):493–540, 2013.

[38] Q. Du, L. Ju, L. Tian, and K. Zhou. A posteriori error analysis of finite element method for linear nonlocal diffusion and peridynamic models. *Math. Comp.*, 82:1889 – 1922, 2013.

[39] Q. Du and K. Zhou. Mathematical analysis for the peridynamic nonlocal continuum theory. *ESAIM: M2AN*, 45(2):217–234, 2011.

[40] M. Duffy. Quadrature over a pyramid or cube of integrands with a singularity at a vertex. *SIAM Journal on Numerical Analysis*, 19(6):1260–1262, 1982.

[41] D.A. Dunavant. High degree efficient symmetrical Gaussian quadrature rules for the triangle. *International Journal for Numerical Methods in Engineering*, 21(6):1129–1148, 1985.

[42] M. D'Elia, Q. Du, M. Gunzburger, and R.B. Lehoucq. Nonlocal convection-diffusion problems on bounded domains and finite-range jump processes. *Computational Methods in Applied Mathematics*, 17(4):707–722, 2017.

[43] M. D'Elia, M. Perego, P. Bochev, and D. Littlewood. A coupling strategy for nonlocal and local diffusion models with mixed volume constraints and boundary conditions. *Computers and Mathematics with Applications*, 71(11):2218 – 2230, 2016. Proceedings of the conference on Advances in Scientific Computing and Applied Mathematics. A special issue in honor of Max Gunzburger's 70th birthday.

[44] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*. Springer-Verlag New York, 2004.

[45] G. Da Fies and M. Vianello. Algebraic cubature on planar lenses and bubbles. *Dolomites Research Notes on Approximation*, 5(1):7–12, 2012.

[46] G. Da Fies and M. Vianello. Trigonometric Gaussian quadrature on subintervals of the period. *Electron. Trans. Numer. Anal*, 39:102 – 112, 2012.

[47] G. Gilboa and S. Osher. Nonlocal operators with applications to image processing. *Multiscale Modeling & Simulation*, 7(3):1005–1028, 2009.

[48] M. Gunzburger and R.B. Lehoucq. A Nonlocal Vector Calculus with Application to Nonlocal Boundary Value Problems. *Multiscale Modeling & Simulation*, 8(5):1581–1598, 2010.

[49] H. Elman and D. Silvester and A. Wathen. *Finite Elements and Fast Iterative Solvers*. Oxford University Press, New York, 2005.

[50] T.C. Hales. The Jordan curve theorem, formally and informally. *Am. Math. Monthly*, 114:882–894, 2007.

[51] M. Hintermüller and W. Ring. A second order shape optimization approach for image segmentation. *SIAM J. Appl. Math.*, 64(2):442–467, 2003.

[52] L.I. Ignat and J.D. Rossi. A nonlocal convection–diffusion equation. *Journal of Functional Analysis*, 251(2):399–437, 2007.

[53] K. Königsberger. *Analysis 2*. Springer-Verlag Berlin Heidelberg, 2004.

[54] M. Kwaśnicki. Ten equivalent definitions of the fractional Laplace operator. *Fractional Calculus and Applied Analysis*, 20(1):7–51, 2017.

[55] S.Z. Levendorskiĭ. Pricing of the American put under lévy processes. *International Journal of Theoretical and Applied Finance*, 07(03):303–335, 2004.

[56] Y. Lou, X. Zhang, S. Osher, and A. Bertozzi. Image recovery via nonlocal operators. *Journal of Scientific Computing*, 42(2):185–197, 2010.

[57] M.M. Meerschaert and C. Tadjeran. Finite difference approximations for fractional advection–dispersion flow equations. *Journal of Computational and Applied Mathematics*, 172(1):65 – 77, 2004.

[58] P.W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS)*, 8:1–48, 2006.

[59] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bulletin des Sciences Mathematiques*, 136(5):521 – 573, 2012.

[60] J. Nocedale and S.J. Wright. *Numerical Optimization*. Springer, New York, 1999.

[61] A. Novotny and J. Sokolowski. *Topological Derivatives in Shape Optimization*. Springer, 2013.

[62] M.A. Olshanskii and E.E. Tyrtyshnikov. *Iterative Methods for Linear Systems: Theory and Applications*. SIAM, 2014.

[63] H.-K. Pang and H.-W. Sun. Multigrid method for fractional diffusion equations. *Journal of Computational Physics*, 231(2):693 – 703, 2012.

[64] E. Parini and A. Salort. Compactness and dichotomy in nonlocal shape optimization. 2018. arXiv:1806.01165.

[65] M.L. Parks, R.B. Lehoucq, S.J. Plimpton, and S.A. Silling. Implementing peridynamics within a molecular dynamics code. *Computer Physics Communications*, 179(11):777 – 783, 2008.

[66] G. Peyré, S. Bougleux, and L. Cohen. *Non-local Regularization of Inverse Problems*, pages 57–68. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[67] M.J.D. Powell. Restart procedures for the conjugate gradient method. *Mathematical Programming*, 12(1):241–254, 1977.

[68] L. Rosasco, M. Belkin, and E. De Vito. On learning with integral operators. *J. Mach. Learn. Res.*, 11:905–934, 2010.

[69] S. Schmidt, C. Ilic, V. Schulz, and N. Gauger. Three dimensional large scale aerodynamic shape optimization based on the shape calculus. *AIAA Journal*, 51(11):2615–2627, 2013.

[70] V. Schulz and M. Siebenborn. Computational comparison of surface metrics for PDE constrained shape optimization. *Computational Methods in Applied Mathematics*, 16(3):485–496, 2016.

[71] V. Schulz, M. Siebenborn, and K. Welker. Structured inverse modeling in parabolic diffusion problems. *SIAM Journal on Control and Optimization*, 53(6):3319–3338, 2015.

[72] V. Schulz, M. Siebenborn, and K. Welker. Efficient PDE constrained shape optimization based on Steklov–Poincaré-type metrics. *SIAM Journal on Optimization*, 26(4):2800–2819, 2016.

[73] V. Schulz, M. Siebenborn, and K. Welker. A novel Steklov-Poincaré type metric for efficient PDE constrained optimization in shape spaces. *SIAM J. Optimization*, 26(4):2800–2819, 2016.

[74] P. Seleson and D.J. Littlewood. Convergence studies in meshfree peridynamic simulations. *Computers and Mathematics with Applications*, 71(11):2432 – 2448, 2016. Proceedings of the conference on Advances in Scientific Computing and Applied Mathematics. A special issue in honor of Max Gunzburger's 70th birthday.

[75] S. Serra-Capizzano and E.E. Tyrtyshnikov. Any circulant-like preconditioner for multilevel matrices is not superlinear. *SIAM J. Matrix Analysis Applications*, 21:431–439, 2000.

[76] M. Siebenborn. A shape optimization algorithm for interface identification allowing topological changes. *J. Optim. Theory Appl.*, 177(2):306–328, 2018.

[77] M. Siebenborn and A. Vogel. A shape optimization algorithm for cellular composites. 2019. arXiv:1904.03860.

[78] S.A. Silling. Reformulation of elasticity theory for discontinuities and long-range forces. *Journal of the Mechanics and Physics of Solids*, 48(1):175 – 209, 2000.

[79] S.A. Silling and E. Askari. A Meshfree Method Based on the Peridynamic Model of Solid Mechanics. *Comput. Struct.*, 83(17-18):1526–1535, 2005.

[80] V. Simoncini. On the convergence of restarted Krylov subspace methods. *SIAM Journal on Matrix Analysis and Applications*, 22(2):430–452, 2000.

[81] Y. Sire, J.L. Vázquez, and B. Volzone. Symmetrization for fractional elliptic and parabolic equations and an isoperimetric application. *Chinese Annals of Mathematics, Series B*, 38(2):661–686, 2017.

[82] J. Sokolowski and J.-P. Zolésio. *An introduction to shape optimization.* Springer, 1992.

[83] P. Tankov. *Financial modelling with jump processes.* CRC Press, 2003.

[84] X. Tian. *Nonlocal models with a finite range of nonlocal interactions*. PhD thesis, Columbia University, 2017. https://doi.org/10.7916/D8ZG6XWN.

[85] X. Tian and Q. Du. Asymptotically compatible schemes and applications to robust discretization of nonlocal models. *SIAM Journal on Numerical Analysis*, 52(4):1641–1665, 2014.

[86] N. Trask, H. You, Y. Yu, and M.L. Parks. An asymptotically compatible meshfree quadrature rule for nonlocal problems with applications to peridynamics. *Computer Methods in Applied Mechanics and Engineering*, 343:151 – 165, 2019.

[87] C. Vollmann and V. Schulz. Exploiting multilevel Toeplitz structures in high dimensional nonlocal diffusion. *Computing and Visualization in Science*, 2019.

[88] V.Olshevsky, I. Oseledets, and E. Tyrtyshnikov. Tensor properties of multilevel Toeplitz and related matrices. *Linear Algebra and its Applications*, 412(1):1–21, 2006.

[89] H. Wang and T.S. Basu. A fast finite difference method for two-dimensional space-fractional diffusion equations. *SIAM Journal on Scientific Computing*, 34(5):A2444–A2458, 2012.

[90] H. Wang and N. Du. A fast finite difference method for three-dimensional time-dependent space-fractional diffusion equations and its efficient implementation. *Journal of Computational Physics*, 253:50 – 63, 2013.

[91] M. Webb. Analysis and approximation of a fractional differential equation. Master's thesis, Oxford University, Oxford, 2012.

[92] K. Welker. *Efficient PDE Constrained Shape Optimization in Shape Spaces*. PhD thesis, Universität Trier, 2016. http://ubt.opus.hbz-nrw.de/volltexte/2017/1024/.

[93] K. Welker. Suitable spaces for shape optimization. 2017. arXiv:1702.07579.

[94] D. Werner. *Funktionalanalysis*. Springer-Verlag Berlin Heidelberg, 2000.

[95] F. Xu. *A Multiscale Implementation of Finite Element Methods for Nonlocal Models of Mechanics and Diffusion*. PhD thesis, The Florida State University, 2015. http://diginole.lib.fsu.edu/islandora/object/fsu%3A360485.

[96] F. Xu, M. Gunzburger, and J. Burkardt. A multiscale method for nonlocal mechanics and diffusion and for the approximation of discontinuous functions. *Computer Methods in Applied Mechanics and Engineering*, 307:117 – 143, 2016.

[97] K. Ye and L.H. Lim. Algorithms for structured matrix-vector product of optimal bilinear complexity. In *2016 IEEE Information Theory Workshop (ITW)*, pages 310–314, 2016.

[98] K. Zhou and Q. Du. Mathematical and Numerical Analysis of Linear Peridynamic Models with Nonlocal Boundary Conditions. *SIAM Journal on Numerical Analysis*, 48(5):1759–1780, 2010.

Appendix A

Implementation details for the Toeplitz solver

In this appendix we aim to provide further details on the implementation of the Toeplitz solver presented in Chapter 5.

A.1 Entries of the stiffness matrix

We analyze the entries

$$a_{kj} = a(\mathbf{v}_{kj}),$$

with $\mathbf{v}_{kj}$ from (5.6), of the stiffness matrix $\mathbf{A}^h$ more closely and derive a representation which can be efficiently implemented.

We first characterize the domain of integration occurring in the integral in $a(\mathbf{v}_{kj})$. Let us define

$$\hat{\mathcal{S}}_{kj} := (\mathbf{v}_{kj} + \hat{\mathcal{S}}),$$

then

$$\begin{aligned}
&\left(\hat{\mathcal{S}}^c \times \hat{\mathcal{S}}^c \cup \hat{\mathcal{S}}_{kj}^c \times \hat{\mathcal{S}}_{kj}^c\right)^c = \left(\hat{\mathcal{S}}^c \times \hat{\mathcal{S}}^c\right)^c \cap \left(\hat{\mathcal{S}}_{kj}^c \times \hat{\mathcal{S}}_{kj}^c\right)^c \\
=&\left((\hat{\mathcal{S}} \times \hat{\mathcal{S}}) \cup (\hat{\mathcal{S}}^c \times \hat{\mathcal{S}}) \cup (\hat{\mathcal{S}} \times \hat{\mathcal{S}}^c)\right) \cap \left((\hat{\mathcal{S}}_{kj} \times \hat{\mathcal{S}}_{kj}) \cup (\hat{\mathcal{S}}_{kj}^c \times \hat{\mathcal{S}}_{kj}) \cup (\hat{\mathcal{S}}_{kj} \times \hat{\mathcal{S}}_{kj}^c)\right) \\
=&(C \times C) \cup (D_k \times C) \cup (C \times D_k) \cup (D_j \times C) \cup (\hat{\mathcal{S}}^c \cap \hat{\mathcal{S}}_{kj}^c \times C) \cup (D_j \times D_k) \\
&\cup (C \times D_j) \cup (D_k \times \hat{\mathcal{S}}_{kj}) \cup (C \times \hat{\mathcal{S}}^c \cap \hat{\mathcal{S}}_{kj}^c),
\end{aligned}$$

where we set

$$C := \hat{\mathcal{S}} \cap \hat{\mathcal{S}}_{kj}, \quad D_k := \hat{\mathcal{S}} \cap \hat{\mathcal{S}}_{kj}^c \quad \text{and} \quad D_j := \hat{\mathcal{S}}^c \cap \hat{\mathcal{S}}_{kj}.$$

According to (5.7) we obtain, by exploiting the symmetry of the integrand,

$$2a_{kj}/h^{2d} = \int_{\hat{\mathcal{S}} \cap \hat{\mathcal{S}}_{kj}} \int_{\hat{\mathcal{S}} \cap \hat{\mathcal{S}}_{kj}} F(\mathbf{x}, \mathbf{y}; \mathbf{v}_{kj}) d\mathbf{y} d\mathbf{x}$$

$$
\begin{aligned}
&+ 2\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_{kj}}\int_{\hat{\mathcal{S}}_{kj}^c\cap\hat{\mathcal{S}}} F(\mathbf{x},\mathbf{y};\mathbf{v}_{kj})d\mathbf{y}d\mathbf{x}\\
&+ 2\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_{kj}}\int_{\hat{\mathcal{S}}_{kj}\cap\hat{\mathcal{S}}^c} F(\mathbf{x},\mathbf{y};\mathbf{v}_{kj})d\mathbf{y}d\mathbf{x}\\
&+ 2\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_{kj}}\int_{\hat{\mathcal{S}}^c\cap\hat{\mathcal{S}}_{kj}^c} F(\mathbf{x},\mathbf{y};\mathbf{v}_{kj})d\mathbf{y}d\mathbf{x}\\
&+ 2\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_{kj}^c}\int_{\hat{\mathcal{S}}^c\cap\hat{\mathcal{S}}_{kj}} F(\mathbf{x},\mathbf{y};\mathbf{v}_{kj})d\mathbf{y}d\mathbf{x},
\end{aligned}
$$

where $F(\mathbf{x},\mathbf{y};\mathbf{z}) = (\hat{\varphi}(\mathbf{y}-\mathbf{z}) - \hat{\varphi}(\mathbf{x}-\mathbf{z}))(\hat{\varphi}(\mathbf{y}) - \hat{\varphi}(\mathbf{x}))\gamma(h\mathbf{y}, h\mathbf{x})$. Note that this representation holds for a general setting without assuming (A1) and (A2) (see also (4.18)).

From now on we also assume that the interaction horizon δ is larger or equal the diameter of the domain Ω, such that $\Omega \subset B_{\delta,\infty}(\mathbf{x})$ for all $\mathbf{x} \in \Omega$, i.e., $\delta \geq \operatorname{diam}(\Omega) = ||\mathbf{b}-\mathbf{a}||_2$. Thus we do not have to deal with partially covered elements (see the discussion in Subsection 4.3.2). This coincides with the application to space-fractional diffusion problems where we aim to model $\delta \to \infty$ (see Section 6.2).

Furthermore, since we can construct the whole stiffness matrix $\mathbf{A}^h$ from its first row $\mathbf{R}^h$, it is convenient to introduce the following $\mathbf{L}^d$-dimensional vectors:

$$
\begin{aligned}
\mathtt{sing}_k &:= h^{2d}\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_k}\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_k} F(\mathbf{x},\mathbf{y};E^{\mathbf{L}}(k))d\mathbf{y}d\mathbf{x}\\
&\quad + 2h^{2d}\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_k}\int_{\hat{\mathcal{S}}_k^c\cap\hat{\mathcal{S}}} F(\mathbf{x},\mathbf{y};E^{\mathbf{L}}(k))d\mathbf{y}d\mathbf{x}\\
&\quad + 2h^{2d}\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_k}\int_{\hat{\mathcal{S}}_k\cap\hat{\mathcal{S}}^c} F(\mathbf{x},\mathbf{y};E^{\mathbf{L}}(k))d\mathbf{y}d\mathbf{x},\\
\mathtt{rad}_k &:= 2h^{2d}\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_k}\int_{(\hat{\mathcal{S}}\cup\hat{\mathcal{S}}_k)^c} F(\mathbf{x},\mathbf{y};E^{\mathbf{L}}(k))\chi_{B_{\delta,\infty}(h\mathbf{x})}(h\mathbf{y})d\mathbf{y}d\mathbf{x},\\
\mathtt{dis}_k &:= 2h^{2d}\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_k^c}\int_{\hat{\mathcal{S}}^c\cap\hat{\mathcal{S}}_k} F(\mathbf{x},\mathbf{y};E^{\mathbf{L}}(k))d\mathbf{y}d\mathbf{x},
\end{aligned}
$$

for $0 \leq k < \mathbf{L}^d$, where

$$\hat{\mathcal{S}}_k := \hat{\mathcal{S}}_{0k} = E^{\mathbf{L}}(k) + \hat{\mathcal{S}}.$$

Since $\mathbf{x}_{k0} = \frac{\mathbf{x}_k - \mathbf{x}_0}{h} = E^{\mathbf{L}}(k)$, we have

$$\mathbf{R}^h = \frac{\mathtt{sing} + \mathtt{rad} + \mathtt{dis}}{2}.$$

As we see in the subsequent program, each of these vectors requires a different numerical handling which justifies this separation. In these premises we point out, that on the one hand we may touch possible singularities of the kernel function along the integration in $\mathtt{sing}_k$. On the other hand, the computation of $\mathtt{rad}_k$ may require the integration over a "large" domain if $\delta \to \infty$ (e.g. fractional kernel). Both are numerically demanding tasks

and complicate the assembling process. In contrast to that, the computation of $\mathtt{dis}_k$ turns out to be numerically viable without requiring a special treatment. However, we fortunately find

$$\hat{\mathcal{S}} \cap \hat{\mathcal{S}}_k = \emptyset \quad \Rightarrow \quad \mathtt{sing}_k = 0 = \mathtt{rad}_k.$$

Hence, it is worth identifying those indices k and treat them differently in the assembling loop. Therefore, from $\hat{\mathcal{S}} = \dot{\bigcup}_{i=0}^{2d}(\square - \mathbf{v}_i)$ we deduce that

$$\hat{\mathcal{S}} \cap \hat{\mathcal{S}}_k \neq \emptyset \quad \Leftrightarrow \quad \exists\, 0 \leq i < 2^d : \; k = E^{\mathbf{L}}(\mathbf{v}_i).$$

As a consequence, we only have to compute $\mathtt{sing}_k$ and $\mathtt{rad}_k$ for

$$k \in \mathtt{idx}_0 := \left\{ E^{\mathbf{L}}(\mathbf{v}_i) \colon 0 \leq i < 2^d \right\}.$$

We can cluster these indices even more. For that reason let us define on $\mathtt{idx}_0$ the equivalence relation

$$k \sim j :\Leftrightarrow \exists \text{ permutation matrix } \mathbf{P} \in \mathbb{R}^{\mathbf{L}^d \times \mathbf{L}^d} : \; E^{-\mathbf{L}}(k) = \mathbf{P} E^{-\mathbf{L}}(j).$$

Then due to Remark 5.1.3 we have to compute the values $\mathtt{sing}_k$ and $\mathtt{rad}_k$ only for

$$k \in \mathtt{idx}_s^k := \{ [j]_\sim \colon j \in \mathtt{idx}_0 \}.$$

In order to make this more precise, we figure out that the quotient set can further be specified as

$$\mathtt{idx}_s^k = \left\{ [E^{\mathbf{L}}(\mathbf{z})]_\sim \colon \mathbf{z} \in S \right\},$$

where

$$S := \{(0,\dots,0), (1,0,\dots,0), (1,1,0,\dots,0), \dots, (1,1,\dots,1)\} \subset \left\{ \mathbf{v}_i \colon 0 \leq i < 2^d \right\}$$

with $|S| = d + 1 \leq 2^d$. With other words, we group those $\mathbf{v}_i$ which are permutations of one another. The associated indices $0 \leq i < 2^d$ are thus given by

$$\left\{ E^{2\mathbf{e}}(\mathbf{z}) \colon \mathbf{z} \in S \right\} =: \mathtt{idx}_s^i.$$

In addition to the preceding considerations, we also want to partition the integration domains $\hat{\mathcal{S}} \cap \hat{\mathcal{S}}_k$, $\hat{\mathcal{S}} \cap \hat{\mathcal{S}}_k^c$ and $\hat{\mathcal{S}}^c \cap \hat{\mathcal{S}}_k$, where $k \in \mathtt{idx}_s^k$, into cubes $(\square - \mathbf{v}_\nu)$, such that we can express the reference basis function $\hat{\varphi}$ with the help of the element basis functions $\hat{\psi}_i$. This is necessary in order to compute the integrals in a vectorized fashion and obtain an efficient implementation. Let us start with

$$\hat{\mathcal{S}} \cap \hat{\mathcal{S}}_k, \quad \text{where} \quad k = E^{\mathbf{L}}(\mathbf{v}_i) \quad \text{for} \quad 0 \leq i < 2^d$$

such that $\hat{\mathcal{S}}_k = \mathbf{v}_i + \hat{\mathcal{S}}$. Then we define the set

$$D_i := \left\{ 0 \leq \mu < 2^d \colon \square - \mathbf{v}_\mu \in \hat{\mathcal{S}} \cap \hat{\mathcal{S}}_k \right\} = \left\{ 0 \leq \mu < 2^d \colon \exists\, \kappa \colon \mathbf{v}_\mu + \mathbf{v}_i = \mathbf{v}_\kappa \right\}$$

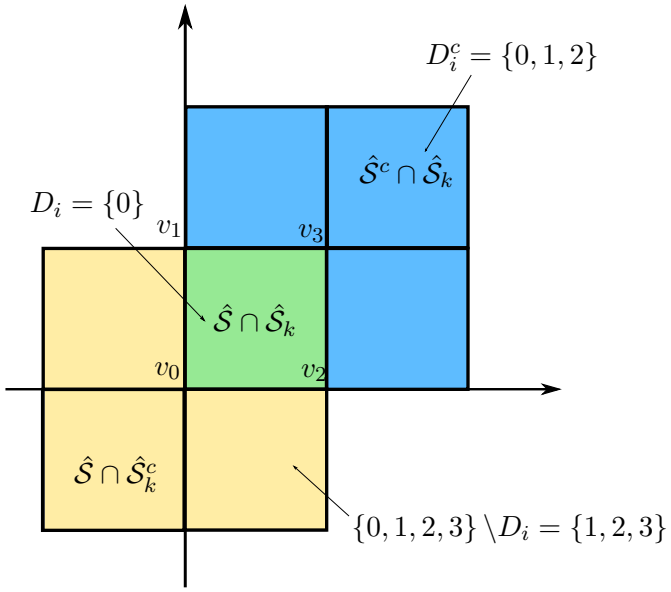


Figure A.1: Illustration of the index sets for $d = 2$, $i = 3$.

$$= \left\{ 0 \leq \mu < 2^d \colon (\mathbf{v}_\mu + \mathbf{v}_i)_j < 2 \quad \text{for all } 0 \leq j < d \right\}.$$

Since $\hat{\mathcal{S}} = \dot{\bigcup}_{i=0}^{2^d} (\square - \mathbf{v}_i)$, we find that

$$\hat{\mathcal{S}} \cap \hat{\mathcal{S}}_k = \bigcup_{\nu \in D_i} (\square - \mathbf{v}_\nu).$$

Hence, we readily recognize that

$$\hat{\mathcal{S}} \cap \hat{\mathcal{S}}_k^c = \bigcup_{\nu \in \{0,\ldots,2^d-1\} \setminus D_i} (\square - \mathbf{v}_\nu).$$

Similarly, we can derive a set D_i^c, such that

$$\hat{\mathcal{S}}^c \cap \hat{\mathcal{S}}_k = \mathbf{v}_i + \bigcup_{\nu \in D_i^c} (\square - \mathbf{v}_\nu).$$

Figure A.1 illustrates the latter considerations and in Table A.1 these sets are listed for dimensions $d \in \{1, 2, 3\}$, respectively.

With this at hand, we can now have a closer look at the vectors `sing`, `rad` and `dis` and put them into a form which is suitable for the implementation.

A.1.1 Vector "sing"

Let $i \in \mathtt{idx}_s^i$, such that $E^{\mathbf{L}}(k) = \mathbf{v}_i$. Since $\hat{\mathcal{S}} \cap \hat{\mathcal{S}}_k = \bigcup_{\nu \in D_i} \square - \mathbf{v}_\nu$, we can transform the first integral in `sing` as follows

$$\int_{\hat{\mathcal{S}} \cap \hat{\mathcal{S}}_k} \int_{\hat{\mathcal{S}} \cap \hat{\mathcal{S}}_k} F(\mathbf{x}, \mathbf{y}; E^{\mathbf{L}}(k)) d\mathbf{y} d\mathbf{x} = \sum_{\nu \in D_i} \sum_{\mu \in D_i} \int_{\square - \mathbf{v}_\nu} \int_{\square - \mathbf{v}_\mu} F(\mathbf{x}, \mathbf{y}; \mathbf{v}_i) d\mathbf{y} d\mathbf{x}$$

d	$i \in \mathtt{idx}_s^i$	$k = E^L(\mathbf{v}_i) \in \mathtt{idx}_s^k$	D_i	$D_i^c = \{0 \le j < i\}$	$\kappa(D_i, i)$
1	0	0	$(0,1)$	$\emptyset$	$(0,1)$
	1	1	0	0	1
2	0	0	$(0,1,2,3)$	$\emptyset$	$(0,1,2,3)$
	2	p_0	$(0,1)$	$(0,1)$	$(2,3)$
	3	$p_0 + p_1$	0	$(0,1,2)$	3
3	0	0	$(0,1,2,3,4,5,6,7)$	$\emptyset$	$(0,1,2,3,4,5,6,7)$
	4	p_0	$(0,1,2,3)$	$(0,1,2,3)$	$(4,5,6,7)$
	6	$p_0 + p_1$	$(0,1)$	$(0,1,2,3,4,5)$	$(6,7)$
	7	$p_0 + p_1 + p_2$	0	$(0,1,2,3,4,5,6)$	7

Table A.1: Index sets for the implementation.

$$= \sum_{\nu \in D_i} \sum_{\mu \in D_i} \int_\square \int_\square F(\mathbf{x} - \mathbf{v}_\nu, \mathbf{y} - \mathbf{v}_\mu; \mathbf{v}_i) d\mathbf{y} d\mathbf{x}.$$

By definition of D_i we have for $\mu \in D_i$ that

$$\mathbf{v}_\mu + \mathbf{v}_i = \mathbf{v}_{\kappa(\mu,i)} \quad \text{with} \quad \kappa(\mu, i) = E^{\mathbf{2e}}(\mathbf{v}_i + \mathbf{v}_\mu)$$

and since

$$\begin{aligned} &F(\mathbf{x} - \mathbf{v}_\nu, \mathbf{y} - \mathbf{v}_\mu; \mathbf{v}_i) \\ =&(\hat{\varphi}(\mathbf{y} - \mathbf{v}_{\kappa(\mu,i)}) - \hat{\varphi}(\mathbf{x} - \mathbf{v}_{\kappa(\nu,i)}))(\hat{\varphi}(\mathbf{y} - \mathbf{v}_\mu) - \hat{\varphi}(\mathbf{x} - \mathbf{v}_\nu))\gamma(h(\mathbf{y} - \mathbf{v}_\mu), h(\mathbf{x} - \mathbf{v}_\nu)) \end{aligned}$$

we find due to (5.3) that

$$\begin{aligned} &\sum_{\nu \in D_i} \sum_{\mu \in D_i} \int_\square \int_\square F(\mathbf{x} - \mathbf{v}_\nu, \mathbf{y} - \mathbf{v}_\mu; \mathbf{v}_i) d\mathbf{y} d\mathbf{x} \\ =& \sum_{\nu \in D_i} \sum_{\mu \in D_i} \int_\square \int_\square ((\hat{\psi}_{\kappa(\mu,i)}(\mathbf{y}) - \hat{\psi}_{\kappa(\nu,i)}(\mathbf{x}))(\hat{\psi}_\mu(\mathbf{y}) - \hat{\psi}_\nu(\mathbf{x})))\gamma(h(\mathbf{y} - \mathbf{v}_\mu), h(\mathbf{x} - \mathbf{v}_\nu)) d\mathbf{y} d\mathbf{x}. \end{aligned}$$

We separate the case $\mu = \nu$ since the kernel may have singularities at $(\mathbf{x}, \mathbf{y})$ with $\mathbf{x} = \mathbf{y}$ and therefore these integrals need a different numerical treatment. At this point we find that

$$\begin{aligned} &\sum_{\nu \in D_i} \int_\square \int_\square ((\hat{\psi}_{\kappa(\nu,i)}(\mathbf{y}) - \hat{\psi}_{\kappa(\nu,i)}(\mathbf{x}))(\hat{\psi}_\nu(\mathbf{y}) - \hat{\psi}_\nu(\mathbf{x})))\gamma(h\mathbf{y}, h\mathbf{x}) d\mathbf{y} d\mathbf{x} \\ =&|D_i| \int_\square \int_\square ((\hat{\psi}_i(\mathbf{y}) - \hat{\psi}_i(\mathbf{x}))(\hat{\psi}_0(\mathbf{y}) - \hat{\psi}_0(\mathbf{x})))\gamma(h\mathbf{y}, h\mathbf{x}) d\mathbf{y} d\mathbf{x}. \end{aligned}$$

This simplification follows from some straightforward transformations exploiting assumption (A2) and the fact that each element basis function $\hat{\psi}_\nu$ can be expressed by $\hat{\psi}_0$ through the relation $\hat{\psi}_\nu \circ g_\nu = \hat{\psi}_0$, where $g_\nu(\mathbf{x}) := \mathbf{v}_\nu + \mathbf{R}_\nu \mathbf{x}$ for an appropriate rotation matrix

$\mathbf{R}_\nu$. With this observation we also find that the other two integrals in `sing` are equal, i.e.,

$$\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_k}\int_{\hat{\mathcal{S}}_k^c\cap\hat{\mathcal{S}}} F(\mathbf{x},\mathbf{y};E^{\mathbf{L}}(k))d\mathbf{y}d\mathbf{x} = \int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_k}\int_{\hat{\mathcal{S}}_k\cap\hat{\mathcal{S}}^c} F(\mathbf{x},\mathbf{y};E^{\mathbf{L}}(k))d\mathbf{y}d\mathbf{x}.$$

By exploiting again that $\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_k = \bigcup_{\nu\in D_i}\square-\mathbf{v}_\nu$ and thus $\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_k^c = \bigcup_{\nu\in\{0,\dots,2^d-1\}\setminus D_i}\square-\mathbf{v}_\nu$ we obtain

$$\begin{aligned}&\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_k}\int_{\hat{\mathcal{S}}_k^c\cap\hat{\mathcal{S}}} F(\mathbf{x},\mathbf{y};E^{\mathbf{L}}(k))d\mathbf{y}d\mathbf{x}\\ &= -\sum_{\nu\in D_i}\sum_{\mu\in\{0,\dots,2^d\}\setminus D_i}\int_\square\int_\square \hat{\psi}_{\kappa(\nu,i)}(\mathbf{x})(\hat{\psi}_\mu(\mathbf{y})-\hat{\psi}_\nu(\mathbf{x}))\gamma(h(\mathbf{v}_\nu-\mathbf{v}_\mu),h(\mathbf{x}-\mathbf{y}))d\mathbf{y}d\mathbf{x}.\end{aligned}$$

Note that by definition of D_i we have that for $\mu\in\{0,\dots,2^d\}\setminus D_i$ there is no $\kappa\in\{0,\dots,2^d\}$ such that $\mathbf{v}_\mu+\mathbf{v}_i=\mathbf{v}_\kappa$ and therefore $\hat{\varphi}(\mathbf{y}-(\mathbf{v}_\mu+\mathbf{v}_i))=0$ for all $\mathbf{y}\in\square$. All in all we have

$$\begin{aligned}&\mathtt{sing}_k/h^{2d} =\\ &|D_i|\int_\square\int_\square((\hat{\psi}_i(\mathbf{y})-\hat{\psi}_i(\mathbf{x}))(\hat{\psi}_0(\mathbf{y})-\hat{\psi}_0(\mathbf{x})))\gamma(h\mathbf{y},h\mathbf{x})d\mathbf{y}d\mathbf{x}\\ &+\sum_{\nu\in D_i}\Bigg[\sum_{\mu\in D_i\mu\neq\nu}\int_\square\int_\square((\hat{\psi}_{\kappa(\mu,i)}(\mathbf{y})-\hat{\psi}_{\kappa(\nu,i)}(\mathbf{x}))(\hat{\psi}_\mu(\mathbf{y})-\hat{\psi}_\nu(\mathbf{x})))\\ &\qquad\qquad\cdot\gamma(h(\mathbf{v}_\mu-\mathbf{v}_\nu),h(\mathbf{y}-\mathbf{x}))d\mathbf{y}d\mathbf{x}\\ &-4\sum_{\mu\in\{0,\dots,2^d\}\setminus D_i}\int_\square\int_\square\hat{\psi}_{\kappa(\nu,i)}(\mathbf{x})(\hat{\psi}_\mu(\mathbf{y})-\hat{\psi}_\nu(\mathbf{x}))\gamma(h(\mathbf{v}_\mu-\mathbf{v}_\nu),h(\mathbf{y}-\mathbf{x}))d\mathbf{y}d\mathbf{x}\Bigg].\end{aligned}$$

A.1.2 Vector "rad"

Let $i\in\mathtt{idx}_s^i$, such that $E^{\mathbf{L}}(k)=\mathbf{v}_i$. Proceeding as above we obtain

$$\begin{aligned}\mathtt{rad}_k &= 2h^{2d}\int_{\hat{\mathcal{S}}\cap\hat{\mathcal{S}}_k}\int_{(\hat{\mathcal{S}}\cup\hat{\mathcal{S}}_k)^c}F(\mathbf{x},\mathbf{y};\mathbf{v}_i)d\mathbf{y}d\mathbf{x}\\ &= 2h^{2d}\sum_{\nu\in D_i}\int_\square\hat{\psi}_\nu(\mathbf{x})\hat{\psi}_{\kappa(\nu,i)}(\mathbf{x})\int_{(\hat{\mathcal{S}}\cup\hat{\mathcal{S}}_k)^c}\gamma(h(\mathbf{x}-\mathbf{v}_\nu),h\mathbf{y})(h\mathbf{y})d\mathbf{y}d\mathbf{x}.\end{aligned}$$

Let us define

$$P_\nu(\mathbf{x}) := \int_{(\hat{\mathcal{S}}\cup\hat{\mathcal{S}}_k)^c}\phi(h(\mathbf{x}-\mathbf{v}_\nu),h\mathbf{y})\chi_{B_{\delta,\infty}(h(\mathbf{x}-\mathbf{v}_\nu))}(h\mathbf{y})d\mathbf{y}, \tag{A.1}$$

where we consider $\gamma(\mathbf{x},\mathbf{y}) = \phi(\mathbf{x},\mathbf{y})\chi_{B_{\delta,\infty}(\mathbf{x})}(\mathbf{y})$. Again, we can use $\hat{\psi}_\nu \circ g_\nu = \hat{\psi}_0$ in order to show by some straightforward transformations that

$$\int_\square \hat{\psi}_0(\mathbf{x})\hat{\psi}_i(\mathbf{x})P_0(\mathbf{x})d\mathbf{x} = \int_\square \hat{\psi}_\nu(\mathbf{x})\hat{\psi}_{\kappa(\nu,i)}(\mathbf{x})P_\nu(\mathbf{x})d\mathbf{x}$$

for all $\nu \in D_i$. Hence, we arrive at

$$\mathtt{rad}_k = 2h^{2d}|D_i| \int_\square \hat{\psi}_0(\mathbf{x})\hat{\psi}_i(\mathbf{x})P_0(\mathbf{x})d\mathbf{x}.$$

A.1.3 Vector "dis"

Now we distinguish between the case where $\mathtt{rad}_k$ and $\mathtt{sing}_k$ are zero and the complement case. First let $i \in \mathtt{idx}_s^i$, such that $E^{-\mathbf{L}}(k) = \mathbf{v}_i$, then we find

$$\begin{aligned}\mathtt{dis}_k &:= 2h^{2d} \int_{\hat{\mathcal{S}}_k^c \cap \hat{\mathcal{S}}} \int_{\hat{\mathcal{S}}_k \cap \hat{\mathcal{S}}^c} F(\mathbf{x},\mathbf{y}; E^{-\mathbf{L}}(k))d\mathbf{y}d\mathbf{x} \\ &= -2h^{2d} \sum_{\nu \in \{0,\dots,2^d\}\setminus D_i} \sum_{\mu \in D_i^c} \int_\square \int_\square \hat{\psi}_\mu(\mathbf{y})\hat{\psi}_\nu(\mathbf{x})\gamma(h(\mathbf{v}_\mu - \mathbf{v}_\nu - \mathbf{v}_i), h(\mathbf{y}-\mathbf{x}))d\mathbf{y}d\mathbf{x}.\end{aligned}$$

Now let $k \neq E^{-\mathbf{L}}(\mathbf{v}_i)$ for any $0 \leq i < 2^d$, then $\hat{\mathcal{S}}_k^c \cap \hat{\mathcal{S}} = \hat{\mathcal{S}}$ and $\hat{\mathcal{S}}_k \cap \hat{\mathcal{S}}^c = \hat{\mathcal{S}}_k$ such that

$$\begin{aligned}\mathtt{dis}_k =& 2h^{2d} \int_{\hat{\mathcal{S}}} \int_{\hat{\mathcal{S}}_k} F(\mathbf{x},\mathbf{y}; E^{-\mathbf{L}}(k))d\mathbf{y}d\mathbf{x} \\ =& -2h^{2d} \int_{\hat{\mathcal{S}}} \int_{\hat{\mathcal{S}}_k} \hat{\varphi}(\mathbf{y} + E^{-\mathbf{L}}(k))\hat{\varphi}(\mathbf{x})\gamma(h(\mathbf{y} + E^{-\mathbf{L}}(k)), h\mathbf{x})d\mathbf{y}d\mathbf{x}.\end{aligned}$$

Since by definition $\hat{\mathcal{S}}_k = E^{-\mathbf{L}}(k) + \hat{\mathcal{S}}$ and $\hat{\mathcal{S}} = \dot{\bigcup}_{i=0}^{2^d}(\square - \mathbf{v}_i)$, we obtain

$$\begin{aligned}&\mathtt{dis}_k \\ =& -2h^{2d} \sum_{0 \leq \nu < 2^d} \sum_{0 \leq \mu < 2^d} \int_\square \int_\square \hat{\psi}_\nu(\mathbf{x})\hat{\psi}_\mu(\mathbf{y})\gamma(h(\mathbf{v}_\mu - E^{-\mathbf{L}}(k) - \mathbf{v}_\nu), h(\mathbf{y}-\mathbf{x}))d\mathbf{y}d\mathbf{x}.\end{aligned}$$

All in all we conclude that, if $k = E^{\mathbf{L}}(\mathbf{v}_i)$, then

$$\begin{aligned}&\frac{\mathtt{dis}_k}{-2h^{2d}} \\ =& \sum_{\nu \in \{0,\dots,2^d\}\setminus D_i} \sum_{\mu \in D_i^c} \int_\square \int_\square \hat{\psi}_\nu(\mathbf{x})\hat{\psi}_\mu(\mathbf{y})\gamma(h(\mathbf{v}_\mu - \mathbf{v}_\nu - E^{-\mathbf{L}}(k)), h(\mathbf{y}-\mathbf{x}))d\mathbf{y}d\mathbf{x}.\end{aligned}$$

Elsewise we have

$$\frac{\mathtt{dis}_k}{-2h^{2d}} = \sum_{0\le\nu<2^d}\sum_{0\le\mu<2^d}\int_\square\int_\square \hat\psi_\nu(\mathbf{x})\hat\psi_\mu(\mathbf{y})\gamma(h(\mathbf{v}_\mu-\mathbf{v}_\nu-E^{-\mathbf{L}}(k)),h(\mathbf{y}-\mathbf{x}))d\mathbf{y}d\mathbf{x}.$$

A.1.4 Source term

We compute $\mathbf{f}^h\in\mathbb{R}^{\mathbf{L}^d}$ by

$$\begin{aligned} f_k^h &= \int_\Omega f\varphi_k d\mathbf{x} \\ &= h^d\sum_{\nu=0}^{2^d-1}\int_\square f(\mathbf{x}_k+h(\mathbf{v}-\mathbf{v}_\nu))\hat\psi_\nu(\mathbf{v})d\mathbf{v} \\ &= h^d 2^d\int_\square f(\mathbf{x}_k+h\mathbf{v})\hat\psi_0(\mathbf{v})d\mathbf{v}, \end{aligned}$$

where the last equality follows again from considering $\hat\psi_\nu\circ g_\nu=\hat\psi_0$.

A.2 Numerical computation of the integrals

In this section we want to point out how we numerically handle the occurring integrals.

A.2.1 Nonsingular integrals

We mainly have to compute integrals of the form

$$\int_\square\int_\square g(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x},$$

where $\square=[0,1]^d$ and $g\colon\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$ is a (typically smooth) function, which we assume to have no singularities in the domain $\square\times\square$. We approximate the value of this integral by employing a n-point Gauss-Legendre quadrature rule in each dimension. More precisely, we built a d-dimensional tensor grid $\mathtt{X}\in\mathbb{R}^{d\times n^d}$ with associated weights $\mathtt{W}^{single}\in\mathbb{R}^{n^d}$, such that $\int_\square g(\mathbf{x},\mathbf{y})d\mathbf{y}\approx\sum_{i=0}^{n^d-1}g(\mathbf{x},\mathtt{X}_i)\mathtt{W}_i^{single}$ for $\mathbf{x}\in\mathtt{X}$. Finally, we define the arrays

$$\mathtt{V}:=(\mathtt{X},\mathtt{X},\dots,\mathtt{X})\in\mathbb{R}^{d\times n^{2d}},$$

$$\mathtt{Q} := (\mathtt{X}_0, \dots, \mathtt{X}_0, \mathtt{X}_1, \dots, \mathtt{X}_1, \dots, \mathtt{X}_{n^d-1}, \dots, \mathtt{X}_{n^d-1}) \in \mathbb{R}^{d \times n^{2d}}$$

with associated weights $\mathtt{W}^{double} \in \mathbb{R}^{n^{2d}}$, such that we finally arrive at the following quadrature rule:

$$\int_{\square}\int_{\square} g(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x} \approx \sum_{i=0}^{n^{2d}-1} g(\mathtt{V}_i, \mathtt{Q}_i)\mathtt{W}_i^{double} = \left(g(\mathtt{V},\mathtt{Q}) \cdot \mathtt{W}^{double}\right).sum().$$

A.2.2 Singular integrals

The kernel function may come along with singularities at $(\mathbf{x},\mathbf{y})$ with $\mathbf{x} = \mathbf{y}$ (see, e.g., the fractional kernel (6.24)). Therefore we start with a general observation, which paves the way for numerically handling these singularities. Let $g\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be a symmetric function, i.e., $g(\mathbf{x},\mathbf{y}) = g(\mathbf{y},\mathbf{x})$, and let us further define the sets

$$M := \{(\mathbf{x},\mathbf{y}) \in \square \times \square \colon y_d \in [0, x_d]\}$$

and

$$M' := \{(\mathbf{x},\mathbf{y}) \in \square \times \square \colon (\mathbf{y},\mathbf{x}) \in M\}.$$

Then it is straightforward to show that $M \cup M' = \square \times \square$ and

$$\begin{aligned} M \cap M' &= \{(\mathbf{x},\mathbf{y}) \in \square \times \square \colon x_d = y_d\} \\ &= \left\{(v^{d-1}, z, w^{d-1}, z) \in \mathbb{R}^{2d} \colon (v^{d-1}, w^{d-1}, z) \in [0,1]^{2d-1}\right\}, \end{aligned}$$

such that $\lambda_{2d}(M \cap M') = 0$. Hence, we find that

$$\begin{aligned} \int_{\square\times\square} g d\lambda_{2d} &= \int_{M\cup M'} g d\lambda_{2d} = \int_M g d\lambda_{2d} + \int_{M'} g d\lambda_{2d} - \int_{M\cap M'} g d\lambda_{2d} \\ &= \int_M g d\lambda_{2d} + \int_{M'} g d\lambda_{2d}. \end{aligned}$$

From the symmetry of g we additionally deduce that

$$\int_M g d\lambda_{2d} = \int_{M'} g d\lambda_{2d}$$

and therefore we finally obtain

$$\int_{\square\times\square} g d\lambda_{2d} = 2\int_M g d\lambda_{2d}.$$

This observation can now be applied to the singular integrals occurring in the vector `sing` such that

$$\begin{aligned} &\int_{\square}\int_{\square}((\hat{\psi}_i(\mathbf{y}) - \hat{\psi}_i(\mathbf{x}))(\hat{\psi}_0(\mathbf{y}) - \hat{\psi}_0(\mathbf{x})))\gamma(h\mathbf{y}, h\mathbf{x})d\mathbf{y}d\mathbf{x} \\ =&2\int_{[0,1]^d}\int_{[0,1]^{d-1}\times[0,x_{d-1}]}((\hat{\psi}_i(\mathbf{y}) - \hat{\psi}_i(\mathbf{x}))(\hat{\psi}_0(\mathbf{y}) - \hat{\psi}_0(\mathbf{x})))\gamma(h\mathbf{y}, h\mathbf{x})d\mathbf{y}d\mathbf{x}. \end{aligned}$$

The essential advantage of this representation relies on the fact that the singularities are now located on the boundary of the integration domain. Thus, we do not evaluate the integrand on its singularities while using quadrature points which lie in the interior. We extend the one-dimensional adaptive (G7,K15)-Gauss-Kronrod quadrature rule to d-dimensional integrals by again tensorizing the one-dimensional rule. Moreover, in order to take full advantage of the Gauss-Kronrod quadrature, we divide the set $[0,1]^{d-1} \times [0, x_{d-1}]$ into 2^{d-1} disjoint rectangular subsets such that the singularity $\mathbf{x}$ is located at a vertex. The latter partitioning reinforces the adaptivity property of the Gauss-Kronrod quadrature rule.

A.2.3 Integrals with large interaction horizon

Now we discuss the quadrature of

$$\begin{aligned} P_0(\mathbf{x}) = & \int_{(\hat{\mathcal{S}} \cup \hat{\mathcal{S}}_k)^c} \phi(h\mathbf{x}, h\mathbf{y}) \chi_{B_{\delta,\infty}(h\mathbf{x})}(h\mathbf{y}) d\mathbf{y} \\ = & \ (1/h^d) \int_{(\mathcal{S}_0 \cup \mathcal{S}_k)^c} \phi(h\mathbf{x}, \mathbf{y} - (\mathbf{a} + h\mathbf{e})) \chi_{B_{\delta,\infty}(h\mathbf{x})}(\mathbf{y} - (\mathbf{a} + h\mathbf{e})) d\mathbf{y} \end{aligned}$$

from (A.1).

Remark A.2.1. *By the time this code has been implemented, the author was not aware of the implementation recipe suggested in Remark 6.2.4, so that a coarsening strategy is considered in the following to cope with the difficulty to numerically integrate over large integration domains.*

We carry out two simplifications for the implementation, which are mainly motivated by the fact that $\gamma(\mathbf{x}, \mathbf{y}) \to 0$ as $y \to \infty$ for kernels such as the fractional Laplace kernel. First, since $h \to 0$, we set

$$B_{\delta,\infty}(h\mathbf{x}) = B_{\delta,\infty}(0).$$

Especially when δ is large and h small, this simplification does not significantly affect the value of the integral. Second, we employ the $||\cdot||_\infty$-norm for the ball $B_{\delta,\infty}(0)$ instead of the $||\cdot||_2$-norm. This is justified in Section 6.2. Hereby we can use our quadrature rules for rectangular elements. Consequently, we are concerned with the quadrature of the integral

$$\int_{B_{\delta,\infty}(\mathbf{a}+h) \setminus (\mathcal{S}_0 \cup \mathcal{S}_k)} \phi(h\mathbf{x}, \mathbf{y} - (\mathbf{a} + h\mathbf{e})) d\mathbf{y}$$

for $\mathbf{x} \in \mathtt{X}$. For this purpose, we define the box

$$\mathcal{B} := \prod_{i=0}^{d-1} [a_i - \lambda, a_i + \lambda]$$

for a constant $\delta \geq \lambda > 2h$ such that $(\mathcal{S}_0 \cup \mathcal{S}_k) \subset \mathcal{B}$ and we partition

$$B_{\delta,\infty}(\mathbf{a} + h) = \mathcal{B} \cup B_{\delta,\infty}(\mathbf{a} + h\mathbf{e}) \setminus \mathcal{B}.$$

Thus, we obtain

$$P_0(\mathbf{x})h^d \approx \int_{\mathcal{B}\setminus(\mathcal{S}_0\cup\mathcal{S}_k)} \phi(h\mathbf{x}, \mathbf{y}-(\mathbf{a}+h\mathbf{e}))d\mathbf{y} + \int_{B_{\delta,\infty}(\mathbf{a}+h\mathbf{e})\setminus\mathcal{B}} \phi(h\mathbf{x}, \mathbf{y}-(\mathbf{a}+h\mathbf{e}))d\mathbf{y}.$$

We discretize $\mathcal{B}$ with the same elements which we used for Ω. This is convenient for two reasons. On the one hand we capture the critical values of the kernel, which in case of singular kernels typically decrease as $\mathbf{y}\to\infty$. On the other hand, we have to leave out the integration over $(\mathcal{S}_0\cup\mathcal{S}_k)$, which then can easily be implemented since we use the same discretization. However, this results in $\mathbf{N}_2^d$, where $N_2^i = 2\lambda/h$, hypercubes with base points

$$\mathbf{y}_j := (\mathbf{a}-\lambda\mathbf{e}) + hE^{-\mathbf{N}_2}(j)$$

such that

$$\begin{aligned}\int_{\mathcal{B}\setminus(\mathcal{S}_0\cup\mathcal{S}_k)} \phi(h\mathbf{x},\mathbf{y})d\mathbf{y} &= \sum_{j=0,j\notin R_i}^{\mathbf{N}_2^d-1} \int_{y_j+h\mathbf{e}\square} \phi(h\mathbf{x}, \mathbf{y}-(\mathbf{a}+h\mathbf{e}))d\mathbf{y} \\ &= h^d \sum_{j=0,j\notin R_i}^{\mathbf{N}_2^d-1} \int_{\square} \phi(\lambda\mathbf{e}+h(\mathbf{e}-E^{-\mathbf{N}_2}(j)), h(\mathbf{y}-\mathbf{x}))d\mathbf{y},\end{aligned}$$

where R_i contains the indices for those elements, which are contained in $(\mathcal{S}_0\cup\mathcal{S}_k)$. This set can be characterized as

$$R_i := \left\{0\le j<\mathbf{N}_2^d-1\colon \mathbf{y}_j+h\square\subset(\mathcal{S}_0\cup\mathcal{S}_k)\right\}.$$

Note that by definition we have $\mathcal{S}_0 = \mathbf{a}+h(\mathbf{e}+\bigcup_{0\le\nu<2^d}\square-\mathbf{v}_\nu)$ and since $E^{-\mathbf{L}}(k) = \mathbf{v}_i$ for $k\in \mathtt{idx}_s^k$ such that $\mathbf{x}_k = \mathbf{a}+h(\mathbf{e}+\mathbf{v}_i)$ we know that

$$\mathcal{S}_0^c\cap\mathcal{S}_k = \mathbf{a}+h(\mathbf{e}+\mathbf{v}_i+\bigcup_{\nu\in D_i^c}\square-\mathbf{v}_\nu).$$

Hence,

$$\begin{aligned}(\mathbf{y}_j+h\square)\subset(\mathcal{S}_0\cup\mathcal{S}_k)\Leftrightarrow\ & \mathbf{y}_j=\mathbf{a}+h(\mathbf{e}-\mathbf{v}_\nu)\ \text{ for }\ 0\le\nu<2^d \\ & \textit{or}\ \ \mathbf{y}_j=\mathbf{a}+h(\mathbf{e}+\mathbf{v}_i-\mathbf{v}_\nu)\ \text{ for }\ \nu\in D_i^c.\end{aligned}$$

Since $\mathbf{y}_j = \mathbf{a}-\lambda\mathbf{e}+hE^{-\mathbf{N}_2}(j)$ we find by equating $\mathbf{y}_j$ with these requirements that

$$R_i = \left\{E^{\mathbf{N}_2}(\mathbf{e}-\mathbf{v}_\nu+\lambda h^{-1}\mathbf{e}) : 0\le\nu<2^d\right\}\cup\left\{E^{\mathbf{N}_2}(\mathbf{e}+\mathbf{v}_i-\mathbf{v}_\nu+\lambda h^{-1}\mathbf{e}) : \nu\in D_i^c\right\}.$$

Now we discuss the quadrature of the second integral

$$\int_{B_{\delta,\infty}(\mathbf{a}+h\mathbf{e})\setminus\mathcal{B}} \phi(h\mathbf{x},\mathbf{y}-(\mathbf{a}+h\mathbf{e}))d\mathbf{y}$$

for $\mathbf{x} \in \mathtt{X}$. Since we want to apply the algorithm to fractional diffusion, we have to get around the computational costs that occur when δ is large. In order to alleviate those costs we follow the idea in [33] and apply a coarsening rule to discretize the domain $B_{\delta,\infty}(\mathbf{a}+h\mathbf{e})\backslash\mathcal{B}$. Assume we have a procedure which outputs a triangulation $(z,\hat{h})$ of $B_{\delta,\infty}(\mathbf{a}+h\mathbf{e})\backslash\mathcal{B}$ consisting of N_3 hyperrectangles with base points $\mathbf{z}_j \in \mathbb{R}^d$ and sides of length $\hat{h}_j \in \mathbb{R}^d$. Then we obtain

$$\begin{aligned}\int_{B_{\delta,\infty}(\mathbf{a}+h\mathbf{e})\backslash\mathcal{B}} \phi(h\mathbf{x}, \mathbf{y}-(\mathbf{a}+h\mathbf{e}))d\mathbf{y} &= \sum_{j=0}^{N_3-1} \int_{\mathbf{z}_j+\hat{h}_j\square} \phi(h\mathbf{x}, \mathbf{y}-(\mathbf{a}+h\mathbf{e}))d\mathbf{y} \\ &= \sum_{j=0}^{N_3-1} \hat{h}_j^d \int_{\square} \phi(-\mathbf{z}_j+\mathbf{a}+h\mathbf{e}, \hat{h}_j\mathbf{y}-h\mathbf{x})d\mathbf{y}.\end{aligned}$$

All in all we thus have

$$\begin{aligned}\mathtt{rad}_k \approx\ & 2|D_i|h^{2d} \sum_{j=0, j\notin R_i}^{\mathbf{N}_2^d-1} \int_{\square}\int_{\square} \hat{\psi}_0(\mathbf{x})\hat{\psi}_i(\mathbf{x})\phi(\lambda\mathbf{e}+h(\mathbf{e}-E^{-\mathbf{N}_2}(j)), h(\mathbf{y}-\mathbf{x}))d\mathbf{y}d\mathbf{x} \\ &+ 2|D_i|h^d \sum_{j=0}^{N_3-1} \hat{h}_j^d \int_{\square}\int_{\square} \hat{\psi}_0(\mathbf{x})\hat{\psi}_i(\mathbf{x})\phi(-\mathbf{z}_j+\mathbf{a}+h\mathbf{e}, \hat{h}_j\mathbf{y}-h\mathbf{x})d\mathbf{y}d\mathbf{x}.\end{aligned}$$

This leaves space for discussion concerning the choice of an optimal coarsening rule. We use the following simple approach in our code: We decompose $B_{\delta,\infty}(\mathbf{a}+h\mathbf{e})\backslash\mathcal{B}$ into (3^d-1) d-dimensional hyperrectangles surrounding the box $\mathcal{B}$. Then we build a tensor grid by employing in each dimension the coarsening strategy

$$\mathbf{v} + i^q h_{min}\mathbf{e},$$

where $\mathbf{v}$ is a vertex of $\mathcal{B}$, $q \geq 1$ the coarsening parameter and $h_{min} > 0$ a minimum grid size. By concatenating all arrays we obtain a triangulation $(z,\hat{h})$ of $B_{\delta,\infty}(\mathbf{a}+h\mathbf{e})\backslash\mathcal{B}$.

A.3 Final remarks

Depending on the dimension, we only have to adapt the index sets $\mathtt{idx}_s^i$, $\mathtt{idx}_s^k$, D_i and D_i^c, the implementation of the quadrature rules discussed above and the 2^d element basis functions. The rest can be implemented in a generic way. In addition to that, we framed above the relevant representations for implementing the assembly of the first row.
Furthermore, along the assembling process of the first row, we first compute the more challenging $(d+1)$ entries

$$2\mathbf{R}_k^h = \mathtt{sing}_k + \mathtt{rad}_k + \mathtt{dis}_k \quad \text{for} \quad k \in \mathtt{idx}_s^k.$$

Here we parallelize the computations of the integrals in $\mathtt{rad}_k$ over the base points $\mathbf{y}_j$ for the box $\mathcal{B}$ and $\mathbf{z}_j$ for the coarsening strategy. For the remaining $\mathbf{L}^d - (d+1)$ indices we have that

$$2\mathbf{R}_k^h = \mathtt{dis}_k$$

and we can simply parallelize the loop over $\left\{0 \leq k < \mathbf{L}^d\right\} \setminus \mathtt{idx}_s^k$.