Franziska Nestler

Efficient Computation of Electrostatic Interactions in Particle Systems Based on Nonequispaced Fast Fourier Transforms

Franziska Nestler

Efficient Computation of Electrostatic Interactions in Particle Systems Based on Nonequispaced Fast Fourier Transforms

TECHNISCHE UNIVERSITÄT
CHEMNITZ

Universitätsverlag Chemnitz
2018

Impressum

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Angaben sind im Internet über http://dnb.d-nb.de abrufbar.

Titelgrafik: Franziska Nestler
Satz/Layout: Franziska Nestler

Technische Universität Chemnitz/Universitätsbibliothek
Universitätsverlag Chemnitz
09107 Chemnitz
http://www.tu-chemnitz.de/ub/univerlag

readbox unipress
in der readbox publishing GmbH
Am Hawerkamp 31
48155 Münster
http://unipress.readbox.net

ISBN 978-3-96100-054-8

http://nbn-resolving.de/urn:nbn:de:bsz:ch1-qucosa2-233760

Efficient Computation of Electrostatic Interactions in Particle Systems Based on Nonequispaced Fast Fourier Transforms

von der Fakultät für Mathematik
der Technischen Universität Chemnitz
genehmigte

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt

von Dipl.-Math. Franziska Nestler

Tag der Einreichung: 30. November 2017

Betreuer: Prof. Dr. Daniel Potts, Technische Universität Chemnitz

Gutachter: Prof. Dr. Daniel Potts, Technische Universität Chemnitz
Prof. Dr. Gabriele Steidl, Universität Kaiserslautern
Prof. Dr. Christian Holm, Universität Stuttgart

Tag der öffentlichen Prüfung: 5. April 2018

Contents

1 Introduction

Computer simulations have become an essential tool for studying complex systems in various branches of science. In order to describe physical phenomena based on appropriate models, mathematics acts as a basic language. In the present age of modern computer technologies, especially the development of efficient algorithms is of high interest. Naive approaches for the computation of certain quantities of interest often result in algorithms with a poor complexity, i.e., large or rather complex systems are not treatable within acceptable time. For this reason, one makes use of adequate tools of applied mathematics in order to develop methods showing a significantly lower complexity. This is usually connected with a loss of accuracy, i.e., one has to be aware of upcoming approximation errors. Having the resulting errors under control brings the user in the profitable position to find an adequate balance between accuracy and required runtime. Finally, applying adequate efficient algorithms in combination with modern computer technology allows the study of highly complex natural phenomena involving a large number of influencing factors. Based on such investigations one may be able to confirm theoretical assumptions or gain new insights in processes, which cannot be studied based on real experiments.

The present thesis is largely concerned with an application originating from the area of electrostatics. We consider problems arising in molecular dynamics simulations, which find their applications in different areas of biology, physics, chemistry and material science, see [FS02, Cai13] for instance. Within the scope of such investigations, one considers systems containing hundreds, thousands or even millions of interacting charged particles or molecules and aims to determine the resulting electrostatic potentials, the acting forces or the overall electrostatic energy, for instance. Based on the results, one may be able to study the dynamic evolution of particle systems, study physical phenomena on atomic level or describe the behavior or properties of materials that have not been developed so far.

Since each particle is interacting with all other particles, doubling the number of particles results in an increase in the number of interactions by a factor of four. Simply speaking, naive approaches to solve this problem will result in an arithmetic complexity of $O(N^2)$, where N denotes the number of present particles, which is unfavorable for large system sizes. In several applications it is, furthermore, reasonable to assume periodic boundary conditions in order to avoid undesired boundary effects. This makes the problem definition even more complex. In addition, the considered electrostatic interactions are long ranged. Thus, naive attempts to reduce the required computational costs by applying simple truncation schemes will fail if a high accuracy is required or large particle systems are to be studied. We remark that the classical N-body problem originating from the work of Isaac Newton in the 17th century, where a system of point masses is considered, is very similar to the above described problem concerning electrostatics in charged particle systems.

In 1921, P.P. Ewald proposed a promising approach [Ewa21] to overcome the problems originating from the long range character of the involved interaction kernels. The correspondingly named Ewald summation approach splits up the singular and at the same time long ranged interaction kernel into a singular short range part and a continuous long range part. The short range part can be handled based on an appropriate truncation scheme and the long range part can be transformed into Fourier space, where it shows an extremely rapid convergence behavior.

In many applications periodic boundary conditions are assumed in all three spatial directions. Surprisingly, this is the case most easily treatable, since the underlying analysis is quiet simple. If all involved parameters are chosen appropriately, the Ewald summation approach achieves an arithmetic complexity of $O(N^{3/2})$, cf. [PPD88, KP92, Fin94]. By making use of the well known fast Fourier transform (FFT) this can be further reduced to $O(N \log N)$, i.e., an almost linear scaling with the number of particles is possible. In this context, a fundamental difficulty results from the fact that the particles are usually not distributed on a regular grid. Thus, the present particle distribution has to be approximated based on a regular grid. Algorithms which are based on this idea are commonly known as particle-mesh methods. In order to perform the grid based approximation, these methods make use of a so called window or charge assignment function.

Several particle-mesh methods make use of a B-spline window function, as for instance the well known particle-particle particle-mesh (P^3M) method [EHL80, HE88, DH98a], the particle-mesh Ewald (PME) method [DYP93] and the smooth particle-mesh Ewald (SPME) method [EPB+95]. Also approximations based on a Gaussian have already been considered, see [SKE+05, LT11]. The particle-particle NFFT (P^2NFFT), as proposed in [HL06, PP11, WHP+14], is based on the FFT for nonequispaced data (NFFT) and allows the application of various types of window functions, including B-splines, the Gaussian and also Bessel functions, for instance. Beside the Fourier based methods there are also several other methods, such as the fast multipole method (FMM) [GR87, Dac10] with an arithmetic complexity of $O(N)$, multigrid based approaches [Bol08] or the Barnes-Hut tree algorithm [BH86], just to mention a few.

All the above mentioned algorithms enable an efficient computation of electrostatic interactions in charged particle systems under 3d-periodic boundary conditions, where only charge-charge interactions are present. We remark that [AFH+13] serves a detailed comparison between various methods for the 3d-periodic Coulomb problem, showing that P^3M and P^2NFFT rank among the best solvers in this field. The open software library ScaFaCoS [ABD+] (Scalable fast Coulomb solvers) provides various algorithms for computing interactions in charged particle systems, including P^3M, P^2NFFT and FMM, for instance.

Beside interactions between charged particles, also dipole-dipole and charge-dipole interactions are of interest for certain applications. For pure dipole systems the dipolar P^3M method has already been proposed and applied successfully, see [CBLH08, CBH11].

The Ewald summation approach is also applicable if 2d-periodic or 1d-periodic boundary conditions are assumed. In these cases, the transformation of the resulting long range part into Fourier space does not separate the particle coordinates with respect to the nonperiodic dimensions. Thus, the derivation of efficient algorithms is more challenging than in the 3d-periodic setting. However, a few efficient algorithms have already been proposed, as for example the spectrally accurate Ewald summation for 2d- and 1d-periodic boundary conditions [LT12, ST17] and the electrostatic layer correction (ELC) method for slab geometries [AdJH02]. In addition, several algorithms resulting in a worse complexity have been proposed, see [WA97, YB99, AH02, TMPM02, Bró04, BS04, MMY+04, AH05]. All the mentioned methods for mixed periodic constraints only take charge-charge interactions into consideration.

The main contributions of this work can be summarized as follows. Following the NFFT based fast summation approach [PS03], the P^2NFFT method [PP11] is generalized in order to be applicable for mixed periodic and open boundary conditions, see also [NPP15a, NPP15b]. As a result, the extended P^2NFFT framework allows a unified treatment of all possible combinations of periodic and nonperiodic boundary conditions in the single dimensions. Furthermore, an extension to particle systems including a mixture of charges and dipoles is derived, see also [Nes16b, HNP17, Nes17]. Finally, an $O(N \log N)$ algorithm for the approximation of electrostatic interactions in mixed charge-dipole systems under

various types of periodic boundary conditions is presented for the first time. The reliability of the resulting algorithm is verified based on numerical examples. As already mentioned, the presented method is very efficient but also approximate, i.e., we have to be aware of upcoming approximation errors. Being able to estimate the resulting errors as accurately as possible is required in order to develop strategies for balancing between accuracy and required computational costs. For this reason, also error estimates are taken into consideration.

The approximation errors resulting from truncating the Ewald sums have already been proposed for charge-charge and dipole-dipole systems, see [KP92] and [WH01], respectively. In order to provide reliable estimates for mixed charge-dipole systems, appropriate estimates for interactions between charges and dipoles are developed. The application of the NFFT causes additional approximation errors. These errors are essentially studied based on the already known error formulas for the P^3M method, see [HE88, DH98b, CBLH08, CBH11]. An appropriate extension to charge-dipole systems is presented for the first time. Within this scope, the application of an oversampling factor, as intended by the NFFT approximation scheme, is taken into consideration. For this purpose, the already known error estimates are adapted appropriately. In addition to the well established B-spline window, also other window functions are considered. Based on the error analysis, we gain new insights concerning an optimal choice of the involved parameters, including the oversampling factor as well as the applied window function. These newly gained insights may lead to significant improvements.

Outline of the Thesis

Chapter 2: Electrostatic interactions in particle systems

This chapter is dedicated to basic definitions and concepts concerning the computation of electrostatic interactions in particle systems. We start by introducing some basic terms such as the electrostatic energy, potentials and forces, defined for electrostatic particle systems including point charges and point dipoles, see Section 2.1. In this context, we especially introduce the considered types of periodic boundary conditions and discuss related convergence characteristics. In Section 2.2 the Ewald summation approach will be introduced and discussed. In particular, we present and derive the resulting Ewald representations of the electrostatic potentials, where we allow arbitrary triclinic box structures in case of 3d-periodic boundary conditions, see Section 2.2.1. The cases of 2d-periodic and 1d-periodic boundary conditions are studied in Sections 2.2.2 and 2.2.3, respectively. The case of 0d-periodic or rather open boundary conditions takes a special role. Since periodicity is not applied at all, a derivation of Ewald summation formulas is not possible in the classical sense, cf. Section 2.2.4. Nevertheless, the so called Ewald splitting is also applied for open constraints in order to form the basis for a tool that enables a unified treatment of all the considered types of boundary conditions. Although the resulting Ewald summation formulas appear to be pretty diverse concerning the different types of periodic boundary conditions, the same mathematical ideas form the basis. Consequently, the different cases may be summarized based on a common framework. Such a universal notation of the Ewald summation formulas is introduced by Theorem 2.16.

If periodic boundary conditions are present in at least one of three dimensions, the Ewald summation technique results in a splitting of the considered electrostatic quantities into a short range part and a Fourier space part, which are infinite sums. Both are rapidly converging and hence may be truncated in order to compute the electrostatic potentials, forces and the like efficiently. Of course, this process causes certain approximation or rather truncation errors. In order to develop reliable algorithms, an estimation or rather prediction of these errors is of high interest. In Section 2.3 we discuss the

estimation of these truncation errors. For this purpose, we restrict our considerations to the 3d-periodic case, which is the simplest case in terms of the underlying analysis. We study the rms error in the forces, which is a common measure of accuracy in the field of molecular dynamics simulations. After reviewing the basic concepts, we consider the short range part and the Fourier space part separately in Sections 2.3.2 and 2.3.3, respectively. We revisit the already known error estimates for charge-charge [KP92] as well as dipole-dipole interactions [WH01] and extend the theory to mixed charge-dipole systems, as introduced in [HNP17]. The Ewald truncation errors are summarized in Theorems 2.18 and 2.20, where the short range parts of the forces are studied, as well as in Theorems 2.23 and 2.24, where the Fourier space parts are estimated. Based on the presented estimates we are for the first time able to tune the Ewald summation parameters for mixed charge-dipole systems with respect to a certain rms force accuracy, see Section 2.3.4.

Chapter 3: Nonequispaced fast Fourier transforms in three variables

Chapter 3 is concerned with the nonuniform FFT (NUFFT) or rather FFT for nonequispaced data (NFFT), cf. [DR93, Bey95, Ste98, DS99, War98, PST01, GL04, KKP09]. After reviewing some basic terms and definitions in Section 3.1, we give an introduction to the NFFT algorithm and some of its variations in Section 3.2. First of all, the basic concepts of the classical NFFT and its adjoint are considered, cf. Section 3.2.1. In accordance with the investigation of rms errors concerning the Ewald summation approach, see Section 2.3, we take into consideration the approximation errors measured in the $\mathcal{L}_2$-norm and revisit results concerning the $\mathcal{L}_2$-optimization of the presented algorithms, cf. [DS99, Jac09, Pip15].

The NFFT and its adjoint act as the main building blocks of the P^2NFFT method for charged particle systems. In order to compute the acting forces, a further variant of the NFFT, namely the so called gradient NFFT, has already been developed. For the computation of interactions in mixed charge-dipole systems, the conception of two new NFFT modules, namely the Hessian NFFT and the adjoint gradient NFFT, is required, cf. [HNP17]. The gradient NFFT, the Hessian NFFT as well as the adjoint gradient NFFT are introduced in Section 3.2.2. All the presented NFFT modules build the basis for the extended P^2NFFT algorithm, which allows an efficient computation of interactions in mixed charge-dipole systems subject to different types of periodic boundary conditions, see Chapter 4.

A basic element of the NFFT algorithms is the window function, which is used in order to map the given irregularly spaced nodes onto a regular grid. Various possible window functions have already been proposed in the literature, see [Bey95, PS03, Pip15]. Following the results presented in [Nes16a], we concentrate on compactly supported window functions, namely the B-spline as well as the Bessel window. Furthermore, we introduce the modified B-spline window and consider the Gaussian as an example for a window function without compact support.

In order to estimate the resulting approximation errors we derive new error bounds concerning the Bessel window and the modified B-spline window, cf. [Nes16a]. After introducing the relevant window functions as well as the corresponding estimates, we consider the one-dimensional case and show how the NFFT approximation errors may be predicted, see 3.3.1. Therefor, we assume that the underlying Fourier coefficients are known and compare the results concerning two different sets of Fourier coefficients. In addition, we compare the theoretical estimates with measured errors and show that usually a very good agreement is achieved. The Bessel, the modified B-spline as well as the Gaussian window function are defined by introducing a so called shape parameter, which is not needed in case of the classical B-spline window. The presented numerical examples show that the choice of this shape parameter may influence the resulting accuracy significantly and that the introduced error

estimates are a helpful tool in order to tune the shape parameter automatically. Fundamentally different optimal shape parameters are obtained for the considered sets of Fourier coefficients. A comparison between the considered window functions shows that the Bessel window is oftentimes the best choice regarding accuracy. In the context of possible applications, the assumption that the underlying Fourier coefficients are known is certainly not always reasonable. However, in case of electrostatic interactions in particle systems under 3d-periodic boundary conditions this is the case, see Section 2.2.1 and 4.1.1.

The multivariate case is considered in Section 3.3.2. We discuss some fundamental difficulties concerning an exact and efficient prediction of the resulting approximation errors in higher dimensions. In this context, we present some numerical examples in three variables, where we apply the classical B-spline window. In addition to the usual NFFT, we also consider the gradient NFFT. A similar concept will be applied later in Section 4.3, where we also take the Bessel window function into consideration.

The introduced NFFT algorithms are suitable tools for the efficient evaluation of periodic functions or rather trigonometric sums involving unequally distributed nodes. However, there are many applications, where an efficient evaluation of nonperiodic expressions is desired. As an example, the task of computing a convolution of given data on a irregular grid with a nonperiodic kernel function, may be taken into consideration. In order to solve this problem efficiently, the NFFT based fast summation approach, as introduced in Section 3.4, may be applied. The basic idea is to embed a given nonperiodic function into a periodic setting in order to enable an application of the NFFT. We discuss three different variants for continuing a nonperiodic function periodically, which are combined in the P^{2}NFFT method for mixed periodic boundary conditions, see Section 4.1.2. Furthermore, the discussed variants enable to range other approaches for 2d- and 1d-periodic constraints, presented in the literature. Independent from the applied periodization technique, the fast summation approach enables an efficient convolution with nonperiodic kernels by combining the NFFT and its adjoint. In Section 3.4.1 we start by considering the one-dimensional case, followed by a short introduction to fast summation for radial kernels in higher dimensions, cf. Section 3.4.2. The NFFT based fast summation approach builds the basis for the NFFT based fast Ewald summation for mixed periodic and open boundary conditions, cf. Section 4.1.2.

Chapter 4: The P^{2}NFFT method for charges and dipoles

In this chapter we combine the presented Ewald summation formulas, see Section 2.2, and the different NFFT modules, as introduced in Chapter 3, for the approximation of interactions in electrostatic systems with charges and dipoles efficiently. In Section 4.1.1 we consider the 3d-periodic case, where the Ewald formulas straightforwardly lead to an efficient algorithm. An adequate combination of the adjoint NFFT, the adjoint gradient NFFT as well as the NFFT serves an efficient way to approximate the electrostatic potentials within $O(N \log N)$ arithmetic operations, cf. [Pip15, PP11, PP13, Nes16b, HNP17]. In order to compute the electrostatic fields and the field gradients, the NFFT has to be replaced by the gradient NFFT and the Hessian NFFT, respectively.

In Section 4.1.2 the generalization of the method to mixed periodic and open boundary conditions is presented. The computation of the short ranged sums is realized analogously for all considered types of boundary conditions. In contrast, the computation of the long range parts seems to follow fundamentally different principles, since the corresponding Fourier space representations are differing. In order to compute the Fourier space parts efficiently, we apply the NFFT based fast summation approach with respect to the nonperiodic dimensions. Whereas the Fourier coefficients are known in an analytically closed form in case of 3d-periodic constraints, the Fourier coefficients for mixed periodic and open boundary conditions are computed numerically based on the FFT in a precomputation step. However,

the structure of the final algorithm remains exactly the same as in case of 3d-periodic boundary conditions, where only the Fourier coefficients have to be replaced, see also [NPP15a, NPP15b]. Finally, the extended P^{2}NFFT framework enables a unified treatment of mixed charge-dipole systems under various types of periodic boundary conditions.

Some details regarding the implementation of the resulting algorithm are given in Section 4.2, where we also summarize the P^{2}NFFT method for charge-dipole systems under various types of periodic boundary conditions in Algorithm 4.3. Our implementation is part of the above mentioned ScaFaCoS library [ABD$^+$] and makes use of the PNFFT software [PP13, Pipb] in order to perform the required nonuniform FFTs. Replacing the fast NFFT modules by naive slow summation procedures (NDFT) serves a pure Ewald summation, which we call the P^{2}NDFT Ewald method. The current implementation implies certain limitations concerning the shape of the primary simulation box in case of mixed periodic and open boundary conditions. However, it is shown in Section 2.2 that a common triclinic box shape may be traced back to the structure suitable for the present implementation.

The approximation of the truncated Fourier space sums based on the NFFT causes further approximation errors, which are studied in Section 4.3. Essentially, the already developed error formulas for the P^{3}M method, cf. [HE88, DH98b, CBLH08, CBH11], are also applicable for the P^{2}NFFT method, where we additionally have to take into account the presence of an oversampling factor, see also [Nes16c]. Furthermore, the theory is extended to mixed charge-dipole systems, where we consider the so called i$\boldsymbol{k}$-differentiation approach, see Section 4.3.1 and especially Theorem 4.13, as well as the analytic differentiation approach, see Section 4.3.2 and especially Theorem 4.20.

Chapter 5: Numerical results

Chapter 5 contains several numerical examples demonstrating the functional efficiency of the developed algorithm. Furthermore, we verify the theoretical error estimates. In Section 5.1 we start by verifying the Ewald truncation errors for 3d-periodic boundary conditions, as considered in Section 2.3. The actual errors are estimated appropriately for simple orthorhombic box shapes as well as in case of tricilinc structures.

Section 5.2 is concerned with the P^{2}NFFT method for 3d-periodic boundary conditions, where mainly the NFFT related approximation errors are taken into consideration. In Section 5.2.1 we present numerical results verifying the theoretical estimates, where the P^{2}NDFT Ewald method served as a reference method in order to compute the actual approximation errors. In our numerical tests we considered the widely used B-spline window function as well as the Bessel window and compared the results. The real approximation errors are properly estimated from above. Furthermore, the results show that an adequate choice of the shape parameter does improve the performance of the Bessel window function a lot. Consequently, the Bessel window is in many cases the better choice with respect to accuracy. In other words, in order to achieve a certain accuracy, the Bessel window is oftentimes better with respect to the required computation time.

In Section 5.2.2 we discuss how all involved parameters may be tuned in advance in order to achieve a desired accuracy. For this purpose, only the near field cutoff radius has to be specified beforehand. Consequently, all other parameters, including the NFFT oversampling factor as well as the shape parameter for the Bessel window function, may be tuned based on the error estimates. An optimization of the parameters with respect to runtime should depend on the used hard- and software. This may be realized based on the above described accuracy tuning and optimizing the runtime with respect to the near field cutoff radius. The presented numerical examples demonstrate how the required computation time is influenced by the choice of the near field cutoff radius. In addition, the proposed complexity of

the P^{2}NFFT method is verified numerically by considering particle systems of increasing size.

Numerical results for mixed periodic and open boundary conditions are presented in Section 5.3. Whereas the 3d-periodic case is analytically well understood since the underlying Fourier coefficients are known, the derivation of exact error estimation formulas in case of mixed periodic or open boundary conditions is, by now, unsought. Thus, we tune the parameters rather heuristically based on the error analysis for 3d-periodic constraints. Finally, we obtain almost constant approximation errors among all considered types of periodic boundary conditions, where the errors are in some cases somewhat larger if mixed periodic or open boundary conditions are applied.

Chapter 6 : Conclusion

The main results of the thesis are briefly summarized in Chapter 6.

Acknowledgments

First of all, I want to express my sincere gratitude to my adviser Prof. Dr. Daniel Potts, who initiated the topic of the dissertation and supervised my work with great commitment for several years. He offered permanent support through fruitful discussions and frequently contributed valuable impulses. Furthermore, I thank Prof. Dr. Gabriele Steidl and Prof. Dr. Christian Holm, who examined the thesis.

Special thanks go to my former colleague Dr. Michael Pippig. His outstanding competence in the area of particle-mesh methods and especially his continuous support concerning the implementation of the algorithms presented in this thesis promoted my academic work essentially. Far beyond the scope of his own dissertation, he was always willing to extend his software regarding new ideas and offered assistance concerning technical issues. It is not an exaggeration to say that, without his support, an extension of the P^{2}NFFT software to mixed periodic and open boundary conditions as well as to mixed charge-dipole systems would not have been possible in that way.

In the same context, I want to thank Dr. Michael Hofmann, who essentially contributed to the extension of the ScaFaCoS interface to dipole particles and invested time in implementing several relevant software components. In addition, he was always available as a contact person in case of technical questions.

Moreover, I thank Prof. Dr. Christian Holm, Dr. Rudolf Weeber and Florian Weik from the Institute of Computational Physics in Stuttgart for the scientific and application-oriented exchange. Insights into the work at the institute and into possible areas of application provided a basis for further developments of our algorithms. Especially first application-oriented tests of the P^{2}NFFT method for dipolar systems under mixed periodic and open boundary conditions provided revealing information about its functional capability.

Furthermore, I would like to thank my colleagues from different work areas for the friendly working atmosphere and the valuable exchange as well as support concerning technical or organizational issues. I thank Dr. Arne-Jens Hempel for reading parts of the thesis and giving valuable hints concerning language.

This work was partly supported by the German Research Foundation (DFG), project PO 711/12-1.

2 Electrostatic interactions in particle systems

The efficient computation of electrostatic interactions in particle systems in three dimensional space is required in the field of molecular dynamics simulations, for instance. Typically, the interactions decrease with growing distance r between two particles with rate r^{-n}, $n \in \mathbb{N}$. In case of Coulombic and dipolar interactions we have $n \in \{1, 3, 5\}$, i.e., the convergence is rather slow. Thus, these types of interactions are widely known as long range interactions.

In many applications one additionally assumes periodic boundary conditions in order to avoid undesirable boundary effects, which would especially dominate in case of small system sizes. On the other hand, in certain applications the focus is set on the investigation of such boundary effects. In order to simulate thin liquid films [KS05] or single stiff polymers [DAH03], for instance, periodicity is only reasonable in two or one of three directions, respectively. However, a naive truncation scheme is not applicable due to the very slow convergence and one is highly interested in more efficient techniques for the computation of the required interaction quantities such as potentials, forces, energies and torques, for instance.

In the following we introduce some basic definitions concerning electrostatic interactions in particle systems consisting of charges and dipoles, see Section 2.1. Next, we consider the so called Ewald summation approach [Ewa21], which is a common tool for the development of efficient algorithms for molecular dynamics simulations. Various types of periodic boundary conditions are taken into account, see Section 2.2. The approximation errors resulting from truncating the underlying infinite sums are studied in Section 2.3, where we restrict our considerations to the 3d-periodic setting.

2.1 Preliminaries and Definitions

We consider electrostatic systems consisting of N_c point charges as well as N_d point dipoles, which we suppose to be located in a primary box spanned by three linearly independent vectors $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3 \in \mathbb{R}^3$.

We denote by

$$\boldsymbol{x}_j \in \mathbb{R}^3 \text{ and } q_j \in \mathbb{R}, j = 1, \ldots, N_\mathrm{c},$$

the positions of the charges as well as the corresponding charge values and by

$$\boldsymbol{x}_j \in \mathbb{R}^3 \text{ and } \boldsymbol{\mu}_j \in \mathbb{R}^3, j = N_\mathrm{c} + 1, \ldots, N_\mathrm{c} + N_\mathrm{d},$$

the positions of the dipoles as well as the corresponding dipole moments. Furthermore, we refer to

$$N := N_\mathrm{c} + N_\mathrm{d}$$

as the total number of present particles. That the particles are located in a box spanned by the three vectors $\boldsymbol{\ell}_1$, $\boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$ can be expressed as follows.

$$\boldsymbol{x}_j = \lambda_{j,1}\boldsymbol{\ell}_1 + \lambda_{j,2}\boldsymbol{\ell}_2 + \lambda_{j,3}\boldsymbol{\ell}_3 + \boldsymbol{\ell}_0 \quad \text{with} \quad \lambda_{j,1\ldots3} \in [-1/2, 1/2) \text{ and } \boldsymbol{\ell}_0 \in \mathbb{R}^3 = \text{const.} \tag{2.1}$$

Note that the interval $[-1/2, 1/2)$ may be replaced by an arbitrary interval of the same length, as for example $[0, 1)$. The upper interval limit is excluded since periodic boundary conditions may be

applied. The offset of the box is denoted by $\boldsymbol{\ell}_0$. Using the notation (2.1), the box has its center in the point of origin for $\boldsymbol{\ell}_0 = (0,0,0)^\top$.

If periodic boundary conditions are applied in some of the three dimensions, the overall electrostatic energy of the particle system can be written as

$$U := \frac{1}{2}\sum_{j=1}^{N} \xi_j \phi(j) \tag{2.2}$$

with the potentials

$$\phi(j) := \sum_{\boldsymbol{n}\in\mathcal{S}}{}' \sum_{i=1}^{N} \frac{\xi_i}{\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|}. \tag{2.3}$$

Here, the index set $\mathcal{S} \subset \mathbb{Z}^3$ is representing the applied periodic boundary conditions, $\|\cdot\|$ denotes the Euclidean norm and the vectors $\boldsymbol{x}_{ij}$ are defined by $\boldsymbol{x}_{ij} := \boldsymbol{x}_i - \boldsymbol{x}_j$. Furthermore, the prime on the infinite sum in (2.3) indicates that for $\boldsymbol{n} = \boldsymbol{0}$ all terms with $i = j$ are omitted, since a particle is not allowed to interact with itself. The regular matrix $\boldsymbol{L}$ is given by $\boldsymbol{L} := [\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3]$ and the operators ξ_j are defined via

$$\xi_j := \begin{cases} q_j & : j \in \{1,\dots,N_\mathrm{c}\}, \\ \boldsymbol{\mu}_j^\top \nabla_{\boldsymbol{x}_j} & : j \in \{N_\mathrm{c}+1,\dots,N\}. \end{cases} \tag{2.4}$$

Note that the electrostatic energy (2.2), the potentials (2.3) as well as the quantities introduced in the following are defined based on Gaussian units. Working with the standard international system of units introduces the prefactor $(4\pi\epsilon_0)^{-1}$, known as Coulomb's constant, where ϵ_0 denotes the vacuum permittivity.

If we assume periodic boundary conditions in the first $\rho \in \{0,1,2,3\}$ directions and open boundary conditions regarding the last $3-\rho$, we choose

$$\mathcal{S} := \mathbb{Z}^\rho \times \{0\}^{3-\rho}.$$

For a graphical illustration see Figures 2.1 and 2.2.

By making use of the one-dimensional torus

$$\mathbb{T} := \mathbb{R}/\mathbb{Z} \simeq [-1/2, 1/2),$$

and the well known Minkowski sum, the positions of the particles may also be regarded as elements of the space

$$\bigoplus_{j=1}^{\rho} \mathbb{T}\boldsymbol{\ell}_j + \bigoplus_{j=\rho+1}^{3} \mathbb{R}\boldsymbol{\ell}_j$$

in order to express the periodic boundary conditions with respect to the first $\rho \in \{1,2\}$ directions and open constraints concerning the last $3-\rho$. As usual, we identify the corresponding equivalence classes $[x]_\sim := \{x+n, n \in \mathbb{Z}\}$ with a single representative and write $x \in \mathbb{T}$ instead of $[x]_\sim \in \mathbb{T}$. For $\rho = 3$, i.e., for fully periodic constraints, we shortly write

$$\mathbb{T}\boldsymbol{\ell}_1 + \mathbb{T}\boldsymbol{\ell}_2 + \mathbb{T}\boldsymbol{\ell}_3 =: \boldsymbol{L}\mathbb{T}^3.$$

Note that the expressions for the energy (2.2) and the potentials (2.3) are the same as for charge-charge systems [Ewa21, dLPS80], where the charges q_j have simply been replaced by the operators ξ_j, as defined via (2.4).

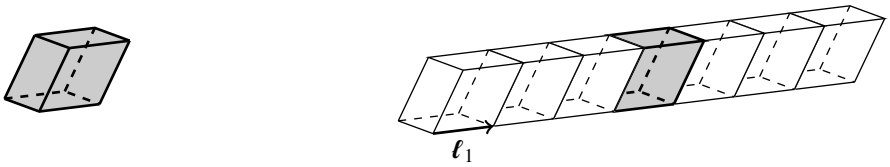


Figure 2.1: For open boundary conditions the particles are distributed in a finite box in $\mathbb{R}^3$ (left). In the 1d-periodic setting the box is duplicated with respect to the vector $\boldsymbol{\ell}_1$ (right).

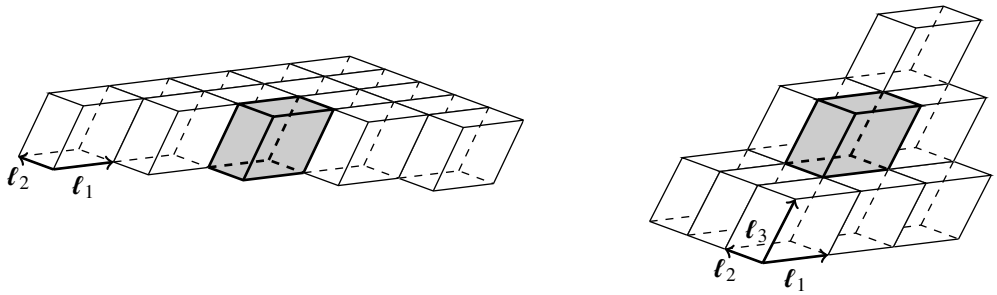


Figure 2.2: For 2d-periodic boundary conditions the primary box is duplicated along two directions (left). In the 3d-periodic case the box is duplicated with respect to all three box vectors $\boldsymbol{\ell}_1$, $\boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$ (right).

Remark 2.1 (Convergence)**.** In this remark we assume $\rho \in \{1, 2, 3\}$, i.e., we exclude the case of open boundary conditions. If we assume charge neutrality, i.e,

$$\sum_{j=1}^{N_c} q_j = 0 \tag{2.5}$$

it is easy to show that the single summands in (2.3) tend to zero like $O(\|\boldsymbol{Ln}\|^{-3})$.

Therefor, we compute the Taylor expansions

$$\frac{1}{\|\boldsymbol{x}+\boldsymbol{Ln}\|} = \frac{1}{\|\boldsymbol{Ln}\|} - \frac{\boldsymbol{x}^\top \boldsymbol{Ln}}{\|\boldsymbol{Ln}\|^3} + \frac{3(\boldsymbol{x}^\top \boldsymbol{Ln})^2}{2\|\boldsymbol{Ln}\|^5} - \frac{\boldsymbol{x}^\top \boldsymbol{x}}{2\|\boldsymbol{Ln}\|^3} + O(\|\boldsymbol{Ln}\|^{-4}), \tag{2.6}$$

$$\frac{\boldsymbol{\mu}^\top \nabla_{\boldsymbol{x}}}{\|\boldsymbol{x}+\boldsymbol{Ln}\|} = -\frac{\boldsymbol{\mu}^\top \boldsymbol{Ln}}{\|\boldsymbol{Ln}\|^3} + \frac{3(\boldsymbol{x}^\top \boldsymbol{Ln})(\boldsymbol{\mu}^\top \boldsymbol{Ln})}{\|\boldsymbol{Ln}\|^5} - \frac{\boldsymbol{\mu}^\top \boldsymbol{x}}{\|\boldsymbol{Ln}\|^3} + O(\|\boldsymbol{Ln}\|^{-4}), \tag{2.7}$$

and proceed as follows. Since the sum of all charges is zero we can subtract a constant in each summand, namely $\|\boldsymbol{Ln}\|^{-1}$. This idea has already been applied in the literature, see [Arn01, AH02]. We obtain

$$\begin{aligned}
&\sum_{\boldsymbol{n}\in\mathcal{S}\setminus\{\boldsymbol{0}\}} \left(\sum_{i=1}^{N_c} \frac{q_i}{\|\boldsymbol{x}_{ij}+\boldsymbol{Ln}\|} + \sum_{i=N_c+1}^{N} \frac{\boldsymbol{\mu}_i^\top \nabla_{\boldsymbol{x}_i}}{\|\boldsymbol{x}_{ij}+\boldsymbol{Ln}\|} \right) \\
&= \frac{1}{2} \sum_{\boldsymbol{n}\in\mathcal{S}\setminus\{\boldsymbol{0}\}} \sum_{i=1}^{N_c} \left(\frac{q_i}{\|\boldsymbol{x}_{ij}+\boldsymbol{Ln}\|} + \frac{q_i}{\|\boldsymbol{x}_{ij}-\boldsymbol{Ln}\|} - \frac{2q_i}{\|\boldsymbol{Ln}\|} \right) \\
&\quad + \frac{1}{2} \sum_{\boldsymbol{n}\in\mathcal{S}\setminus\{\boldsymbol{0}\}} \sum_{i=N_c+1}^{N} \left(\frac{\boldsymbol{\mu}_i^\top \nabla_{\boldsymbol{x}_i}}{\|\boldsymbol{x}_{ij}+\boldsymbol{Ln}\|} + \frac{\boldsymbol{\mu}_i^\top \nabla_{\boldsymbol{x}_i}}{\|\boldsymbol{x}_{ij}-\boldsymbol{Ln}\|} \right),
\end{aligned}$$

where we insert the Taylor expansions as introduced above to get

$$\begin{aligned}&\frac{q_i}{\|\boldsymbol{x}_{ij}+\boldsymbol{Ln}\|}+\frac{q_i}{\|\boldsymbol{x}_{ij}-\boldsymbol{Ln}\|}-\frac{2q_i}{\|\boldsymbol{Ln}\|}\\&=\frac{2q_i}{\|\boldsymbol{Ln}\|}-\frac{\boldsymbol{x}_{ij}^\top\boldsymbol{Ln}}{\|\boldsymbol{Ln}\|^3}-\frac{\boldsymbol{x}_{ij}^\top(-\boldsymbol{Ln})}{\|\boldsymbol{Ln}\|^3}-\frac{2q_i}{\|\boldsymbol{Ln}\|}+O(\|\boldsymbol{Ln}\|^{-3})=O(\|\boldsymbol{Ln}\|^{-3})\end{aligned}\tag{2.8}$$

and

$$\frac{\boldsymbol{\mu}_i^\top\nabla_{\boldsymbol{x}_i}}{\|\boldsymbol{x}_{ij}+\boldsymbol{Ln}\|}+\frac{\boldsymbol{\mu}_i^\top\nabla_{\boldsymbol{x}_i}}{\|\boldsymbol{x}_{ij}-\boldsymbol{Ln}\|}=-\frac{\boldsymbol{\mu}_i^\top\boldsymbol{Ln}}{\|\boldsymbol{Ln}\|^3}-\frac{\boldsymbol{\mu}_i^\top(-\boldsymbol{Ln})}{\|\boldsymbol{Ln}\|^3}+O(\|\boldsymbol{Ln}\|^{-3})=O(\|\boldsymbol{Ln}\|^{-3}).\tag{2.9}$$

In order to calculate the lattice sum in (2.3), a certain order of how the periodic duplicates of the primary box should be added has to be specified.

Generally, a spherical order of summation is applied, i.e., the values of the potentials are defined via the limit

$$\phi(j):=\lim_{C\to\infty}{\sum_{\boldsymbol{n}\in\mathcal{S}:\|\boldsymbol{Ln}\|\le C}}'\sum_{i=1}^{N}\frac{\xi_i}{\|\boldsymbol{x}_{ij}+\boldsymbol{Ln}\|},$$

see [dLPS80] for instance. Note that only entire boxes are added by increasing norm of the lattice vectors $\boldsymbol{Ln}$. Furthermore, we have

$${\sum_{\boldsymbol{n}\in\mathcal{S}:\|\boldsymbol{Ln}\|\le C}}'\sum_{i=1}^{N}\frac{\xi_i}{\|\boldsymbol{x}_{ij}+\boldsymbol{Ln}\|}=\sum_{\substack{i=1\\i\neq j}}^{N}\frac{\xi_i}{\|\boldsymbol{x}_{ij}\|}+\frac{1}{2}\sum_{\substack{\boldsymbol{n}\in\mathcal{S}\\0<\|\boldsymbol{Ln}\|\le C}}\left(\sum_{i=1}^{N}\frac{\xi_i}{\|\boldsymbol{x}_{ij}\pm\boldsymbol{Ln}\|}-\frac{2}{\|\boldsymbol{Ln}\|}\sum_{i=1}^{N_\mathrm{c}}q_i\right),$$

so that the single summands can again be considered to tend to zero like $O(\|\boldsymbol{Ln}\|^{-3})$.

This shows absolute convergence for 2d-periodic as well as for 1d-periodic boundary conditions, since

$$\sum_{\boldsymbol{n}\in\mathbb{Z}^2\setminus\{\boldsymbol{0}\}}\frac{1}{\|\boldsymbol{n}\|^3}<\infty\quad\text{and}\quad\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{1}{|n|^3}<\infty.$$

The absolute convergence applies to all orders of summation, were only entire boxes are added successively.

Of course, this argumentation is not sufficient in case of 3d-periodic boundary conditions. In order to prove the convergence for fully periodic constraints we have to show that for fixed $\boldsymbol{x},\boldsymbol{\mu}\in\mathbb{R}^3$ the limits

$$\lim_{C\to\infty}\sum_{\substack{\boldsymbol{n}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}\\\|\boldsymbol{n}\|\le C}}\left(\frac{3(\boldsymbol{x}^\top\boldsymbol{n})^2}{\|\boldsymbol{n}\|^5}-\frac{\boldsymbol{x}^\top\boldsymbol{x}}{\|\boldsymbol{n}\|^3}\right)\tag{2.10}$$

and

$$\lim_{C\to\infty}\sum_{\substack{\boldsymbol{n}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}\\\|\boldsymbol{n}\|\le C}}\left(\frac{3(\boldsymbol{x}^\top\boldsymbol{n})(\boldsymbol{\mu}^\top\boldsymbol{n})}{\|\boldsymbol{n}\|^5}-\frac{\boldsymbol{\mu}^\top\boldsymbol{x}}{\|\boldsymbol{n}\|^3}\right)\tag{2.11}$$

exist. The first terms of the Taylor expansions (2.6) and (2.7) cancel out due to the underlying symmetry as well as the charge neutrality, see (2.8) and (2.9). The remaining parts tend to zero like $O(\|\boldsymbol{n}\|^{-4})$, which is sufficient for absolute convergence in three dimensions. Note that we for simplicity omit the scaling by $\boldsymbol{L}$ and consider box vectors $\boldsymbol{n}$.

We consider the limit (2.10) and compute

$$\frac{3(\boldsymbol{x}^\top \boldsymbol{n})^2}{\|\boldsymbol{n}\|^5} - \frac{\boldsymbol{x}^\top \boldsymbol{x}}{\|\boldsymbol{n}\|^3} = \|\boldsymbol{x}\|^2 \left(\frac{3\cos^2 \theta_{\boldsymbol{x},\boldsymbol{n}}}{\|\boldsymbol{n}\|^3} - \frac{1}{\|\boldsymbol{n}\|^3} \right),$$

where we denote by $\theta_{\boldsymbol{x},\boldsymbol{n}}$ the angle between the vectors $\boldsymbol{x}$ and $\boldsymbol{n}$. For two positive numbers $R_1 < R_2$ we get

$$\int_{R_1 \le \|\boldsymbol{r}\| \le R_2} \frac{3\cos^2 \theta_{\boldsymbol{x},\boldsymbol{r}} - 1}{\|\boldsymbol{r}\|^3} \mathrm{d}\boldsymbol{r} = \int_{R_1}^{R_2} \int_0^{\pi} \int_0^{2\pi} \frac{3\cos^2 \theta - 1}{r} \sin\theta \, \mathrm{d}\phi \, \mathrm{d}\theta \, \mathrm{d}r = 0,$$

where the spherical coordinates are chosen such that θ describes the angle between the fixed vector $\boldsymbol{x} \in \mathbb{R}^3$ and $\boldsymbol{r} \in \mathbb{R}^3$.

In the following we set $R_2 := 1/2$ and $R_1 := a$, where $a \in (0, 1/2)$. For $L \in \mathbb{N}$ we subdivide the cube $[-1/2, 1/2)^3$ into small cubes $B_{\boldsymbol{n},L}$ of size $(2L+1)^{-3}$, i.e.

$$[-1/2, 1/2)^3 = \bigcup_{n_1,n_2,n_3=-L}^{L} B_{\boldsymbol{n},L}$$

with $\boldsymbol{n} = (n_1, n_2, n_3) \in \mathbb{Z}^3$ and

$$B_{\boldsymbol{n},L} := \left[\tfrac{n_1}{2L+1} - \tfrac{1}{4L+2}, \tfrac{n_1}{2L+1} + \tfrac{1}{4L+2} \right) \times \cdots \times \left[\tfrac{n_3}{2L+1} - \tfrac{1}{4L+2}, \tfrac{n_3}{2L+1} + \tfrac{1}{4L+2} \right).$$

Since the integrated function is continuous and smooth, the considered integral is well approximated by a simple quadrature rule. For this purpose, the integration domain may be approached by little cubes. Note that this approach is based on the Jordan measure of the integration domain, which is simply a spherical shell and therefore Jordan measurable. Thus, for increasing number of summands L the sum will converge to the real value of the integral, i.e.,

$$\begin{aligned} 0 = \int_{a \le \|\boldsymbol{r}\| \le 1/2} \frac{3\cos^2 \theta_{\boldsymbol{x},\boldsymbol{r}} - 1}{\|\boldsymbol{r}\|^3} \mathrm{d}\boldsymbol{r} &= \lim_{L\to\infty} \frac{1}{(2L+1)^3} \sum_{a \le (2L+1)^{-1}\|\boldsymbol{n}\| \le 1/2} \frac{3\cos^2 \theta_{\boldsymbol{x},(2L+1)^{-1}\boldsymbol{n}} - 1}{(2L+1)^{-3}\|\boldsymbol{n}\|^3} \\ &= \lim_{L\to\infty} \sum_{(2L+1)a \le \|\boldsymbol{n}\| \le \frac{2L+1}{2}} \frac{3\cos^2 \theta_{\boldsymbol{x},\boldsymbol{n}} - 1}{\|\boldsymbol{n}\|^3}. \end{aligned}$$

Since $a \in (0, 1/2)$ is arbitrary, we can write equivalently

$$\forall \epsilon > 0 \; \exists R(\epsilon) : \left| \sum_{R_1 \le \|\boldsymbol{n}\| \le R_2} \frac{3\cos^2 \theta_{\boldsymbol{x},\boldsymbol{n}} - 1}{\|\boldsymbol{n}\|^3} \right| < \epsilon \quad \forall \, R_2 > R_1 > R(\epsilon),$$

which means that the spherical limit (2.10) exists.

In order to show the existence of the limit (2.11) we proceed similarly. The spherical coordinates $\boldsymbol{n} \mapsto \|\boldsymbol{n}\|(\sin\theta_{\boldsymbol{n}} \sin\phi_{\boldsymbol{n}}, \sin\theta_{\boldsymbol{n}} \cos\phi_{\boldsymbol{n}}, \cos\theta_{\boldsymbol{n}})^\top$ are chosen such that $\theta_{\boldsymbol{n}}$ denotes the angle between $\boldsymbol{x}$ and $\boldsymbol{n}$, i.e., $\boldsymbol{x} \mapsto (0, 0, \|\boldsymbol{x}\|)^\top$, and $\phi_{\boldsymbol{n}} = 0$ for all $\boldsymbol{n} = c \cdot \boldsymbol{\mu}$ with $c > 0$, i.e., $\boldsymbol{\mu} \mapsto \|\boldsymbol{\mu}\|(\sin\theta_{\boldsymbol{\mu}}, 0, \cos\theta_{\boldsymbol{\mu}})^\top$. Thus, we obtain

$$\frac{3(\boldsymbol{x}^\top \boldsymbol{n})(\boldsymbol{\mu}^\top \boldsymbol{n})}{\|\boldsymbol{n}\|^5} - \frac{\boldsymbol{\mu}^\top \boldsymbol{x}}{\|\boldsymbol{n}\|^3} = \frac{\|\boldsymbol{\mu}\|\|\boldsymbol{x}\|}{\|\boldsymbol{n}\|^3} \left(\cos\theta_{\boldsymbol{\mu}} (3\cos^2\theta_{\boldsymbol{n}} - 1) + \sin\theta_{\boldsymbol{\mu}} \cos\theta_{\boldsymbol{n}} \sin\theta_{\boldsymbol{n}} \sin\phi_{\boldsymbol{n}} \right).$$

Since the two associated integrals

$$\cos\theta_{\boldsymbol{\mu}} \int_{R_1}^{R_2} \int_0^{\pi} \int_0^{2\pi} \frac{3\cos^2\theta - 1}{r} \sin\theta \,\mathrm{d}r \,\mathrm{d}\theta \,\mathrm{d}\phi = 0$$

and

$$\sin\theta_{\boldsymbol{\mu}} \int_{R_1}^{R_2} \int_0^{\pi} \int_0^{2\pi} \frac{\cos\theta \sin^2\theta \sin\phi}{r} \,\mathrm{d}r \,\mathrm{d}\theta \,\mathrm{d}\phi = 0$$

are equal to zero, the convergence of the spherical limit (2.11) is verified in an analogue manner. □

We are also interested in the computation of the fields $\boldsymbol{E}(j)$, energies $U(j)$ and forces $\boldsymbol{F}(j)$ of the single particles, which are given by

$$\boldsymbol{E}(j) := -\nabla_{\boldsymbol{x}_j} \phi(j), \tag{2.12}$$

$$U(j) := \xi_j \phi(j) = \begin{cases} q_j \phi(j) & : j \in \{1, \dots, N_\mathrm{c}\}, \\ -\boldsymbol{\mu}_j^\top \boldsymbol{E}(j) & : j \in \{N_\mathrm{c}+1, \dots, N\}, \end{cases} \tag{2.13}$$

and

$$\boldsymbol{F}(j) := -\nabla_{\boldsymbol{x}_j} U(j) = \begin{cases} q_j \boldsymbol{E}(j) & : j \in \{1, \dots, N_\mathrm{c}\}, \\ \boldsymbol{G}(j)\boldsymbol{\mu}_j & : j \in \{N_\mathrm{c}+1, \dots, N\}, \end{cases} \tag{2.14}$$

where $\boldsymbol{G}(j)$ is the negative electric field gradient, i.e.,

$$\boldsymbol{G}(j) := -\nabla_{\boldsymbol{x}_j} \nabla_{\boldsymbol{x}_j}^\top \phi(j).$$

For the dipoles, i.e., for $j = N_\mathrm{c}+1, \dots, N$, the torques are given by

$$\boldsymbol{\tau}(j) := \boldsymbol{\mu}_j \times \boldsymbol{E}(j). \tag{2.15}$$

Note that we can further simplify the expression of the overall energy (2.2) via

$$U = \frac{1}{2} \sum_{j=1}^{N_\mathrm{c}} q_j \phi(j) - \frac{1}{2} \sum_{j=N_\mathrm{c}+1}^{N} \boldsymbol{\mu}_j^\top \boldsymbol{E}(j), \tag{2.16}$$

i.e., the overall energy can be obtained easily after having computed the potentials of the charges as well as the fields of the dipoles.

As already discussed before, the considered interactions are long ranged so that an evaluation of the underlying lattice sums via truncation is not practicable. Thus, more advanced methods are required in order to compute the electrostatic quantities of interest in an efficient fashion. A large number of existing particle-mesh methods are based on the so called Ewald summation approach [Ewa21]. In the following we consider particle systems containing both point charges as well as point dipoles, and revisit the Ewald summation approach for different types of periodic as well as open boundary conditions.

2.2 Ewald summation for electrostatic systems with charges and dipoles

The basic idea of the Ewald summation technique is explained as follows. The long ranged interaction kernel r^{-1} is split into two parts via

$$\frac{1}{r} = \frac{\operatorname{erfc}(\alpha r)}{r} + \frac{\operatorname{erf}(\alpha r)}{r}, \tag{2.17}$$

where $\alpha > 0$ is called Ewald or splitting parameter,

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-z^2} \mathrm{d}z \tag{2.18}$$

is the well known error function and $\operatorname{erfc}(x) := 1 - \operatorname{erf}(x)$ is the complementary error function. Based on this splitting, we will end up with a so called short range part converging absolutely in real space as well as a so called long range part converging exponentially fast in Fourier space.

Note that the singularity at $r = 0$ is passed along the first part in (2.17), whereas the second term has a finite limit for $r \to 0$, which is given by

$$\lim_{r \to 0} \frac{\operatorname{erf}(\alpha r)}{r} = \frac{2\alpha}{\sqrt{\pi}}. \tag{2.19}$$

In other words, the second kernel function is continuous, but still converging to zero very slowly as $r \to \infty$. The complementary error function, in contrast, tends to zero exponentially fast with growing distance r. More precisely, for $x \neq 0$ we have

$$\begin{aligned} \operatorname{erfc}(x) &= \frac{2}{\sqrt{\pi}} \int_x^\infty \mathrm{e}^{-z^2} \mathrm{d}z = \frac{2}{\sqrt{\pi}} \int_x^\infty z\, \mathrm{e}^{-z^2} \cdot \frac{1}{z} \mathrm{d}z \\ &= \frac{2}{\sqrt{\pi}} \left(\frac{\mathrm{e}^{-x^2}}{2x} - \int_x^\infty \frac{\mathrm{e}^{-z^2}}{2z^2} \mathrm{d}z \right) < \frac{\mathrm{e}^{-x^2}}{\sqrt{\pi} x}. \end{aligned} \tag{2.20}$$

If we continue with applying the partial integration repeatedly, we obtain the well known asymptotic expansion formula for the complementary error function [AS72, Eq. 7.1.23]

$$\operatorname{erfc}(x) = \frac{\mathrm{e}^{-x^2}}{x\sqrt{\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{(2n-1)!!}{2^n x^{2n}} \approx \frac{\mathrm{e}^{-x^2}}{x\sqrt{\pi}} \quad \text{for } x \text{ large enough.} \tag{2.21}$$

If we apply the splitting in (2.3) by setting $r := \|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|$ in each summand, the energy U is split into an absolutely converging short range part

$$U^{\text{short}} := \frac{1}{2} \sum_{j=1}^{N} \xi_j \phi^{\text{short}}(j) \quad \text{with} \quad \phi^{\text{short}}(j) := \sum_{\boldsymbol{n} \in \mathcal{S}}{}' \sum_{i=1}^{N} \xi_i \frac{\operatorname{erfc}(\alpha \|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|)}{\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|} \tag{2.22}$$

as well as a long range part

$$U^{\text{long}} := \frac{1}{2} \sum_{j=1}^{N} \xi_j \phi^{\text{long}}(j) \quad \text{with} \quad \phi^{\text{long}}(j) := \sum_{\boldsymbol{n} \in \mathcal{S}} \sum_{i=1}^{N} \xi_i \frac{\operatorname{erf}(\alpha \|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|)}{\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|} - \phi^{\text{self}}(j). \tag{2.23}$$

Since the kernel function $r^{-1}\mathrm{erfc}(\alpha r)$ in the short range part tends to zero exponentially fast, the short range potentials can be computed efficiently by truncating the infinite sum, i.e., only distances $\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|$ smaller than some appropriate cutoff radius r_{cut} may be considered. Of course, also the splitting parameter α has to be chosen appropriately.

Since the long ranged kernel function $r^{-1}\mathrm{erf}(\alpha r)$ is continuous in $r = 0$, we can include all terms with $\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\| = 0$, i.e., the prime on the first sum in (2.23) is omitted. The thereby included self potential, which we denote by $\phi^{\mathrm{self}}(j)$, is simply subtracted afterward. The self potentials adopt the same values for all types of periodic boundary conditions and are given by

$$\phi^{\mathrm{self}}(j) = \begin{cases} \frac{2\alpha}{\sqrt{\pi}} q_j & : j = 1, \ldots, N_{\mathrm{c}}, \\ 0 & : j = N_{\mathrm{c}} + 1, \ldots, N, \end{cases} \tag{2.24}$$

which is obtained by computing the limit for $\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\| \to 0$, see (2.19). That the self potentials are equal to zero for the dipole particles follows from the underlying symmetry.

A periodic and continuous function can be represented by a Fourier series. Thus, it seems reasonable to transform the long range part, see (2.23), into Fourier space, which has to be done according to the given boundary conditions.

In general, the Poisson summation formula is needed in order to obtain the Fourier space representations of the long range parts. The Poisson summation formula is stated in the following lemma.

Lemma 2.2. *We consider the periodization*

$$\sum_{\boldsymbol{n} \in \mathbb{Z}^d} f(\boldsymbol{x} + \boldsymbol{n})$$

of a function f, which converges almost everywhere if $f \in \mathcal{L}_1(\mathbb{R})$. If f is also continuous, then its periodization converges to a continuous, 1-periodic function f_1. If, in addition, the Fourier series of this function f_1 converges everywhere to f_1, we have, cf. [Zyg93, *Theorem* 13.5]*,*

$$f_1(\boldsymbol{x}) := \sum_{\boldsymbol{n} \in \mathbb{Z}^d} f(\boldsymbol{x} + \boldsymbol{n}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} \hat{f}(\boldsymbol{k})\, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{x}}, \quad \forall\, \boldsymbol{x} \in \mathbb{R}^d, \tag{2.25}$$

where we denote by

$$\hat{f}(\boldsymbol{\xi}) := \int_{\mathbb{R}^d} f(\boldsymbol{x})\ \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{\xi}^\top \boldsymbol{x}} \mathrm{d}\boldsymbol{x}$$

the Fourier transform of the function f. In other words, the $\boldsymbol{k}$-th Fourier coefficient of the periodic function f_1 coincides with $\hat{f}(\boldsymbol{k})$.

Corollary 2.3. *Let the regular matrix $\boldsymbol{A} \in \mathbb{R}^{d \times d}$ be given and suppose that the function $f : \mathbb{R}^d \to \mathbb{R}$ satisfies the conditions of Lemma* 2.2*. Then, we have*

$$\sum_{\boldsymbol{n} \in \mathbb{Z}^d} f(\boldsymbol{x} + \boldsymbol{A}\boldsymbol{n}) = \frac{1}{|\det(\boldsymbol{A})|} \sum_{\boldsymbol{k} \in \mathbb{Z}^d} \hat{f}(\boldsymbol{A}^{-\top} \boldsymbol{k})\, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{A}^{-1} \boldsymbol{x}}. \tag{2.26}$$

Proof. We define the function $g(\boldsymbol{x}) := f(\boldsymbol{A}\boldsymbol{x})$ and obtain

$$\hat{g}(\boldsymbol{\xi}) = \int_{\mathbb{R}^d} f(\boldsymbol{A}\boldsymbol{x})\, \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{\xi}^\top \boldsymbol{A}^{-1} \boldsymbol{A}\boldsymbol{x}} \mathrm{d}\boldsymbol{x} = \frac{1}{|\det(\boldsymbol{A})|} \hat{f}(\boldsymbol{A}^{-\top} \boldsymbol{\xi}),$$

where we simply substitute $\boldsymbol{A}\boldsymbol{x}$ in the integral. Furthermore, we have

$$\sum_{\boldsymbol{n} \in \mathbb{Z}^d} f(\boldsymbol{x} + \boldsymbol{A}\boldsymbol{n}) = \sum_{\boldsymbol{n} \in \mathbb{Z}^d} g(\boldsymbol{A}^{-1}\boldsymbol{x} + \boldsymbol{n}),$$

and the application of the Poisson summation formula (2.25) gives (2.26). ∎

Example 2.4. For a Gaussian of the form

$$f(\boldsymbol{x}) = \mathrm{e}^{-\beta\|\boldsymbol{x}\|^2}$$

with $\boldsymbol{x} \in \mathbb{R}^d$ and $\beta > 0$ we have

$$\hat{f}(\boldsymbol{\xi}) = \left(\frac{\pi}{\beta}\right)^{d/2} \mathrm{e}^{-\pi^2\|\boldsymbol{\xi}\|^2/\beta}, \tag{2.27}$$

as is well known. By applying (2.26) we obtain

$$\sum_{\boldsymbol{n}\in\mathbb{Z}^d} \mathrm{e}^{-\beta\|\boldsymbol{x}+\boldsymbol{A}\boldsymbol{n}\|^2} = \frac{1}{|\det(\boldsymbol{A})|}\left(\frac{\pi}{\beta}\right)^{d/2} \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \mathrm{e}^{-\pi^2\|\boldsymbol{A}^{-\top}\boldsymbol{k}\|^2/\beta}\, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{A}^{-1}\boldsymbol{x}}. \tag{2.28}$$

□

In the following sections we apply the Poisson summation formula for the periodized Gaussian (2.28) in order to compute a Fourier space representation of the long range potential, as given in (2.23). The easiest case is that periodic boundary conditions are applied in all three dimensions, where $\mathcal{S} = \mathbb{Z}^3$, which is going to be considered in Section 2.2.1. Here, the underlying infinite sum is transformed into Fourier space with respect to all three dimensions. In contrast to the original series in spatial domain, the Fourier series converges exponentially fast, which makes it possible to compute a good approximation via truncation.

For 2d- and 1d-periodic boundary conditions, where $\mathcal{S} = \mathbb{Z}^2 \times \{0\}$ and $\mathcal{S} = \mathbb{Z} \times \{0\}^2$, respectively, only the dimensions subject to periodic boundary conditions may be transformed into Fourier space analytically. The mixed periodic boundary conditions are considered in Sections 2.2.2 and 2.2.3. Analogously to the fully periodic case, the Fourier coefficients underlie an exponential decay. For open boundary conditions a transformation into Fourier space is not readily possible, see Section 2.2.4. Finally, we summarize the results in Section 2.2.5, where Theorem 2.16 presents the Ewald summation formulas in a universal notation, merging all considered types of periodic boundary conditions.

2.2.1 3d-periodic boundary conditions

In Lemma 2.5 we present the Fourier space representation of the long range potential for 3d-periodic boundary conditions [Ewa21], where we denote by

$$V := |\det(\boldsymbol{L})|$$

the volume of the primary simulation box.

The result is obtained based on the theory of convergence factors, as in introduced in [dLPS80]. For more information on the theory of summable series and convergence factors we refer to [Moo38].

Lemma 2.5. *For a charge neutral particle system, cf.* (2.5)*, we have*

$$\lim_{C\to\infty} \sum_{\boldsymbol{n}\in\mathbb{Z}^3:\|\boldsymbol{L}\boldsymbol{n}\|\le C} \sum_{i=1}^{N} \xi_i \frac{\operatorname{erf}(\alpha\|\boldsymbol{x}_{ij}+\boldsymbol{L}\boldsymbol{n}\|)}{\|\boldsymbol{x}_{ij}+\boldsymbol{L}\boldsymbol{n}\|} = \phi^{\mathrm{F}}(j) + \phi^{\boldsymbol{0}}(j), \tag{2.29}$$

where the Fourier space potentials $\phi^{\mathrm{F}}(j)$ are given by

$$\phi^{\mathrm{F}}(j) = \frac{1}{\pi V} \sum_{\boldsymbol{k}\in\mathbb{Z}^3} \hat{\psi}(\boldsymbol{k}) \sum_{i=1}^{N} \xi_i\, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}} \tag{2.30}$$

with the Fourier coefficients

$$\hat{\psi}(\boldsymbol{k}) := \begin{cases} \dfrac{\mathrm{e}^{-\pi^2\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2/\alpha^2}}{\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2} & : \boldsymbol{k} \neq \boldsymbol{0}, \\ 0 & : \boldsymbol{k} = \boldsymbol{0}. \end{cases} \tag{2.31}$$

The $\boldsymbol{k} = \boldsymbol{0}$ parts $\phi^{\boldsymbol{0}}(j)$ are given by

$$\phi^{\boldsymbol{0}}(j) = -\frac{2\pi}{3V} \sum_{i=1}^{N} \xi_i \|\boldsymbol{x}_{ij}\|^2. \tag{2.32}$$

Proof. In order to compute the limit of a conditionally convergent series, convergence factors can be applied, see [dLPS80]. The convergence factors have to be chosen according to the supposed order of summation. For the computation of the spherical limit, as defined above, the convergence factors $\mathrm{e}^{-s\|\boldsymbol{L}\boldsymbol{n}\|^2}$, where $s > 0$, are suitable.

Thus, we transform the absolutely convergent series,

$$\Phi_s(j) := \sum_{\boldsymbol{n}\in\mathbb{Z}^3} \sum_{i=1}^{N} \xi_i \, \mathrm{e}^{-s\|\boldsymbol{L}\boldsymbol{n}\|^2} \frac{\operatorname{erf}(\alpha\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|)}{\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|} \tag{2.33}$$

into Fourier space and then compute the limit for $s \to 0$, which is approached uniformly, cf. [dLPS80, Section 2]. First, we make use of the definition of the error function (2.18) and obtain

$$\Phi_s(j) = \frac{2}{\sqrt{\pi}} \sum_{\boldsymbol{n}\in\mathbb{Z}^3} \sum_{i=1}^{N} \xi_i \int_0^{\alpha} \mathrm{e}^{-\|\boldsymbol{x}_{ij}+\boldsymbol{L}\boldsymbol{n}\|^2 z^2 - s\|\boldsymbol{L}\boldsymbol{n}\|^2} \mathrm{d}z.$$

By using the identity

$$z^2\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|^2 + s\|\boldsymbol{L}\boldsymbol{n}\|^2 = \frac{z^4}{s+z^2} \left\| \frac{z^2+s}{z^2}\boldsymbol{L}\boldsymbol{n} + \boldsymbol{x}_{ij} \right\|^2 + \frac{z^2 s}{z^2+s}\|\boldsymbol{x}_{ij}\|^2,$$

cf. [dLPS80, Eq. 3.12], and making use of the Poisson summation formula for a Gaussian in $d = 3$ dimensions, see (2.28), we end up with

$$\begin{aligned} \Phi_s(j) &= \frac{2}{\sqrt{\pi}} \int_0^{\alpha} \sum_{i=1}^{N} \xi_i \, \mathrm{e}^{-\frac{z^2 s}{z^2+s}\|\boldsymbol{x}_{ij}\|^2} \sum_{\boldsymbol{n}\in\mathbb{Z}^3} \mathrm{e}^{-\frac{z^4}{s+z^2}\left\|\frac{z^2+s}{z^2}\boldsymbol{L}\boldsymbol{n}+\boldsymbol{x}_{ij}\right\|^2} \mathrm{d}z \\ &= \frac{2}{\sqrt{\pi}} \int_0^{\alpha} \sum_{i=1}^{N} \xi_i \, \mathrm{e}^{-\frac{z^2 s}{z^2+s}\|\boldsymbol{x}_{ij}\|^2} \frac{\pi^{3/2}}{V(z^2+s)^{3/2}} \sum_{\boldsymbol{k}\in\mathbb{Z}^3} \mathrm{e}^{-\pi^2\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2/(z^2+s)} \, \mathrm{e}^{\frac{2\pi \mathrm{i} z^2}{z^2+s}\boldsymbol{k}^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}} \mathrm{d}z. \end{aligned}$$

Note that we simply changed the order of summation, which is allowed since the series converges absolutely for $s > 0$. For $\boldsymbol{k} \neq \boldsymbol{0}$ we may simply shift the limit for $s \to 0$ into the integrals. With

$$\int_0^{\alpha} \frac{\mathrm{e}^{-\pi^2\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2/z^2}}{z^3} \mathrm{d}z = \frac{\mathrm{e}^{-\pi^2\|\boldsymbol{x}_{ij}+\boldsymbol{L}\boldsymbol{n}\|^2/\alpha^2}}{2\pi^2\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2} = \frac{\hat{\psi}(\boldsymbol{k})}{2\pi^2} \tag{2.34}$$

we obtain

$$\lim_{s\to 0} \Phi_s(j) = \frac{1}{\pi V} \sum_{\boldsymbol{k}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}} \hat{\psi}(\boldsymbol{k}) \, \mathrm{e}^{2\pi \mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}} + \lim_{s\to 0} \frac{2\pi}{V} \sum_{i=1}^{N} \xi_i \int_0^{\alpha} \frac{\mathrm{e}^{-\frac{z^2 s}{z^2+s}\|\boldsymbol{x}_{ij}\|^2}}{(z^2+s)^{3/2}} \mathrm{d}z. \tag{2.35}$$

With the substitution $y = \frac{z^2 s}{z^2+s}$ and by replacing the exponential function by its Taylor series we finally get

$$\int_0^\alpha \frac{\mathrm{e}^{-\frac{z^2 s}{z^2+s}\|\boldsymbol{x}_{ij}\|^2}}{(z^2+s)^{3/2}}\mathrm{d}z = \int_0^{\frac{\alpha^2}{\alpha^2+s}} \frac{\mathrm{e}^{-ys\|\boldsymbol{x}_{ij}\|^2}}{2s\sqrt{y}}\mathrm{d}y = \frac{1}{s}\sqrt{\frac{\alpha^2}{\alpha^2+s}} - \frac{\|\boldsymbol{x}_{ij}\|^2}{3} + O(s).$$

The first part is not depending on the particle positions. Since the sum over all charges is zero and the application of the differential operators $\boldsymbol{\mu}_i^\top \nabla_{\boldsymbol{x}_i}$ gives also zero, we obtain

$$\lim_{s\to 0} \frac{2\pi}{V} \sum_{i=1}^N \xi_i \int_0^\alpha \frac{\mathrm{e}^{-\frac{z^2 s}{z^2+s}\|\boldsymbol{x}_{ij}\|^2}}{(z^2+s)^{3/2}}\mathrm{d}z = -\frac{2\pi}{3V}\sum_{i=1}^N \xi_i \|\boldsymbol{x}_{ij}\|^2.$$

Inserting this into (2.35) gives the final result $\lim_{s\to 0}\Phi_s(j) = \phi^{\mathrm{F}}(j) + \phi^{\mathbf{0}}(j)$. ■

Remark 2.6. We can see that also the Fourier space parts of the potentials $\phi^{\mathrm{F}}(j)$, as given in (2.30), converge exponentially fast, since the Fourier coefficients tend to zero like $O(k^{-2}\,\mathrm{e}^{-k^2})$ for $k := \|\boldsymbol{L}^{-\top}\boldsymbol{k}\| \to \infty$. Thus, an efficient evaluation of the Fourier space parts is possible by truncating the infinite sum, i.e, only vectors $\|\boldsymbol{L}^{-\top}\boldsymbol{k}\| \le k_{\mathrm{c}}$, where $k_{\mathrm{c}} > 0$ is an appropriate cutoff radius, may be considered. The tradeoff between computational effort needed for the short range part and the Fourier space part is regulated by the splitting parameter α.

For charge-charge systems it has been shown in [KP92, Appendix B] that an optimal choice of the splitting parameter results in an arithmetic complexity of $O(N^{3/2})$, where the constant depends on the required accuracy. □

Remark 2.7. The $\boldsymbol{k} = \mathbf{0}$ part $\phi^{\mathbf{0}}(j)$, see (2.32), is the only term, which depends on the applied summation order as well as the nature of the surrounding medium. The result of Lemma 2.5 represents the spherical summation order, as explained above, combined with vacuum boundary conditions. A more detailed discussion on the origin of this correction term can be found in [Bal14].

In general, the prefactor in (2.32) is

$$\frac{-2\pi}{3V(1+\epsilon)},$$

where we denote by ϵ the dielectric constant of the surrounding medium, cf. [FS02]. For vacuum boundary conditions, where $\epsilon = 0$, we obtain the representation as given in (2.32). In contrast, for metallic boundary conditions, where $\epsilon = \infty$, the $\boldsymbol{k} = \mathbf{0}$ term vanishes.

By utilizing the charge neutrality condition (2.5) we obtain

$$\begin{aligned}\phi^{\mathbf{0}}(j) &= -\frac{2\pi}{3V}\left(\sum_{i=1}^{N_{\mathrm{c}}} q_i \boldsymbol{x}_{ij}^\top \boldsymbol{x}_{ij} + \sum_{i=N_{\mathrm{c}}+1}^{N} \boldsymbol{\mu}_i^\top \nabla_{\boldsymbol{x}_i}\|\boldsymbol{x}_{ij}\|^2\right)\\ &= \frac{4\pi}{3V}\left(\boldsymbol{D}^\top \boldsymbol{x}_j - \frac{1}{2}\sum_{i=1}^{N_{\mathrm{c}}} q_i\|\boldsymbol{x}_i\|^2 - \sum_{i=N_{\mathrm{c}}+1}^{N}\boldsymbol{\mu}_i^\top \boldsymbol{x}_i\right),\end{aligned} \tag{2.36}$$

where we define the total dipole moment

$$\boldsymbol{D} := \sum_{i=1}^{N_{\mathrm{c}}} q_i \boldsymbol{x}_i + \sum_{i=N_{\mathrm{c}}+1}^{N} \boldsymbol{\mu}_i.$$

□

2.2.2 2d-periodic boundary conditions

We set $\mathcal{S} := \mathbb{Z}^2 \times \{0\}$ and denote by $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ the lattice vectors indicating the periodic directions, i.e., without loss of generality we assume the first two directions to be the periodic ones, see left hand side in Figure 2.2.

In the following lemma we present the Ewald formulas for the long range potentials under 2d-periodic boundary conditions, which has already been derived in [GGB00]. The authors consider both charge-charge as well as dipole-dipole interactions and restrict their considerations to cubic box structures, concerning the two periodic dimensions. The following proof, which is of a similar structure as the one presented in [GGB00, Nes12, NPP15a], applies to systems consisting of charges and dipoles and covers the case of a triclinic box structure with respect to the periodic dimensions. In order to derive the Ewald summation formulas, we make use of the theory of convergence factors, see [dLPS80]. The proof presented in [GGB00] follows a quite similar argumentation.

Note that Lemma 2.8 requires a special structure of the box vectors $\boldsymbol{\ell}_j$, but we will see later that each configuration can be traced back to the one considered here, see Lemma 2.10. We remark that the Ewald summation formulas for 2d-periodic boundary conditions are stated in a similar form in [Pip15, Section 4.4.3].

Lemma 2.8. *Suppose that the given particles are distributed in a primary simulation box, spanned by the three linearly independent vectors $\boldsymbol{\ell}_1$, $\boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$, cf.* (2.1)*. Assume that periodic boundary conditions are applied with respect to $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$, which are of the form*

$$\boldsymbol{\ell}_1 = \begin{pmatrix} l_{11} \\ l_{12} \\ 0 \end{pmatrix}, \boldsymbol{\ell}_2 = \begin{pmatrix} l_{21} \\ l_{22} \\ 0 \end{pmatrix}. \tag{2.37}$$

Let further the matrix $\tilde{\boldsymbol{L}} \in \mathbb{R}^{2\times 2}$ be defined by

$$\tilde{\boldsymbol{L}} := \begin{pmatrix} l_{11} & l_{21} \\ l_{12} & l_{22} \end{pmatrix}$$

and the particle positions be written in the form

$$\boldsymbol{x}_j = \begin{pmatrix} \tilde{\boldsymbol{x}}_j \\ x_{j,3} \end{pmatrix}.$$

Then, we have for a charge neutral system

$$\lim_{C\to\infty} \sum_{\boldsymbol{n}\in\mathbb{Z}^2\times\{0\}:\|\boldsymbol{L}\boldsymbol{n}\|\le C} \sum_{i=1}^{N} \xi_i \frac{\operatorname{erf}(\alpha\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|)}{\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|} = \phi^{\mathrm{F}}(j),$$

where the Fourier space representations $\phi^{\mathrm{F}}(j)$ of the long range potentials are given by

$$\phi^{\mathrm{F}}(j) = \frac{1}{\|\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2\|} \sum_{\boldsymbol{k}\in\mathbb{Z}^2} \sum_{i=1}^{N} \xi_i \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^\top \tilde{\boldsymbol{L}}^{-1} \tilde{\boldsymbol{x}}_{ij}} \Theta^{\mathrm{p2}}\left(\|\tilde{\boldsymbol{L}}^{-\top}\boldsymbol{k}\|, x_{ij,3}\right), \tag{2.38}$$

and the function Θ^{p2} is defined via

$$\Theta^{\mathrm{p2}}(k,r) := \begin{cases} \dfrac{1}{2k}\left[\mathrm{e}^{2\pi k r}\operatorname{erfc}\left(\dfrac{\pi k}{\alpha} + \alpha r\right) + \mathrm{e}^{-2\pi k r}\operatorname{erfc}\left(\dfrac{\pi k}{\alpha} - \alpha r\right)\right] & : k \neq 0, \\ -2\sqrt{\pi}\left(\dfrac{\mathrm{e}^{-\alpha^2 r^2}}{\alpha} + \sqrt{\pi} r \operatorname{erf}(\alpha r)\right) & : k = 0. \end{cases} \tag{2.39}$$

Proof. At first, we write $\boldsymbol{n} \in \mathbb{Z}^2 \times \{0\}$ in the form

$$\boldsymbol{n} = \begin{pmatrix} \tilde{\boldsymbol{n}} \\ 0 \end{pmatrix}$$

with $\tilde{\boldsymbol{n}} \in \mathbb{Z}^2$. We obtain

$$\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|^2 = \|\tilde{\boldsymbol{x}}_{ij} + \tilde{\boldsymbol{L}}\tilde{\boldsymbol{n}}\|^2 + x_{ij,3}^2.$$

Analogously to the 3d-periodic case, we make use of the convergence factors $\mathrm{e}^{-s\|\boldsymbol{L}\boldsymbol{n}\|^2}$ and compute $\lim_{s\to 0} \phi_s(j)$, which is defined analogously to 3d-periodic boundary conditions, see (2.33). The same convergence factor is applied in the proof presented in [GGB00]. We obtain

$$\begin{aligned}
\Phi_s(j) &= \frac{2}{\sqrt{\pi}} \int_0^\alpha \sum_{i=1}^N \xi_i \, \mathrm{e}^{-z^2 x_{ij,3}^2} \, \mathrm{e}^{-\frac{z^2 s}{z^2+s}\|\boldsymbol{x}_{ij}\|^2} \sum_{\tilde{\boldsymbol{n}}\in\mathbb{Z}^2} \mathrm{e}^{-\frac{z^4}{s+z^2}\left\|\frac{z^2+s}{z^2}\tilde{\boldsymbol{L}}\tilde{\boldsymbol{n}}+\tilde{\boldsymbol{x}}_{ij}\right\|^2} \mathrm{d}z \\
&= \frac{2}{\sqrt{\pi}} \int_0^\alpha \sum_{i=1}^N \xi_i \frac{\pi \, \mathrm{e}^{-z^2 x_{ij,3}^2} \, \mathrm{e}^{-\frac{z^2 s}{z^2+s}\|\boldsymbol{x}_{ij}\|^2}}{|\det(\tilde{\boldsymbol{L}})|(z^2+s)} \sum_{\boldsymbol{k}\in\mathbb{Z}^2} \mathrm{e}^{-\pi^2\|\tilde{\boldsymbol{L}}^{-\top}\boldsymbol{k}\|^2/(z^2+s)} \, \mathrm{e}^{\frac{2\pi i z^2}{z^2+s}\boldsymbol{k}^\top\tilde{\boldsymbol{L}}^{-1}\tilde{\boldsymbol{x}}_{ij}} \mathrm{d}z, \qquad (2.40)
\end{aligned}$$

where $|\det(\tilde{\boldsymbol{L}})| = \|\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2\|$. In order to obtain (2.40), we applied the Poisson summation formula (2.28) with $d = 2$. Again, we can simply shift the limit for $s \to 0$ into the integrals for $\boldsymbol{k} \neq \boldsymbol{0}$ and compute

$$\Theta^{\mathrm{p2}}(\|\tilde{\boldsymbol{L}}^{-\top}\boldsymbol{k}\|, x_{ij,3}) := 2\sqrt{\pi} \int_0^\alpha \frac{1}{z^2} \, \mathrm{e}^{-z^2 x_{ij,3}^2 - \frac{\pi^2\|\tilde{\boldsymbol{L}}^{-\top}\boldsymbol{k}\|^2}{z^2}} \mathrm{d}z = 2\sqrt{\pi} \int_{1/\alpha}^\infty \mathrm{e}^{-\frac{x_{ij,3}^2}{y^2} - \pi^2\|\tilde{\boldsymbol{L}}^{-\top}\boldsymbol{k}\|^2 y^2} \mathrm{d}y. \qquad (2.41)$$

By using

$$1 = \frac{1}{2\pi k}\left(\pi k - \frac{r}{y}\right) + \frac{1}{2\pi k}\left(\pi k + \frac{r}{y}\right)$$

one finds

$$\begin{aligned}
\int_{1/\alpha}^\infty \mathrm{e}^{-\pi^2 k^2 y^2 - \frac{r^2}{y^2}} \mathrm{d}y &= \frac{1}{2\pi k} \int_{1/\alpha}^\infty \left(\pi k - \frac{r}{y}\right) \mathrm{e}^{\left(-\pi k y + \frac{r}{y}\right)^2 + 2\pi k r} \mathrm{d}y \\
&\quad + \frac{1}{2\pi k} \int_{1/\alpha}^\infty \left(\pi k + \frac{r}{y}\right) \mathrm{e}^{\left(-\pi k y - \frac{r}{y}\right)^2 - 2\pi k r} \mathrm{d}y \\
&= \frac{1}{4\sqrt{\pi} k} \left(\mathrm{e}^{2\pi k r} \mathrm{erfc}\left(\frac{\pi k}{\alpha} + \alpha r\right) + \mathrm{e}^{-2\pi k r} \mathrm{erfc}\left(\frac{\pi k}{\alpha} - \alpha r\right)\right). \qquad (2.42)
\end{aligned}$$

For $\boldsymbol{k} = \boldsymbol{0}$ we obtain

$$\phi^{\boldsymbol{0}}(j) := \frac{2\sqrt{\pi}}{\|\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2\|} \lim_{s\to 0} \sum_{i=1}^N \xi_i \int_0^\alpha \frac{\mathrm{e}^{-z^2 x_{ij,3}^2} \, \mathrm{e}^{-\frac{z^2 s}{z^2+s}\|\tilde{\boldsymbol{x}}_{ij}\|^2}}{z^2+s} \mathrm{d}z.$$

We use the Taylor representation of the exponential function and obtain

$$\frac{1}{z^2+s} \, \mathrm{e}^{-z^2 x_{ij,3}^2} \, \mathrm{e}^{-\frac{z^2 s}{z^2+s}\|\tilde{\boldsymbol{x}}_{ij}\|^2} = \frac{1}{z^2+s} \sum_{n=0}^\infty \frac{(-1)^n}{n!} z^{2n} \left(x_{ij,3}^2 + \frac{s\|\tilde{\boldsymbol{x}}_{ij}\|^2}{z^2+s} \right)^n.$$

For each $m \in \{1, \dots, n-1\}$ we compute

$$\lim_{s\to 0} s^m \int_0^\alpha \frac{z^{2n}}{(z^2+s)^{m+1}} \mathrm{d}z = \lim_{s\to 0} s^m \int_0^\alpha z^{2n-2m-2} \mathrm{d}z = 0.$$

Thus, we have

$$\lim_{s\to 0} \int_0^\alpha \frac{\mathrm{e}^{-z^2 x_{ij,3}^2} \mathrm{e}^{-\frac{z^2 s}{z^2+s}\|\tilde{\boldsymbol{x}}_{ij}\|^2}}{z^2+s} \mathrm{d}z = \lim_{s\to 0} \int_0^\alpha \frac{1}{z^2+s} \sum_{n=0}^\infty \frac{(-1)^n}{n!} z^{2n} \left(x_{ij,3}^{2n} + \frac{s^n \|\tilde{\boldsymbol{x}}_{ij}\|^{2n}}{(z^2+s)^n} \right) \mathrm{d}z.$$

Via partial integration we easily see

$$\int_0^\alpha \frac{z^{2n}}{(z^2+s)^{n+1}} \mathrm{d}z = \begin{cases} \dfrac{\arctan(\alpha/\sqrt{s})}{\sqrt{s}} & : n = 0, \\ -\dfrac{\alpha^{2n-1}}{2n(z^2+s)^n} + \dfrac{2n-1}{2n} \displaystyle\int_0^\alpha \frac{z^{2n-2}}{(z^2+s)^n} \mathrm{d}z & : n > 0, \end{cases}$$

i.e., we can write

$$s^n \int_0^\alpha \frac{z^{2n}}{(z^2+s)^{n+1}} \mathrm{d}z = \beta_{0,n} \frac{s^n \arctan(\alpha/\sqrt{s})}{\sqrt{s}} + \sum_{m=1}^n \beta_{m,n} \frac{s^n \alpha^{2m-1}}{(\alpha^2+s)^m}.$$

with some constant coefficients $\beta_{m,n}$. The limit for $s \to 0$ is equal to zero, except for $n = 0$, where we only have a single term with $\beta_{0,0} = 1$. This contribution is independent from $\tilde{\boldsymbol{x}}_{ij}$ and vanishes due to charge neutrality as well as the applied differentiation for the dipole moments. We obtain

$$\begin{aligned} \phi^{\mathbf{0}}(j) &= \frac{2\sqrt{\pi}}{\|\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2\|} \lim_{s\to 0} \sum_{i=1}^N \xi_i \left(\int_0^\alpha \sum_{n=0}^\infty \frac{(-1)^n x_{ij,3}^{2n} z^{2n}}{n!(z^2+s)} \mathrm{d}z + \frac{\arctan(\alpha/\sqrt{s})}{\sqrt{s}} \right) \\ &= \frac{2\sqrt{\pi}}{\|\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2\|} \lim_{s\to 0} \sum_{i=1}^N \xi_i \frac{1}{\sqrt{s}} \sum_{n=0}^\infty \frac{1}{n!} \int_0^{\alpha/\sqrt{s}} \frac{(-1)^n x_{ij,3}^{2n} s^n y^{2n}}{y^2+1} \mathrm{d}y, \end{aligned}$$

where we substitute $y := z/\sqrt{s}$. By utilizing

$$\frac{1}{\sqrt{s}} \int_0^{\alpha/\sqrt{s}} \frac{y^{2n}}{1+y^2} \mathrm{d}y = (-1)^n \frac{\arctan(\alpha/\sqrt{s})}{\sqrt{s}} + (-1)^{n+1} \sum_{k=0}^{n-1} (-1)^k \frac{\alpha^{2k+1}}{(2k+1)s^{k+1}},$$

which is found by partial integration and induction in n, we finally obtain

$$\phi^{\mathbf{0}}(j) = \frac{2\sqrt{\pi}}{\|\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2\|} \lim_{s\to 0} \sum_{i=1}^N \xi_i \left(\frac{1}{\sqrt{s}} \arctan\left(\frac{\alpha}{\sqrt{s}}\right) + \sum_{n=1}^\infty \frac{(-1)^n x_{ij,3}^{2n} \alpha^{2n-1}}{n!(2n-1)} \right) + O(\sqrt{s}) \tag{2.43}$$

$$= \frac{2\sqrt{\pi}}{\|\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2\|} \sum_{i=1}^N \xi_i \left(\frac{-\mathrm{e}^{-\alpha^2 x_{ij,3}^2}}{\alpha} - \sqrt{\pi} x_{ij,3} \operatorname{erf}(\alpha x_{ij,3}) \right). \tag{2.44}$$

The constant terms are not relevant due to the applied differentiation operators as well as the assumed charge neutrality. The final result is obtained by putting together (2.40), (2.41), (2.42) and (2.44). ■

In summary we obtain in the case of 2d-periodic boundary conditions

$$\phi(j) = \phi^{\text{short}}(j) + \phi^{\text{F}}(j) + \phi^{\text{self}}(j),$$

with the short range potential as given in (2.22), where $\mathcal{S} := \mathbb{Z}^2 \times \{0\}$, the Fourier space part (2.38) and the self potential (2.24). The $\boldsymbol{k} = \mathbf{0}$ contribution is included in $\phi^{\text{F}}(j)$.

Remark 2.9. It is easy to see that the infinite Fourier sum (2.38) converges exponentially fast, similar to the 3d-periodic case. This can be shown as follows. For $k \neq 0$ we have, see (2.41) and (2.20),

$$\Theta^{\mathrm{p2}}(k,r) = 2\sqrt{\pi}\int_{1/\alpha}^{\infty} \mathrm{e}^{-\frac{r^2}{y^2}-\pi^2k^2y^2}\,\mathrm{d}y \leq \int_{1/\alpha}^{\infty} \mathrm{e}^{-\pi^2k^2y^2}\,\mathrm{d}y = \frac{\operatorname{erfc}\left(\frac{\pi k}{\alpha}\right)}{2k\sqrt{\pi}} \sim \frac{\alpha}{2\pi^2k^2}\,\mathrm{e}^{-\pi^2k^2/\alpha^2}.$$

□

For the considered case of 2d-periodic boundary conditions we indeed only require two linear independent vectors in order to described the underlying periodicity. However, the particle positions are oftentimes given as elements of a primary simulation box, spanned by three linearly independent vectors, cf. (2.1). Only two of these vectors are essential in order to describe the periodicity. Obviously, the particle system may be rotated appropriately in order to end up with a structure of the form (2.37). The third vector may be replaced in order to end up with a preferably simple box shape, which might be more favorable in consideration of a possible implementation. This is the purpose of the following lemma. The simple box shape, which is constructed hereby, will be exploited for the implementation presented in Section 4.1.2.

Lemma 2.10. *For each system of particles at positions*

$$\boldsymbol{x}_j \in [-1/2, 1/2)\boldsymbol{\ell}_1 \times [-1/2, 1/2)\boldsymbol{\ell}_2 \times [-1/2, 1/2]\boldsymbol{\ell}_3,$$

where $j = 1, \ldots, N$ and $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3$ are the linearly independent spanning vectors of the primary simulation box, there is an, in terms of 2d-periodic boundary conditions with respect to $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$, equivalent particle system, where the primary simulation box is spanned by three vectors of the form

$$\tilde{\boldsymbol{\ell}}_1 = \begin{pmatrix} \tilde{l}_{11} \\ \tilde{l}_{12} \\ 0 \end{pmatrix}, \tilde{\boldsymbol{\ell}}_2 = \begin{pmatrix} \tilde{l}_{21} \\ \tilde{l}_{22} \\ 0 \end{pmatrix}, \text{ indicating the periodic directions, and } \tilde{\boldsymbol{\ell}}_3 = \begin{pmatrix} 0 \\ 0 \\ \tilde{l}_{33} \end{pmatrix}.$$

In particular, this fulfills the conditions of Lemma 2.8.

Proof. Obviously, we do not change the interaction quantities if we simply rotate the particle system around the origin. Therefor, we consider the vector $\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2$ and compute its representation in spherical coordinates, i.e.,

$$\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2 = \|\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2\| \begin{pmatrix} \sin\vartheta\cos\varphi \\ \sin\vartheta\sin\varphi \\ \cos\vartheta \end{pmatrix}, \tag{2.45}$$

where $\varphi \in [0, 2\pi)$ and $\vartheta \in [0, \pi]$. In the following we define the rotation matrix

$$R := \begin{pmatrix} \cos\vartheta & 0 & -\sin\vartheta \\ 0 & 1 & 0 \\ \sin\vartheta & 0 & \cos\vartheta \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.46}$$

We simply compute

$$R(\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2) = \begin{pmatrix} 0 \\ 0 \\ \|\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2\| \end{pmatrix}$$

and for $j \in \{1, 2\}$ we get

$$\tilde{\boldsymbol{\ell}}_j := R\boldsymbol{\ell}_j = \begin{pmatrix} * \\ * \\ 0 \end{pmatrix} =: \begin{pmatrix} \tilde{l}_{j1} \\ \tilde{l}_{j2} \\ 0 \end{pmatrix},$$

which follows from the fact that $\boldsymbol{\ell}_1^\top(\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2) = \boldsymbol{\ell}_2^\top(\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2) = 0$.

Furthermore, we set

$$\tilde{\boldsymbol{\ell}}_3 := R\boldsymbol{\ell}_3 =: \begin{pmatrix} \tilde{l}_{31} \\ \tilde{l}_{32} \\ \tilde{l}_{33} \end{pmatrix}.$$

The particle positions as well as the dipole moments are rotated analogously, i.e, we set $\tilde{\boldsymbol{x}}_j := R\boldsymbol{x}_j$ for all $j = 1, \ldots, N$ and $\tilde{\boldsymbol{\mu}}_j := R\boldsymbol{\mu}_j$ for all $j = N_\mathrm{c} + 1, \ldots, N$.

The original positions can be written in the form

$$\boldsymbol{x}_j = \lambda_{j,1}\boldsymbol{\ell}_1 + \lambda_{j,2}\boldsymbol{\ell}_2 + \lambda_{j,3}\boldsymbol{\ell}_3$$

with $\lambda_{j,1}, \lambda_{j,2} \in [-1/2, 1/2)$ and $\lambda_{j,3} \in [-1/2, 1/2]$. We obtain

$$\begin{aligned} \tilde{\boldsymbol{x}}_j = R\boldsymbol{x}_j &= \lambda_{j,1}R\boldsymbol{\ell}_1 + \lambda_{j,2}R\boldsymbol{\ell}_2 + \lambda_{j,3}R\boldsymbol{\ell}_3 \\ &= \lambda_{j,1}\tilde{\boldsymbol{\ell}}_1 + \lambda_{j,2}\tilde{\boldsymbol{\ell}}_2 + \lambda_{j,3}\tilde{\boldsymbol{\ell}}_3 = \begin{pmatrix} \lambda_{j,1}\tilde{l}_{11} + \lambda_{j,2}\tilde{l}_{21} + \lambda_{j,3}\tilde{l}_{31} \\ \lambda_{j,1}\tilde{l}_{12} + \lambda_{j,2}\tilde{l}_{22} + \lambda_{j,3}\tilde{l}_{32} \\ \lambda_{j,3}\tilde{l}_{33} \end{pmatrix}. \end{aligned}$$

Since $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are linearly independent we can find a representation of the form

$$\begin{pmatrix} \tilde{l}_{31} \\ \tilde{l}_{32} \end{pmatrix} = \alpha_1 \begin{pmatrix} \tilde{l}_{11} \\ \tilde{l}_{12} \end{pmatrix} + \alpha_2 \begin{pmatrix} \tilde{l}_{21} \\ \tilde{l}_{22} \end{pmatrix}$$

with $\alpha_1, \alpha_2 \in \mathbb{R}$, and we obtain

$$\tilde{\boldsymbol{x}}_j = \begin{pmatrix} (\lambda_{j,1} + \alpha_1\lambda_{j,3})\tilde{l}_{11} + (\lambda_{j,2} + \alpha_2\lambda_{j,3})\tilde{l}_{21} \\ (\lambda_{j,1} + \alpha_1\lambda_{j,3})\tilde{l}_{12} + (\lambda_{j,2} + \alpha_2\lambda_{j,3})\tilde{l}_{22} \\ \lambda_{j,3}\tilde{l}_{33} \end{pmatrix}.$$

Now, we can find for each j some numbers $z_{j,1}, z_{j,2} \in \mathbb{Z}$ such that

$$\eta_{j,1} := \lambda_{j,1} + \alpha_1\lambda_{j,3} + z_{j,1} \in [-1/2, 1/2) \quad \text{and} \quad \eta_{j,2} := \lambda_{j,2} + \alpha_2\lambda_{j,3} + z_{j,2} \in [-1/2, 1/2).$$

Since periodic boundary conditions are applied with respect to $\tilde{\boldsymbol{\ell}}_1$ and $\tilde{\boldsymbol{\ell}}_2$, we obtain an equivalent particle system with

$$\breve{\boldsymbol{x}}_j := \eta_{j,1}\tilde{\boldsymbol{\ell}}_1 + \eta_{j,2}\tilde{\boldsymbol{\ell}}_2 + \lambda_{j,3}\begin{pmatrix} 0 \\ 0 \\ \tilde{l}_{33} \end{pmatrix},$$

which are the positions in the primary box. ■

A graphical illustration is given in Figure 2.3.

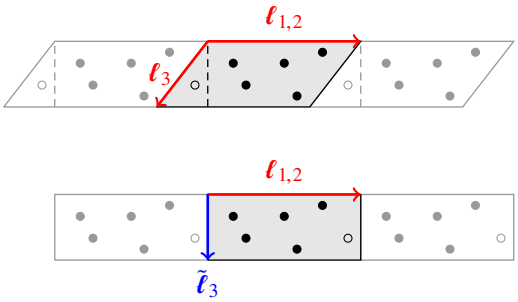


Figure 2.3: The box vector regarding the nonperiodic dimension may be chosen orthogonal to the periodic directions.

Algorithm 2.1 (Equivalent 2d-periodic system)**.**
Input: Particle positions $\boldsymbol{x}_j$, $j = 1, \ldots, N$, in the primary box spanned by the vectors $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$, dipole moments $\boldsymbol{\mu}_j$, $j = N_\mathrm{c} + 1, \ldots, N$. Periodic boundary conditions are assumed with respect to $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$, i.e., $\boldsymbol{x}_j \in \mathbb{T}\boldsymbol{\ell}_1 + \mathbb{T}\boldsymbol{\ell}_2 + [-1/2, 1/2)\boldsymbol{\ell}_3$.

i) Compute $\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2$ and its representation in spherical coordinates, see (2.45).

ii) Compute the rotation matrix R, as defined in (2.46).

iii) Set $\tilde{\boldsymbol{\ell}}_i := R\boldsymbol{\ell}_i$ for all $i \in \{1, 2, 3\}$.

iv) Rotate the dipole moments: Set $\tilde{\boldsymbol{\mu}}_j := R\boldsymbol{\mu}_j$ for all $j = N_\mathrm{c} + 1, \ldots, N$.

v) Compute $\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} := \begin{pmatrix} \tilde{l}_{11} & \tilde{l}_{21} \\ \tilde{l}_{12} & \tilde{l}_{22} \end{pmatrix}^{-1} \begin{pmatrix} \tilde{l}_{32} \\ \tilde{l}_{32} \end{pmatrix}$.

vi) For $j = 1, \ldots, N$:

a) Set $\boldsymbol{\lambda}_j := \boldsymbol{L}^{-1}\boldsymbol{x}_j$, i.e., we have $\boldsymbol{x}_j = \lambda_{j,1}\boldsymbol{\ell}_1 + \lambda_{j,2}\boldsymbol{\ell}_2 + \lambda_{j,3}\boldsymbol{\ell}_3$.

b) Find $z_{j,1} \in \mathbb{Z}$ such that $\eta_{j,1} := \lambda_{j,1} + \alpha_1\lambda_{j,3} + z_{j,1} \in [-1/2, 1/2)$.

c) Find $z_{j,2} \in \mathbb{Z}$ such that $\eta_{j,2} := \lambda_{j,2} + \alpha_2\lambda_{j,3} + z_{j,2} \in [-1/2, 1/2)$.

d) Set $\tilde{\boldsymbol{x}}_j := \eta_{j,1}\tilde{\boldsymbol{\ell}}_1 + \eta_{j,2}\tilde{\boldsymbol{\ell}}_2 + \lambda_{j,3}\begin{pmatrix} 0 \\ 0 \\ \tilde{l}_{33} \end{pmatrix}$.

Output: Particle positions $\tilde{\boldsymbol{x}}_j$, $j = 1, \ldots, N$, dipole moments $\tilde{\boldsymbol{\mu}}_j$, $j = N_\mathrm{c} + 1, \ldots, N$, and new basis vectors $\tilde{\boldsymbol{\ell}}_1 = (\tilde{l}_{11}, \tilde{l}_{12}, 0)^\top$, $\tilde{\boldsymbol{\ell}}_2 = (\tilde{l}_{21}, \tilde{l}_{22}, 0)^\top$, $\tilde{\boldsymbol{\ell}}_3 = (0, 0, \tilde{l}_{33})^\top$. □

Example 2.11. We consider a small particle system containing $N_\mathrm{c} = 10$ charges $q_j = (-1)^j$ and $N_\mathrm{d} = 10$ dipoles with randomly oriented dipole moments $\boldsymbol{\mu}_j \in \mathbb{R}^3$, $\|\boldsymbol{\mu}_j\| = 1$. The particles have random positions $\boldsymbol{x}_j$ in a box spanned by

$$\boldsymbol{\ell}_1 := \begin{pmatrix} 10 \\ 10 \\ 0 \end{pmatrix}, \boldsymbol{\ell}_2 := \begin{pmatrix} 0 \\ 10 \\ 10 \end{pmatrix}, \boldsymbol{\ell}_3 := \begin{pmatrix} 10 \\ 0 \\ 10 \end{pmatrix},$$

where periodic boundary conditions are applied with respect to $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$.

The equivalent particle system was computed via Algorithm 2.1, resulting in

$$\tilde{\boldsymbol{\ell}}_1 = \begin{pmatrix} 0 \\ 2\sqrt{10} \\ 0 \end{pmatrix} \approx \begin{pmatrix} 0 \\ 14.14 \\ 0 \end{pmatrix}, \tilde{\boldsymbol{\ell}}_2 = \begin{pmatrix} -5\sqrt{6} \\ 2\sqrt{5} \\ 0 \end{pmatrix} \approx \begin{pmatrix} -12.25 \\ 7.07 \\ 0 \end{pmatrix}, \tilde{\boldsymbol{\ell}}_3 = \begin{pmatrix} 0 \\ 0 \\ \frac{20}{3}\sqrt{3} \end{pmatrix} \approx \begin{pmatrix} 0 \\ 0 \\ 11.55 \end{pmatrix}.$$

The total energies of the systems should be equal. In order to verify the equivalence of the two particle systems, we approximated the total energies via

$$U \approx \frac{1}{2} \sum_{i,j=1}^{N} {\sum_{\substack{\boldsymbol{n} \in \mathbb{Z}^2 \times \{0\} \\ -C \le n_1, n_2 \le C}}}' \frac{\xi_i \xi_j}{\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|} =: U_{\approx}, \tag{2.47}$$

i.e., we computed the slowly converging series via a cubic truncation scheme using some cutoff $C \in \mathbb{N}$ in each periodic direction. Indeed, the two energies approach the same value for growing cutoff parameter C, see Figure 2.4. □

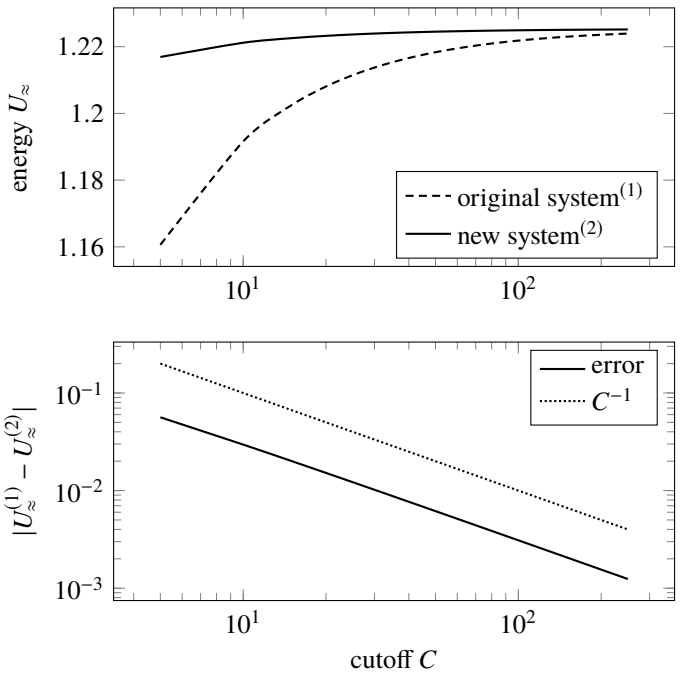


Figure 2.4: The electrostatic energy U has been approximated based on a cubic truncation scheme for growing cutoff $C \in \mathbb{N}$, see (2.47). The same value is approached for the original system$^{(1)}$ and the computed equivalent system$^{(2)}$ with rate C^{-1}.

2.2.3 1d-periodic boundary conditions

We set $\mathcal{S} := \mathbb{Z} \times \{0\}^2$ and denote by $\boldsymbol{\ell}_1$ the lattice vector indicating the periodic direction, i.e., without loss of generality we a assume the first direction to be the periodic one, see right hand side of Figure 2.1.

In the following we denote by

$$\Gamma(s,x) := \int_x^\infty t^{s-1}\,\mathrm{e}^{-t}\mathrm{d}t \tag{2.48}$$

the upper incomplete Gamma function. In the case $s = 0$ we have [AS72, Eq. 6.5.32]

$$\Gamma(0,x) = -\gamma - \ln(x) - \sum_{k=1}^{\infty}\frac{(-1)^k}{k!k}x^k, \tag{2.49}$$

where γ is the Euler-Mascheroni constant, and via partial integration we obtain the asymptotic expansion

$$\Gamma(0,x) = \frac{\mathrm{e}^{-x}}{x} - \int_0^\infty \frac{\mathrm{e}^{-t}}{t^2}\mathrm{d}t = \frac{\mathrm{e}^{-x}}{x} - \frac{\mathrm{e}^{-x}}{x^2} + \int_x^\infty \frac{2\,\mathrm{e}^{-t}}{t^3}\mathrm{d}t = \frac{\mathrm{e}^{-x}}{x}\sum_{n=0}^{\infty}\frac{(-1)^n n!}{x^n}, \tag{2.50}$$

see [AS72, Eq. 5.1.11]. Furthermore we denote by

$$K_\nu(x,y) := \int_1^\infty t^{\nu-1}\,\mathrm{e}^{-xt-y/t}\mathrm{d}t$$

the incomplete modified Bessel function of the second kind [Har08].

In the following lemma we present the Ewald formulas for the long range potentials under 1d-periodic boundary conditions, which has already been derived in [Por00]. The paper covers both, the Coulomb case as well as dipole-dipole interactions. The following proof is of a similar structure as the one presented in [Nes12, NPP15b] and applies to systems consisting of charges and dipoles. Similarly to the case of 3d-periodic boundary conditions, we make use of the theory of convergence factors, see [dLPS80]. The proof presented in [Por00] is not based on convergence factor but ends up with the same result.

Lemma 2.12. *Suppose that the given particles are distributed in a primary simulation box, spanned by the three linearly independent vectors $\boldsymbol{\ell}_1$, $\boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$, cf.* (2.1)*. Assume that periodic boundary conditions are applied with respect to $\boldsymbol{\ell}_1$, which is of the form*

$$\boldsymbol{\ell}_1 = \begin{pmatrix} l_{11} \\ 0 \\ 0 \end{pmatrix}. \tag{2.51}$$

Let further the particle positions be denoted by

$$\boldsymbol{x}_j = \begin{pmatrix} x_{j,1} \\ x_{j,2} \\ x_{j,3} \end{pmatrix} =: \begin{pmatrix} x_{j,1} \\ \tilde{\boldsymbol{x}}_j \end{pmatrix}.$$

Then, we have for a charge neutral system

$$\lim_{C\to\infty} \sum_{\boldsymbol{n}\in\mathbb{Z}\times\{0\}^2:\|\boldsymbol{L}\boldsymbol{n}\|\le C} \sum_{i=1}^{N} \xi_i \frac{\operatorname{erf}(\alpha\|\boldsymbol{x}_{ij}+\boldsymbol{L}\boldsymbol{n}\|)}{\|\boldsymbol{x}_{ij}+\boldsymbol{L}\boldsymbol{n}\|} = \phi^{\mathrm{F}}(j),$$

where the Fourier space representations $\phi^{\mathrm{F}}(j)$ of the long range potentials are given by

$$\phi^{\mathrm{F}}(j) = \frac{1}{|l_{11}|}\sum_{k\in\mathbb{Z}}\sum_{i=1}^{N}\xi_i\,\mathrm{e}^{2\pi\mathrm{i}kx_{ij,1}/l_{11}}\Theta^{\mathrm{p1}}\left(k/l_{11}, \|\tilde{\boldsymbol{x}}_{ij}\|\right). \tag{2.52}$$

The function Θ^{p1} *is defined via*

$$\Theta^{\mathrm{p1}}(k,r) := \begin{cases} K_0\left(\frac{\pi^2 k^2}{\alpha^2}, \alpha^2 r^2\right) & : k \neq 0, \\ -\left(\gamma + \Gamma(0, \alpha^2 r^2) + \ln(\alpha^2 r^2)\right) & : k = 0. \end{cases} \tag{2.53}$$

Proof. The lattice vectors $\boldsymbol{n} \in \mathbb{Z} \times \{0\}^2$ can be written in the form

$$\boldsymbol{n} = \begin{pmatrix} n_1 \\ 0 \\ 0 \end{pmatrix}.$$

Since the lattice vector $\boldsymbol{\ell}_1$ is of the same structure we have

$$\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|^2 = \|x_{ij,1} + l_{11} n_1\|^2 + \|\tilde{\boldsymbol{x}}_{ij}\|^2.$$

Analogously to the other types of periodic boundary conditions, we make use of the convergence factors $\mathrm{e}^{-s\|\boldsymbol{L}\boldsymbol{n}\|^2} = \mathrm{e}^{-s\, l_{11}^2 n_1^2}$ and obtain

$$\begin{aligned} \Phi_s(j) &= \frac{2}{\sqrt{\pi}} \int_0^\alpha \sum_{i=1}^N \xi_i \, \mathrm{e}^{-z^2\|\tilde{\boldsymbol{x}}_{ij}\|^2} \, \mathrm{e}^{-\frac{z^2 s}{z^2+s} x_{ij,1}^2} \sum_{n_1 \in \mathbb{Z}} \mathrm{e}^{-\frac{z^4}{s+z^2}\left\| \frac{z^2+s}{z^2} l_{11} n_1 + x_{ij,1}\right\|^2} \\ &= \frac{2}{\sqrt{\pi}} \int_0^\alpha \sum_{i=1}^N \xi_i \frac{\sqrt{\pi}\, \mathrm{e}^{-z^2\|\tilde{\boldsymbol{x}}_{ij}\|^2} \, \mathrm{e}^{-\frac{z^2 s}{z^2+s} x_{ij,1}^2}}{l_{11}\sqrt{z^2+s}} \sum_{k \in \mathbb{Z}} \mathrm{e}^{-\frac{\pi^2 k^2 z^2}{l_{11}^2 (s+z^2)}} \, \mathrm{e}^{\frac{2\pi \mathrm{i} z^2 k x_{ij,1}}{(s+z^2) l_{11}}}, \end{aligned}$$

where we apply the Poisson summation formula (2.28) with $d = 1$. For $k \neq 0$ we simply shift the limit for $s \to 0$ into the integrals and get

$$\begin{aligned} \phi^{\mathrm{F}}(j) = \frac{1}{|l_{11}|} \sum_{k \in \mathbb{Z}\setminus\{0\}} \sum_{i=1}^N \xi_i \, \mathrm{e}^{2\pi \mathrm{i} k x_{ij,1}/l_{11}} \int_0^\alpha \frac{2}{z} \, \mathrm{e}^{-z^2\|\tilde{\boldsymbol{x}}_{ij}\|^2 - \pi^2 k^2/(l_{11}^2 z^2)} \mathrm{d}z \\ + \lim_{s \to 0} \frac{2}{|l_{11}|} \sum_{i=1}^N \xi_i \int_0^\alpha \frac{\mathrm{e}^{-z^2\|\tilde{\boldsymbol{x}}_{ij}\|^2} \, \mathrm{e}^{-\frac{z^2 s}{z^2+s} x_{ij,1}^2}}{\sqrt{z^2+s}} \mathrm{d}z, \end{aligned} \tag{2.54}$$

where

$$\int_0^\alpha \frac{2}{z} \, \mathrm{e}^{-z^2\|\tilde{\boldsymbol{x}}_{ij}\|^2 - \frac{\pi^2 k^2}{l_{11}^2 z^2}} \mathrm{d}z \overset{y=\alpha^2/z^2}{=} \int_1^\infty \frac{1}{y} \, \mathrm{e}^{-\frac{\alpha^2\|\tilde{\boldsymbol{x}}_{ij}\|^2}{y} - \frac{\pi^2 k^2 y}{l_{11}^2 \alpha^2}} \mathrm{d}y = K_0\left(\frac{\pi^2 k^2}{\alpha^2 l_{11}^2}, \alpha^2 \|\tilde{\boldsymbol{x}}_{ij}\|^2\right) \tag{2.55}$$

for all $k \neq 0$. In the case $k = 0$ we see that the terms depending on $x_{ij,1}$ become not relevant for $s \to 0$ since

$$\frac{\mathrm{e}^{-z^2\|\tilde{\boldsymbol{x}}_{ij}\|^2} \, \mathrm{e}^{-\frac{z^2 s}{z^2+s} x_{ij,1}^2}}{\sqrt{z^2+s}} = \frac{1}{\sqrt{z^2+s}} \sum_{n=0}^\infty \frac{(-1)^n}{n!} z^{2n} \left(\|\tilde{\boldsymbol{x}}_{ij}\|^2 + \frac{s}{z^2+s} x_{ij,1}^2\right)^n$$

and the limits

$$\lim_{s \to 0} \int_0^\alpha \frac{z^{2n}}{(z^2+s)^{m+1/2}} \mathrm{d}z = \int_0^\alpha z^{2n-2m-1/2} \mathrm{d}z \quad \forall\, m \leq n$$

exist.

Thus, we have

$$\begin{aligned}\phi^0(j) &:= \lim_{s\to 0}\frac{2}{|l_{11}|}\sum_{i=1}^{N}\xi_i\int_0^{\alpha}\sum_{n=0}^{\infty}\frac{(-1)^n\|\tilde{\boldsymbol{x}}_{ij}\|^{2n}z^{2n}}{n!\sqrt{z^2+s}}\mathrm{d}z\\ &= \lim_{s\to 0}\frac{2}{|l_{11}|}\sum_{i=1}^{N}\xi_i\int_0^{\alpha/\sqrt{s}}\sum_{n=0}^{\infty}\frac{(-1)^n\|\tilde{\boldsymbol{x}}_{ij}\|^{2n}s^n y^{2n}}{n!\sqrt{y^2+1}}\mathrm{d}y.\end{aligned}$$

By utilizing

$$\int_0^{\beta}\frac{x^{2n}}{\sqrt{x^2+1}}\mathrm{d}x = (-1)^n\frac{(2n-1)!!}{(2n)!!}\left(\sqrt{1+\beta^2}\sum_{k=1}^{n}(-1)^k\frac{(2k-2)!!}{(2k-1)!!}\beta^{2k-1}+\ln(\beta+\sqrt{1+\beta^2})\right)$$

with $\beta := {}^{\alpha}/_{\sqrt{s}}$, we see that only the terms $k = n$ are relevant when considering the limit for $s \to 0$ and compute

$$\begin{aligned}\phi^0(j) &= \lim_{s\to 0}\frac{2}{|l_{11}|}\sum_{i=1}^{N}\xi_i\left(\sum_{n=1}^{\infty}(-1)^n\|\tilde{\boldsymbol{x}}_{ij}\|^{2n}\sqrt{s+\alpha^2}\frac{\alpha^{2n-1}}{n!(2n)}+\ln\left(\frac{\alpha}{\sqrt{s}}+\sqrt{1+\frac{\alpha^2}{s}}\right)\right)+O(s)\\ &= \frac{2}{|l_{11}|}\sum_{i=1}^{N}\xi_i\sum_{n=1}^{\infty}\frac{(-1)^n\|\tilde{\boldsymbol{x}}_{ij}\|^{2n}\alpha^{2n}}{n!(2n)} \qquad (2.56)\\ &= -\frac{1}{|l_{11}|}\sum_{i=1}^{N}\xi_i\left(\gamma+\ln(\alpha^2\|\tilde{\boldsymbol{x}}_{ij}\|^2)+\Gamma(0,\alpha^2\|\tilde{\boldsymbol{x}}_{ij}\|^2)\right). \qquad (2.57)\end{aligned}$$

The constant terms are not relevant due to charge neutrality. The final result is obtained by putting together (2.54), (2.55) and (2.57). ■

Remark 2.13. It is easy to see that the infinite Fourier sum (2.52) converges exponentially fast for $k \to \infty$, similar to the 3d-periodic case. This can be shown as follows. For $k \neq 0$ we have

$$\begin{aligned}\Theta^{\mathrm{p1}}(k,r) &= \int_1^{\infty}\frac{1}{t}\,\mathrm{e}^{-\frac{\pi^2k^2t}{\alpha^2}-\frac{\alpha^2r^2}{t}}\mathrm{d}t \qquad (2.58)\\ &\leq \int_1^{\infty}\frac{1}{t}\,\mathrm{e}^{-\frac{\pi^2k^2t}{\alpha^2}}\mathrm{d}t = \Gamma\left(0,\frac{\pi^2k^2}{\alpha^2}\right)\sim\frac{\alpha^2}{\pi^2k^2}\,\mathrm{e}^{-\pi^2k^2/\alpha^2},\end{aligned}$$

cf. (2.50). □

The particle positions may be given as elements of a primary simulation box spanned by three linear independent vectors, cf. (2.1), although only one vector is required in order to describe the underlying periodicity in case of 1d-periodic boundary conditions. Obviously, the particle system may be rotated appropriately in order to end up with a structure of the form (2.51). The two remaining vectors may be replaced in order to achieve a preferably simple box shape, which might be more favorable when it comes to implementing an efficient algorithm. The simple box shape, which is constructed based on the following lemma, will be exploited for the implementation presented in Section 4.1.2.

Lemma 2.14. *For each system of particles at positions*

$$\boldsymbol{x}_j \in [-1/2, 1/2)\boldsymbol{\ell}_1 \times [-1/2, 1/2]\boldsymbol{\ell}_2 \times [-1/2, 1/2]\boldsymbol{\ell}_3,$$

where $j = 1, \dots, N$ and $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3$ are the linearly independent spanning vectors of the primary simulation box, there is an, in terms of 1*d-periodic boundary conditions with respect to $\boldsymbol{\ell}_1$, equivalent particle system, where the primary simulation box is spanned by three vectors of the form*

$$\tilde{\boldsymbol{\ell}}_1 = \begin{pmatrix} \tilde{l}_{11} \\ 0 \\ 0 \end{pmatrix}, \text{ indicating the periodic direction, } \tilde{\boldsymbol{\ell}}_2 = \begin{pmatrix} 0 \\ \tilde{l}_{22} \\ \tilde{l}_{23} \end{pmatrix} \text{ and } \tilde{\boldsymbol{\ell}}_3 = \begin{pmatrix} 0 \\ \tilde{l}_{32} \\ \tilde{l}_{33} \end{pmatrix}.$$

In particular, this fulfills the conditions of Lemma 2.12*.*

Proof. Obviously, we do not change the interaction quantities if we simply rotate the particle system around the origin. Therefor, we consider the vector $\boldsymbol{\ell}_1$ and compute its representation in spherical coordinates, i.e.,

$$\boldsymbol{\ell}_1 = \|\boldsymbol{\ell}_1\| \begin{pmatrix} \sin\vartheta\cos\varphi \\ \sin\vartheta\sin\varphi \\ \cos\vartheta \end{pmatrix}, \tag{2.59}$$

where $\varphi \in [0, 2\pi)$ and $\vartheta \in [0, \pi]$. In the following we define the rotation matrix

$$R := \begin{pmatrix} \sin\vartheta & 0 & \cos\vartheta \\ 0 & 1 & 0 \\ -\cos\vartheta & 0 & \sin\vartheta \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.60}$$

We simply compute

$$\tilde{\boldsymbol{\ell}}_1 := R\boldsymbol{\ell}_1 = \begin{pmatrix} \|\boldsymbol{\ell}_1\| \\ 0 \\ 0 \end{pmatrix}$$

and for $j \in \{2, 3\}$ we set

$$\tilde{\boldsymbol{\ell}}_j := R\boldsymbol{\ell}_j =: \begin{pmatrix} \tilde{l}_{j1} \\ \tilde{l}_{j2} \\ \tilde{l}_{j3} \end{pmatrix}.$$

The particle positions and the dipole moments are rotated analogously, i.e, we set $\tilde{\boldsymbol{x}}_j := R\boldsymbol{x}_j$ for all $j = 1, \dots, N$ as well as $\tilde{\boldsymbol{\mu}}_j := R\boldsymbol{\mu}_j$ for all $j = N_c + 1, \dots, N$.

The original positions can be written in the form

$$\boldsymbol{x}_j = \lambda_{j,1}\boldsymbol{\ell}_1 + \lambda_{j,2}\boldsymbol{\ell}_2 + \lambda_{j,3}\boldsymbol{\ell}_3$$

with $\lambda_{j,1} \in [-1/2, 1/2)$ and $\lambda_{j,2}, \lambda_{j,3} \in [-1/2, 1/2]$. We obtain

$$\begin{aligned} \tilde{\boldsymbol{x}}_j = R\boldsymbol{x}_j &= \begin{pmatrix} \lambda_{j,1}\|\boldsymbol{\ell}_1\| + \lambda_{j,2}\tilde{l}_{21} + \lambda_{j,3}\tilde{l}_{31} \\ \lambda_{j,2}\tilde{l}_{22} + \lambda_{j,3}\tilde{l}_{32} \\ \lambda_{j,2}\tilde{l}_{23} + \lambda_{j,3}\tilde{l}_{33} \end{pmatrix} \\ &= \left(\lambda_{j,1} + \frac{\lambda_{j,2}\tilde{l}_{21}}{\|\boldsymbol{\ell}_1\|} + \frac{\lambda_{j,3}\tilde{l}_{31}}{\|\boldsymbol{\ell}_1\|}\right) \begin{pmatrix} \|\boldsymbol{\ell}_1\| \\ 0 \\ 0 \end{pmatrix} + \lambda_{j,2} \begin{pmatrix} 0 \\ \tilde{l}_{22} \\ \tilde{l}_{23} \end{pmatrix} + \lambda_{j,3} \begin{pmatrix} 0 \\ \tilde{l}_{32} \\ \tilde{l}_{33} \end{pmatrix}. \end{aligned}$$

Now, we can find for each j some $z_j \in \mathbb{Z}$ such that

$$\eta_j := \lambda_{j,1} + \frac{\lambda_{j,2}\tilde{l}_{21}}{\|\boldsymbol{\ell}_1\|} + \frac{\lambda_{j,3}\tilde{l}_{31}}{\|\boldsymbol{\ell}_1\|} + z_j \in [-1/2, 1/2).$$

Since periodic boundary conditions are applied with respect to $\boldsymbol{\ell}_1$, we obtain an equivalent particle system with

$$\breve{\boldsymbol{x}}_j := \eta_j \begin{pmatrix} \|\boldsymbol{\ell}_1\| \\ 0 \\ 0 \end{pmatrix} + \lambda_{j,2} \begin{pmatrix} 0 \\ \tilde{l}_{22} \\ \tilde{l}_{23} \end{pmatrix} + \lambda_{j,3} \begin{pmatrix} 0 \\ \tilde{l}_{32} \\ \tilde{l}_{33} \end{pmatrix}, \tag{2.61}$$

which are the positions in the primary box. ■

Algorithm 2.2 (Equivalent 1d-periodic system)**.**
Input: Particle positions $\boldsymbol{x}_j$, $j = 1, \ldots, N$, in the primary box spanned by the vectors $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$, dipole moments $\boldsymbol{\mu}_j$, $j = N_\mathrm{c} + 1, \ldots, N$. Periodic boundary conditions are assumed with respect to $\boldsymbol{\ell}_1$, i.e., $\boldsymbol{x}_j \in \mathbb{T}\boldsymbol{\ell}_1 + [-1/2, 1/2]\boldsymbol{\ell}_2 + [-1/2, 1/2]\boldsymbol{\ell}_3$.

i) Compute the representation of $\boldsymbol{\ell}_1$ in spherical coordinates, see (2.59).

ii) Compute the rotation matrix R, as defined in (2.60).

iii) Set $\tilde{\boldsymbol{\ell}}_i := R\boldsymbol{\ell}_i$ for all $i \in \{1, 2, 3\}$.

iv) Rotate the dipole moments: Set $\tilde{\boldsymbol{\mu}}_j := R\boldsymbol{\mu}_j$ for all $j = N_\mathrm{c} + 1, \ldots, N$.

v) For $j = 1, \ldots, N$:

   a) Set $\boldsymbol{\lambda}_j := \boldsymbol{L}^{-1}\boldsymbol{x}_j$, i.e., we have $\boldsymbol{x}_j = \lambda_{j,1}\boldsymbol{\ell}_1 + \lambda_{j,2}\boldsymbol{\ell}_2 + \lambda_{j,3}\boldsymbol{\ell}_3$.

   b) Find $z_j \in \mathbb{Z}$ such that $\eta_{j,1} := \lambda_{j,1} + \frac{\lambda_{j,2}\tilde{l}_{21}}{\|\boldsymbol{\ell}_1\|} + \frac{\lambda_{j,3}\tilde{l}_{31}}{\|\boldsymbol{\ell}_1\|} + z_j \in [-1/2, 1/2)$.

   c) Set $\tilde{\boldsymbol{x}}_j := \eta_{j,1}\tilde{\boldsymbol{\ell}}_1 + \lambda_{j,2}\begin{pmatrix} 0 \\ \tilde{l}_{22} \\ \tilde{l}_{23} \end{pmatrix} + \lambda_{j,3}\begin{pmatrix} 0 \\ \tilde{l}_{32} \\ \tilde{l}_{33} \end{pmatrix}$.

Output: Particle positions $\tilde{\boldsymbol{x}}_j$, $j = 1, \ldots, N$, dipole moments $\tilde{\boldsymbol{\mu}}_j$, $j = N_\mathrm{c} + 1, \ldots, N$, and new basis vectors $\tilde{\boldsymbol{\ell}}_1 = (|\tilde{l}_{11}|, 0, 0)^\top$, $\tilde{\boldsymbol{\ell}}_2 = (0, \tilde{l}_{22}, \tilde{l}_{23})^\top$, $\tilde{\boldsymbol{\ell}}_3 = (0, \tilde{l}_{32}, \tilde{l}_{33})^\top$. □

Example 2.15. We consider a small particle system containing $N_\mathrm{c} = 10$ charges $q_j = (-1)^j$ and $N_\mathrm{d} = 10$ dipoles with randomly oriented dipole moments $\boldsymbol{\mu}_j \in \mathbb{R}^3$, $\|\boldsymbol{\mu}_j\| = 1$. The particles have random positions $\boldsymbol{x}_j$ in a box spanned by

$$\boldsymbol{\ell}_1 := \begin{pmatrix} 10 \\ 10 \\ 0 \end{pmatrix}, \boldsymbol{\ell}_2 := \begin{pmatrix} 0 \\ 10 \\ 10 \end{pmatrix}, \boldsymbol{\ell}_3 := \begin{pmatrix} 10 \\ 0 \\ 10 \end{pmatrix},$$

where periodic boundary conditions are applied with respect to $\boldsymbol{\ell}_1$.

The equivalent particle system was computed via Algorithm 2.2, resulting in

$$\tilde{\boldsymbol{\ell}}_1 = \begin{pmatrix} 10\sqrt{2} \\ 0 \\ 0 \end{pmatrix} \approx \begin{pmatrix} 14.14 \\ 0 \\ 0 \end{pmatrix}, \tilde{\boldsymbol{\ell}}_2 = \begin{pmatrix} 0 \\ 5\sqrt{2} \\ 10 \end{pmatrix} \approx \begin{pmatrix} 0 \\ 7.07 \\ 10 \end{pmatrix}, \tilde{\boldsymbol{\ell}}_3 = \begin{pmatrix} 0 \\ -5\sqrt{2} \\ 10 \end{pmatrix} \approx \begin{pmatrix} 0 \\ -7.07 \\ 10 \end{pmatrix}.$$

The total energies of the systems should be equal. In order to verify the equivalence of the two particle systems, we approximated the total energies via

$$U \approx \frac{1}{2} \sum_{i,j=1}^{N} \sum_{\substack{\boldsymbol{n} \in \mathbb{Z} \times \{0\}^2 \\ -C \le n_1 \le C}}' \frac{\xi_i \xi_j}{\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|} =: U_{\approx}, \tag{2.62}$$

i.e., we computed the slowly converging series via truncation using some cutoff $C \in \mathbb{N}$. Indeed, the two energies approach the same value for growing cutoff parameter C, see Figure 2.5. □

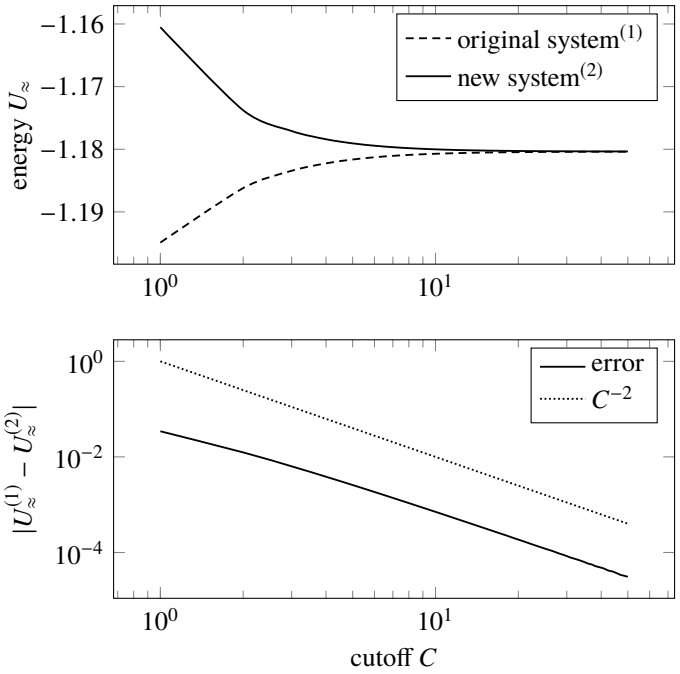


Figure 2.5: The electrostatic energy U has been approximated via truncation for growing cutoff $C \in \mathbb{N}$, see (2.62). The same value is approached for the original system$^{(1)}$ and the computed equivalent system$^{(2)}$ with rate C^{-2}.

2.2.4 0d-periodic boundary conditions

In the case of open boundary conditions, i.e., periodicity is not assumed in any direction, the potentials of the particles are given by

$$\phi(j) = \sum_{\substack{i=1 \\ i \neq j}}^{N} \frac{\xi_i}{\|\boldsymbol{x}_{ij}\|}. \tag{2.63}$$

The sums, which have to be evaluated are finite but the computation of the potentials and forces nevertheless scales like $O(N^2)$, which is unfavorable in terms of numerical efficiency.

In order to develop fast summation algorithms a separation into a short range part and a long range part is nevertheless profitable, see Section 4.1.2.

We obtain

$$\phi(j) = \phi^{\text{short}}(j) + \phi^{\text{long}}(j) - \phi^{\text{self}}(j),$$

where

$$\phi^{\text{short}}(j) = \sum_{\substack{i=1 \\ i\neq j}}^{N} \xi_i \frac{\operatorname{erfc}(\alpha\|\boldsymbol{x}_{ij}\|)}{\|\boldsymbol{x}_{ij}\|}, \tag{2.64}$$

$$\phi^{\text{long}}(j) = \sum_{i=1}^{N} \xi_i \frac{\operatorname{erf}(\alpha\|\boldsymbol{x}_{ij}\|)}{\|\boldsymbol{x}_{ij}\|}. \tag{2.65}$$

The self potential is defined via (2.24). Note that the short range and the long range part of the potentials are obtained from (2.22) and (2.23), respectively, by setting $\mathcal{S} := \{0\}^3$.

2.2.5 Summary

The Ewald summation formulas for the different types of periodic boundary conditions, see Lemmas 2.5, 2.8 and 2.12 are unified in Theorem 2.16. In order to include the case of open boundary conditions, we set $\mathcal{X}^0 := \{0\}$ for $\mathcal{X} \in \{\mathbb{R}, \mathbb{T}, \mathbb{Z}\}$ and $\mathbb{R}^{0\times 0} := \{\mathrm{Id}\}$.

Theorem 2.16. *Suppose that a particle system, fulfilling the charge neutrality condition* (2.5)*, is given. Let further the first $\rho \in \{0, 1, 2, 3\}$ coordinates be affected by periodic boundary conditions, i.e., the particle positions may be regarded as elements of the space*

$$\boldsymbol{L}\mathbb{T}^\rho \times \mathbb{R}^{3-\rho} \text{ with a regular matrix } \boldsymbol{L} \in \mathbb{R}^{\rho\times\rho}.$$

If we write the particle positions $\boldsymbol{x}_j$ in the form

$$\boldsymbol{x}_j = \left(\boldsymbol{x}_j^{\mathrm{p}}, \boldsymbol{x}_j^{\mathrm{np}}\right)^\top \text{ with } \boldsymbol{x}_j^{\mathrm{p}} \in \boldsymbol{L}\mathbb{T}^\rho \text{ and } \boldsymbol{x}_j^{\mathrm{np}} \in \mathbb{R}^{3-\rho},$$

we obtain for the spherical limit

$$\phi(j) := \lim_{C\to\infty} \sum_{\boldsymbol{n}\in\mathbb{Z}^\rho : \|\boldsymbol{L}\boldsymbol{n}\|\le C} \sum_{i=1}^{N} \frac{\xi_i}{\|(\boldsymbol{x}_{ij}^{\mathrm{p}} + \boldsymbol{L}\boldsymbol{n}, \boldsymbol{x}_{ij}^{\mathrm{np}})^\top\|} = \phi^{\text{short}}(j) + \phi^{\mathrm{F}}(j) + \phi^{\boldsymbol{0}}(j) - \phi^{\text{self}}(j).$$

The short range part of the potential is defined via

$$\phi^{\text{short}}(j) := {\sum_{\boldsymbol{n}\in\mathbb{Z}^\rho}}' \sum_{i=1}^{N} \xi_i \frac{\operatorname{erfc}(\alpha\|(\boldsymbol{x}_{ij}^{\mathrm{p}} + \boldsymbol{L}\boldsymbol{n}, \boldsymbol{x}_{ij}^{\mathrm{np}})^\top\|)}{\|(\boldsymbol{x}_{ij}^{\mathrm{p}} + \boldsymbol{L}\boldsymbol{n}, \boldsymbol{x}_{ij}^{\mathrm{np}})^\top\|}.$$

Furthermore, the long range or rather Fourier space part is given by

$$\phi^{\mathrm{F}}(j) := \frac{1}{|\det(\boldsymbol{L})|} \sum_{\boldsymbol{k}\in\mathbb{Z}^\rho} \sum_{i=1}^{N} \xi_i \, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}^{\mathrm{p}}} \Theta(\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|, \|\boldsymbol{x}_{ij}^{\mathrm{np}}\|),$$

where we set $\det(\boldsymbol{L}) = \det(\mathrm{Id}) := 1$ *for $\rho = 0$, and the function Θ is defined via*

$$\Theta(k, r) := 2\pi^{\frac{\rho-1}{2}} \int_0^\alpha \frac{\mathrm{e}^{-\pi^2k^2/z^2 - z^2r^2}}{z^\rho} \mathrm{d}z \quad \text{for } k \neq 0, \rho > 0. \tag{2.66}$$

Furthermore, the $\boldsymbol{k} = \mathbf{0}$ *contributions are described by*

$$\Theta(0,r) := \begin{cases} 2\pi^{\frac{\rho-1}{2}} \displaystyle\sum_{n=\min\{\rho,1\}}^{\infty} \frac{(-1)^n r^{2n} \alpha^{2n-\rho+1}}{n!(2n-\rho+1)} & : \rho \in \{0,1,2\}, \\ 0 & : \textit{else}, \end{cases} \tag{2.67}$$

and

$$\phi^{\mathbf{0}}(j) := \begin{cases} -\dfrac{2\pi}{3|\det(\boldsymbol{L})|} \displaystyle\sum_{i=1}^{N} \xi_i \|\boldsymbol{x}_{ij}\|^2 & : \rho = 3, \\ 0 & : \textit{else}. \end{cases}$$

The self potential $\phi^{\mathrm{self}}(j)$ *is defined via* (2.24)*. For open boundary conditions, where* $\rho = 0$*, the charge neutrality condition* (2.5) *is not necessary.*

Proof. We refer to the Lemmas 2.5, 2.8 and 2.12 as well as the corresponding proofs. For 3d-periodic boundary conditions, i.e., $\rho = 3$, the integrals (2.66) are given by

$$\Theta(k,r) = 2\pi \int_0^{\alpha} \frac{\mathrm{e}^{-\pi^2 k^2/z^2}}{z^3} \mathrm{d}z = \frac{\hat{\psi}(\boldsymbol{k})}{\pi},$$

cf. (2.34). We end up with the representation based on the analytic Fourier coefficients, cf. (2.30) and (2.31). The volume of the primary simulation box is give by $V = |\det(\boldsymbol{L})|$.

For $\rho = 2$ we obtain

$$2\sqrt{\pi} \int_0^{\alpha} \frac{\mathrm{e}^{-\pi^2 k^2/z^2 - z^2 r^2}}{z^2} \mathrm{d}z = \Theta^{\mathrm{p2}}(k,r),$$

see (2.41). The Taylor series (2.67) is included in (2.43), where we have $\|\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2\| = |\det(\boldsymbol{L})|$ due to the special structure of the vectors $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$, cf. (2.37).

In the case $\rho = 1$ we get

$$2 \int_0^{\alpha} \frac{\mathrm{e}^{-\pi^2 k^2/z^2 - z^2 r^2}}{z} \mathrm{d}z = \Theta^{\mathrm{p1}}(k,r),$$

see (2.55). The Taylor series (2.67) has already been used in order to represent the contributions for $k = 0$, cf. (2.56). Here, we have $|l_{11}| = |\det(\boldsymbol{L})|$, cf. (2.51).

In case of open boundary conditions or rather $\rho = 0$ we refer to (2.64) and (2.65). The Taylor series (2.67) simply represents the function $r^{-1}\mathrm{erf}(\alpha r)$. Furthermore, we obtain

$$\frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n r^{2n} \alpha^{2n+1}}{n!(2n+1)} = \frac{\mathrm{erf}(\alpha r)}{r} = \frac{2}{\sqrt{\pi}} \int_0^{\alpha} \mathrm{e}^{-r^2 z^2} \mathrm{d}z,$$

which is in addition consistent with (2.66), where $\rho = 0$ and $k = 0$. ∎

2.3 Truncation errors in the Ewald summation formulas

In the field of molecular dynamics simulations the root mean square (rms) error in the forces is a very common and adequate measure of accuracy.

We will consider the rms force error of the charges

$$\Delta F_{\mathrm{c}} := \sqrt{\frac{1}{N_{\mathrm{c}}} \sum_{j=1}^{N_{\mathrm{c}}} \|\boldsymbol{F}(j) - \boldsymbol{F}_{\approx}(j)\|^2} \tag{2.68}$$

and the rms force error of the dipole particles

$$\Delta F_{\mathrm{d}} := \sqrt{\frac{1}{N_{\mathrm{d}}} \sum_{j=N_{\mathrm{c}}+1}^{N} \|\boldsymbol{F}(j) - \boldsymbol{F}_{\approx}(j)\|^2} \tag{2.69}$$

separately. Note that in (2.68) and (2.69) we denote by $\boldsymbol{F}(j)$ the exact forces and by $\boldsymbol{F}_{\approx}(j)$ certain approximations of the forces. For the overall rms force error we obtain

$$\Delta F^2 := \frac{1}{N} \sum_{j=1}^{N} \|\boldsymbol{F}(j) - \boldsymbol{F}_{\approx}(j)\|^2 = \frac{1}{N} \sum_{j=1}^{N_{\mathrm{c}}} \|\boldsymbol{F}(j) - \boldsymbol{F}_{\approx}(j)\|^2 + \frac{1}{N} \sum_{j=N_{\mathrm{c}}+1}^{N} \|\boldsymbol{F}(j) - \boldsymbol{F}_{\approx}(j)\|^2$$
$$\Rightarrow \Delta F^2 \leq \Delta F_{\mathrm{c}}^2 + \Delta F_{\mathrm{d}}^2 \iff \Delta F \leq \sqrt{\Delta F_{\mathrm{c}}^2 + \Delta F_{\mathrm{d}}^2}.$$

In order to develop parameter tuning strategies one is interested in reliable a priori estimates of the resulting rms errors (2.68) and (2.69). This is especially reasonable in case of uncorrelated particle systems. In the following sections we discuss the estimation of the considered rms errors for the case of 3d-periodic boundary conditions, which is best analytically tractable.

2.3.1 Common approach

In this section we introduce a common approach to the estimation of rms errors, which has already been successfully applied, see [DH98b, WH01] or also [KP92, Appendix A], for instance.

Let two sets of particles located at positions $\mathcal{X}_I := \{\boldsymbol{x}_i \in \mathbb{R}^3, i \in I\}$ and $\mathcal{X}_J := \{\boldsymbol{x}_j \in \mathbb{R}^3, j \in J\}$, where I and J are some index sets, be given. For each $j \in J$ we consider a vector valued error $\boldsymbol{\varepsilon}_j$ of the form

$$\boldsymbol{\varepsilon}_j := \eta_j \sum_{i \in I \setminus \{j\}} \eta_i \boldsymbol{\chi}_{ij},$$

where η_j are real numbers and the vectors $\boldsymbol{\chi}_{ij} \in \mathbb{C}^3$ may only depend, respectively, on $\boldsymbol{x}_{ij}$. We remark that especially $I \cap J \neq \emptyset$ or even $I = J$ is allowed. Note that the technical assumption $i \in I \setminus \{j\}$ or rather $i \neq j$ ensures that no systematic contributions with $\boldsymbol{x}_{ij} = \boldsymbol{0}$, which are independent from the particle positions at all, are present.

We aim to estimate the rms error

$$\sqrt{\frac{1}{|J|} \sum_{j \in J} \|\boldsymbol{\varepsilon}_j\|^2}.$$

If we assume that the contributions from different particles are uncorrelated, we have

$$\left\langle \overline{\boldsymbol{\chi}_{ik}}^{\top} \boldsymbol{\chi}_{kj} \right\rangle = \delta_{ij} \left\langle \overline{\boldsymbol{\chi}_{ij}}^{\top} \boldsymbol{\chi}_{ij} \right\rangle =: \delta_{ij} \chi^2, \tag{2.70}$$

where the angular brackets denote that the average over all possible configurations is considered and δ_{ij} is the well-known Kronecker delta. The assumption of a homogeneous particle distribution is

not always reasonable, but the estimation nevertheless enables a good prediction of the resulting approximation errors in many cases. We further obtain

$$\left\langle \|\boldsymbol{\varepsilon}_j\|^2 \right\rangle = \eta_j^2 \sum_{i\in I\setminus\{j\}} \sum_{k\in I\setminus\{j\}} \eta_i\eta_k \left\langle \overline{\boldsymbol{\chi}_{ij}}^{\top} \boldsymbol{\chi}_{jk} \right\rangle \approx \eta_j^2 \chi^2 \sum_{i\in I\setminus\{j\}} \eta_i^2 \le \eta_j^2 \chi^2 \sum_{i\in I} \eta_i^2$$

and

$$\sqrt{\frac{1}{|J|}\sum_{j\in J} \|\boldsymbol{\varepsilon}_j\|^2} \approx \chi \sqrt{\frac{1}{|J|}\sum_{i\in I} \eta_i^2 \sum_{j\in J} \eta_j^2}. \tag{2.71}$$

In the next sections we will use the symbols

$$Q := \sum_{j=1}^{N_c} q_j^2$$

and

$$\mathcal{M} := \sum_{j=N_c+1}^{N} \|\boldsymbol{\mu}_j\|^2.$$

2.3.2 Rms force errors in the short range part

For 3d-periodic boundary conditions the short range parts of the potentials are defined by (2.22), where we have to set $\mathcal{S} := \mathbb{Z}^3$. Based on this, the short range parts of the forces are given by

$$\boldsymbol{F}^{\text{short}}(j) = \begin{cases} -q_j \nabla_{\boldsymbol{x}_j} \phi^{\text{short}}(j) & : j = 1, \dots, N_c, \\ -\nabla_{\boldsymbol{x}_j} (\boldsymbol{\mu}_j^{\top} \nabla_{\boldsymbol{x}_j}) \phi^{\text{short}}(j) & : j = N_c + 1, \dots, N, \end{cases}$$

see (2.14).

The rms errors in the short range parts of the forces are studied in the following. As mentioned before, we consider the charges and the dipole particles separately.

In order to keep the notation short, we denote the short ranged kernel function by

$$s_\alpha(r) := \frac{\operatorname{erfc}(\alpha r)}{r}$$

and introduce the abbreviations

$$\boldsymbol{x}_{ij,\boldsymbol{n}} := \boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n} \quad \text{and} \quad r_{ij,\boldsymbol{n}} := \|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|.$$

We will make use of the asymptotic expansion formula, as introduced by the following lemma, see also [KP92].

Lemma 2.17. *Let $B > 0$ and $f(x)\,\mathrm{e}^{-Bx^2}$ be an integrable function on $[A, \infty)$. If*

$$\left| \frac{\mathrm{d}}{\mathrm{d}x} \frac{f(x)}{2Bx} \right| \ll |f(x)| \quad \forall\, x \ge A$$

then

$$\int_A^\infty f(x)\,\mathrm{e}^{-Bx^2} \mathrm{d}x \approx \frac{f(A)}{2AB}\,\mathrm{e}^{-BA^2}. \tag{2.72}$$

Proof. The result is easily seen by partial integration, since

$$\int_A^\infty \frac{f(x)}{x} \cdot x\, \mathrm{e}^{-Bx^2} \mathrm{d}x = \left[-\frac{f(x)}{2Bx}\, \mathrm{e}^{-Bx^2} \right]_A^\infty + \frac{1}{2B} \int_A^\infty \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{f(x)}{x} \right) \mathrm{e}^{-Bx^2} \mathrm{d}x$$
$$\iff \int_A^\infty \left(f(x) - \frac{\mathrm{d}}{\mathrm{d}x} \frac{f(x)}{2Bx} \right) \mathrm{e}^{-Bx^2} \mathrm{d}x = \frac{f(A)}{2AB}\, \mathrm{e}^{-BA^2}.$$

■

In the special case that $f(x) \equiv 2\pi^{-1/2}$ and $B = 1$ we obtain the asymptotic expansion formula for the complementary error function, as introduced in (2.21).

Short range errors for the charges

For the charges, i.e., for $j = 1, \ldots, N_\mathrm{c}$, we rewrite $\boldsymbol{F}^{\mathrm{short}}(j)$ via

$$\boldsymbol{F}^{\mathrm{short}}(j) = \boldsymbol{F}^{\mathrm{short}}_{\mathrm{c.c.}}(j) + \boldsymbol{F}^{\mathrm{short}}_{\mathrm{c.d.}}(j),$$

where

$$\begin{aligned}
\boldsymbol{F}^{\mathrm{short}}_{\mathrm{c.c.}}(j) &:= -q_j \nabla_{\boldsymbol{x}_j} {\sum_{\boldsymbol{n}\in\mathcal{S}}}' \sum_{i=1}^{N_\mathrm{c}} q_i s_\alpha(r_{ij,\boldsymbol{n}}) = q_j {\sum_{\boldsymbol{n}\in\mathcal{S}}}' \sum_{i=1}^{N_\mathrm{c}} q_i s'_\alpha(r_{ij,\boldsymbol{n}}) \frac{\boldsymbol{x}_{ij,\boldsymbol{n}}}{r_{ij,\boldsymbol{n}}}, \\
\boldsymbol{F}^{\mathrm{short}}_{\mathrm{c.d.}}(j) &:= -q_j \nabla_{\boldsymbol{x}_j} \sum_{\boldsymbol{n}\in\mathcal{S}} \sum_{j=N_\mathrm{c}+1}^{N} \boldsymbol{\mu}_i^\top \nabla_{\boldsymbol{x}_i} s_\alpha(r_{ij,\boldsymbol{n}}) \\
&= q_j \sum_{\boldsymbol{n}\in\mathcal{S}} \sum_{j=N_\mathrm{c}+1}^{N} s''_\alpha(r_{ij,\boldsymbol{n}}) \boldsymbol{\mu}_i^\top \boldsymbol{x}_{ij,\boldsymbol{n}} \frac{\boldsymbol{x}_{ij,\boldsymbol{n}}}{r_{ij,\boldsymbol{n}}^2} + s'_\alpha(r_{ij,\boldsymbol{n}}) \left(\frac{\boldsymbol{\mu}_i}{r_{ij,\boldsymbol{n}}} - \boldsymbol{\mu}_i^\top \boldsymbol{x}_{ij,\boldsymbol{n}} \frac{\boldsymbol{x}_{ij,\boldsymbol{n}}}{r_{ij,\boldsymbol{n}}^3} \right).
\end{aligned} \tag{2.73}$$

Since the function s_α and its derivatives tend to zero exponentially fast, we can truncate the infinite sum. In other words, for some sufficiently large cutoff radius $r_\mathrm{cut} > 0$ we may only consider the terms with $\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\| \le r_\mathrm{cut}$ and compute the finite sums

$$\boldsymbol{F}^{\mathrm{short}}_{\mathrm{c.c.}}(j) \approx \boldsymbol{F}^{\mathrm{short}}_{\mathrm{c.c.}}(j)_\mathrm{t} := -q_j \nabla_{\boldsymbol{x}_j} \sum_{i=1}^{N_\mathrm{c}} q_i \sum_{\substack{\boldsymbol{n}\in\mathcal{S} \\ 0<r_{ij,\boldsymbol{n}}\le r_\mathrm{cut}}} s_\alpha(r_{ij,\boldsymbol{n}}),$$

$$\boldsymbol{F}^{\mathrm{short}}_{\mathrm{c.d.}}(j) \approx \boldsymbol{F}^{\mathrm{short}}_{\mathrm{c.d.}}(j)_\mathrm{t} := -q_j \nabla_{\boldsymbol{x}_j} \sum_{i=N_\mathrm{c}+1}^{N} \boldsymbol{\mu}_i^\top \sum_{\substack{\boldsymbol{n}\in\mathcal{S} \\ 0<r_{ij,\boldsymbol{n}}\le r_\mathrm{cut}}} \nabla_{\boldsymbol{x}_i} s_\alpha(r_{ij,\boldsymbol{n}}).$$

In the following we denote by $\Delta F^{\mathrm{short}}_{\mathrm{c.c.}}$ and $\Delta F^{\mathrm{short}}_{\mathrm{c.d.}}$ the corresponding rms force errors. We present estimations of these errors for uncorrelated systems in Theorem 2.18.

Theorem 2.18. *If the cutoff parameter $r_\mathrm{cut} > 0$ is chosen sufficiently large and the particles are sufficiently homogeneously distributed, the rms force errors in the short range parts of the charges can*

be approximated as follows, cf. [*KP*92, *HNP*17].

$$\Delta F_{\text{c.c.}}^{\text{short}} := \sqrt{\frac{1}{N_{\text{c}}}\sum_{j=1}^{N_{\text{c}}} \|\boldsymbol{F}_{\text{c.c.}}^{\text{short}}(j) - \boldsymbol{F}_{\text{c.c.}}^{\text{short}}(j)_{\text{t}}\|^2} \approx \frac{Q\sqrt{r_{\text{cut}}}}{\sqrt{N_{\text{c}}V}\alpha}\, \mathrm{e}^{-\alpha^2 r_{\text{cut}}^2}\left(\frac{2\alpha}{\sqrt{\pi}r_{\text{cut}}} + \frac{1}{\sqrt{\pi}\alpha r_{\text{cut}}^3}\right), \tag{2.74}$$

$$\begin{aligned}\Delta F_{\text{c.d.}}^{\text{short}} &:= \sqrt{\frac{1}{N_{\text{c}}}\sum_{j=1}^{N_{\text{c}}} \|\boldsymbol{F}_{\text{c.d.}}^{\text{short}}(j) - \boldsymbol{F}_{\text{c.d.}}^{\text{short}}(j)_{\text{t}}\|^2} \\ &\approx \frac{\mathrm{e}^{-\alpha^2 r_{\text{cut}}^2}}{\alpha^2 r_{\text{cut}}^3}\sqrt{\frac{2Q\mathcal{M}}{3N_{\text{c}}V r_{\text{cut}}}\left(3B_{\alpha,\text{cut}}^2 + 8\alpha^4 r_{\text{cut}}^4 B_{\alpha,\text{cut}} + 8\alpha^8 r_{\text{cut}}^8\right)},\end{aligned} \tag{2.75}$$

where we set $B_{\alpha,\text{cut}} := 1 + 2\alpha^2 r_{\text{cut}}^2$.

Proof. The rms force error in the short range part has already been considered in [KP92], where only the charge-charge case is treated. For the sake of completeness, we present a short summary of the calculations in the following.

Using the notation as introduced above, we obtain

$$\boldsymbol{F}_{\text{c.c.}}^{\text{short}}(j) - \boldsymbol{F}_{\text{c.c.}}^{\text{short}}(j)_{\text{t}} = q_j \sum_{\substack{i=1\\ i\neq j}}^{N_{\text{c}}} q_i \boldsymbol{\chi}_{ij}^{\text{c.c.}}$$

with

$$\boldsymbol{\chi}_{ij}^{\text{c.c.}} = \sum_{\substack{\boldsymbol{n}\in\mathcal{S}\\ r_{ij,\boldsymbol{n}} > r_{\text{cut}}}} s'_\alpha(r_{ij,\boldsymbol{n}})\frac{\boldsymbol{x}_{ij,\boldsymbol{n}}}{r_{ij,\boldsymbol{n}}}.$$

It can be seen easily that $\boldsymbol{\chi}_{jj}^{\text{c.c.}} = \boldsymbol{0}$, since $r^{-1}s'_\alpha(r)$ is an even function. The derivative s'_α is given by

$$\begin{aligned}s'_\alpha(r) &= -\frac{2\alpha}{\sqrt{\pi}r}\,\mathrm{e}^{-\alpha^2 r^2} - \frac{\operatorname{erfc}(\alpha r)}{r^2} \\ &\approx -\frac{2\alpha}{\sqrt{\pi}r}\,\mathrm{e}^{-\alpha^2 r^2} - \frac{1}{\sqrt{\pi}\alpha r^3}\,\mathrm{e}^{-\alpha^2 r^2} \quad \text{for } r \text{ large enough.}\end{aligned}$$

The quadratic mean $\chi_{\text{c.c.}}^2$, as introduced by (2.70), is approximated as follows. As indicated by (2.70), it is defined as the average value of $\|\boldsymbol{\chi}_{ij}^{\text{c.c.}}\|^2$ over all possible configurations, i.e., we set

$$\chi_{\text{c.c.}}^2 := \frac{1}{V^2}\int_{\boldsymbol{L}\mathbb{T}^3}\int_{\boldsymbol{L}\mathbb{T}^3} \|\boldsymbol{\chi}_{ij}^{\text{c.c.}}\|^2 \,\mathrm{d}\boldsymbol{x}_i\,\mathrm{d}\boldsymbol{x}_j.$$

Here, the integration over $\boldsymbol{L}\mathbb{T}^3$ is to be understood as the integration over the representatives in the set $\{\lambda_1\boldsymbol{\ell}_1 + \lambda_2\boldsymbol{\ell}_2 + \lambda_3\boldsymbol{\ell}_3 : \lambda_j \in [-1/2, 1/2), j = 1,2,3\}$. The term $\|\boldsymbol{\chi}_{ij}^{\text{c.c.}}\|^2$ is given by

$$\|\boldsymbol{\chi}_{ij}^{\text{c.c.}}\|^2 = \sum_{\substack{\boldsymbol{n}\in\mathbb{Z}^3\\ r_{ij,\boldsymbol{n}} > r_{\text{cut}}}} |s'_\alpha(r_{ij,\boldsymbol{n}})|^2 + 2\sum_{\substack{\boldsymbol{n}\in\mathbb{Z}^3\\ r_{ij,\boldsymbol{n}} > r_{\text{cut}}}}\sum_{\substack{\boldsymbol{m}\in\mathbb{Z}^3\setminus\{\boldsymbol{n}\}\\ r_{ij,\boldsymbol{m}} > r_{\text{cut}}}} \frac{s'_\alpha(r_{ij,\boldsymbol{n}})s'_\alpha(r_{ij,\boldsymbol{m}})}{r_{ij,\boldsymbol{n}}r_{ij,\boldsymbol{m}}}\boldsymbol{x}_{ij,\boldsymbol{n}}^\top\boldsymbol{x}_{ij,\boldsymbol{m}},$$

where we separated the diagonal and the off-diagonal terms.

The double sum includes summands of alternating signs and if the particles are sufficiently homogeneously distributed these contributions may be neglected. Thus, we compute

$$\begin{aligned}\chi^2_{\text{c.c.}} &\approx \frac{1}{V^2}\int_{\boldsymbol{L}\mathbb{T}^3}\int_{\boldsymbol{L}\mathbb{T}^3}\sum_{\substack{\boldsymbol{n}\in\mathbb{Z}^3\\ r_{ij,\boldsymbol{n}}>r_{\text{cut}}}} |s'_\alpha(r_{ij,\boldsymbol{n}})|^2\,\mathrm{d}\boldsymbol{x}_i\,\mathrm{d}\boldsymbol{x}_j\\ &= \frac{1}{V^2}\int_{\boldsymbol{L}\mathbb{T}^3}\int_{\|\boldsymbol{x}-\boldsymbol{y}\|>r_{\text{cut}}} |s'_\alpha(\|\boldsymbol{x}-\boldsymbol{y}\|)|^2\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} = \frac{1}{V}\int_{\|\boldsymbol{r}\|>r_{\text{cut}}} |s'_\alpha(\|\boldsymbol{r}\|)|^2\,\mathrm{d}\boldsymbol{r},\end{aligned}$$

cf. [KP92, Eq. (17)]. Finally, we obtain

$$\begin{aligned}\chi^2_{\text{c.c.}} &\approx \frac{1}{V}\int_0^\pi\int_0^{2\pi}\int_{r_{\text{cut}}}^\infty \mathrm{e}^{-2\alpha^2r^2}\left(\frac{2\alpha}{\sqrt{\pi}r}+\frac{1}{\sqrt{\pi}\alpha r^3}\right)^2 r^2\sin\theta\,\mathrm{d}r\,\mathrm{d}\phi\,\mathrm{d}\theta\\ &= \frac{4\pi}{V}\int_{r_{\text{cut}}}^\infty \mathrm{e}^{-2\alpha^2r^2}\left(\frac{2\alpha}{\sqrt{\pi}r}+\frac{1}{\sqrt{\pi}\alpha r^3}\right)^2 r^2\mathrm{d}r\\ &\approx \frac{r_{\text{cut}}}{V\alpha^2}\,\mathrm{e}^{-2\alpha^2r_{\text{cut}}^2}\left(\frac{2\alpha}{\sqrt{\pi}r_{\text{cut}}}+\frac{1}{\sqrt{\pi}\alpha r_{\text{cut}}^3}\right)^2,\end{aligned}$$

where we use the asymptotic expansion formula (2.72). By applying (2.71) we obtain (2.74).

We continue with the charge-dipole interactions and consider Equation (2.73). Based on this representation, we start by considering an expression of the form

$$s''_\alpha(r)\boldsymbol{\mu}_i^\top\boldsymbol{r}\frac{\boldsymbol{r}}{r^2}+s'_\alpha(r)\left(\frac{\boldsymbol{\mu}_i}{r}-\boldsymbol{\mu}_i^\top\boldsymbol{r}\frac{\boldsymbol{r}}{r^3}\right) = \|\boldsymbol{\mu}_i\|\left[s''_\alpha(r)r\cos\theta\frac{\boldsymbol{r}}{r^2}+s'_\alpha(r)\left(\frac{\tilde{\boldsymbol{\mu}}_i}{r}-r\cos\theta\frac{\boldsymbol{r}}{r^3}\right)\right],$$

where we denote by $\theta := \sphericalangle(\boldsymbol{\mu}_i,\boldsymbol{r})$ the angle between a vector $\boldsymbol{r}$ and the dipole moment $\boldsymbol{\mu}_i$. Furthermore, we set $\tilde{\boldsymbol{\mu}}_i := \|\boldsymbol{\mu}_i\|^{-1}\boldsymbol{\mu}_i$.

Thus, we can write

$$\boldsymbol{F}^{\text{short}}_{\text{c.d.}}(j)-\boldsymbol{F}^{\text{short}}_{\text{c.d.}}(j)_{\text{t}} = q_j\sum_{i=N_{\text{c}}+1}^{N}\|\boldsymbol{\mu}_i\|\boldsymbol{\chi}^{\text{c.d.}}_{ij},$$

where

$$\boldsymbol{\chi}^{\text{c.d.}}_{ij} = \sum_{\substack{\boldsymbol{n}\in\mathcal{S}\\ r_{ij,\boldsymbol{n}}>r_{\text{cut}}}} s''_\alpha(r_{ij,\boldsymbol{n}})\cos(\theta_{ij,\boldsymbol{n}})\frac{\boldsymbol{x}_{ij,\boldsymbol{n}}}{r_{ij,\boldsymbol{n}}}+s'_\alpha(r_{ij,\boldsymbol{n}})\left(\frac{\tilde{\boldsymbol{\mu}}_i}{r_{ij,\boldsymbol{n}}}-\cos(\theta_{ij,\boldsymbol{n}})\frac{\boldsymbol{x}_{ij,\boldsymbol{n}}}{r^2_{ij,\boldsymbol{n}}}\right). \tag{2.76}$$

Here, we define $\theta_{ij,\boldsymbol{n}} := \sphericalangle(\boldsymbol{\mu}_i,\boldsymbol{x}_{ij,\boldsymbol{n}})$. We compute the squared Euclidean norm of the single summands, which reads as

$$\begin{aligned}&\left\|s''_\alpha(r)\cos(\theta)\frac{\boldsymbol{r}}{r}+s'_\alpha(r)\left(\frac{\tilde{\boldsymbol{\mu}}_i}{r}-\cos(\theta)\frac{\boldsymbol{r}}{r^2}\right)\right\|^2\\ &= \left(\tfrac{s''_\alpha(r)}{r}-\tfrac{s'_\alpha(r)}{r^2}\right)^2 r^2\cos^2\theta+\tfrac{s'_\alpha(r)^2}{r^2}+2\left(\tfrac{s''_\alpha(r)}{r}-\tfrac{s'_\alpha(r)}{r^2}\right)\tfrac{s'_\alpha(r)}{r}\cos^2(\theta)r\\ &= \tfrac{s'_\alpha(r)^2}{r^2}+\cos^2\theta\left(s''_\alpha(r)^2-\tfrac{s'_\alpha(r)^2}{r^2}\right).\end{aligned}$$

Now, we proceed similarly as for the approximation of $\chi^2_{\text{c.c.}}$ and approximate the error sum by an integral using spherical coordinates. The polar coordinates are chosen such that $\boldsymbol{\mu}_i$ points from the center to the north pole, i.e. θ is the polar angle.

We obtain

$$\begin{aligned}\chi^2_{\text{c.d.}} &\approx \frac{2\pi}{V}\int_0^\pi \int_{r_{\text{cut}}}^\infty \left(\tfrac{s'_\alpha(r)^2}{r^2} + \cos^2\theta\left(s''_\alpha(r)^2 - \tfrac{s'_\alpha(r)^2}{r^2}\right)\right) r^2 \sin\theta\,\mathrm{d}r\,\mathrm{d}\theta \\ &= \frac{2\pi}{V}\int_{r_{\text{cut}}}^\infty \tfrac{4}{3}s'_\alpha(r)^2 + \tfrac{2}{3}r^2 s''_\alpha(r)^2 \mathrm{d}r,\end{aligned}$$

where differentiation and application of the asymptotic expansion formula (2.72) yields

$$s''_\alpha(r) = \frac{4\alpha^3}{\sqrt{\pi}}\,\mathrm{e}^{-\alpha^2 r^2} + \frac{4\alpha}{\sqrt{\pi}r^2}\,\mathrm{e}^{-\alpha^2 r^2} + 2\frac{\operatorname{erfc}(\alpha r)}{r^3} \approx \frac{\mathrm{e}^{-\alpha^2 r^2}}{\sqrt{\pi}\alpha r^4}\left(4\alpha^4 r^4 + 4\alpha^2 r^2 + 2\right).$$

The approximation holds if r is large enough. By making use of the asymptotic expansion formula (2.72) once again we obtain

$$\chi^2_{\text{c.d.}} \approx \frac{\mathrm{e}^{-2\alpha^2 r_{\text{cut}}^2}}{3V\alpha^4 r_{\text{cut}}^7}\left(2B^2_{\alpha,\text{cut}} + (4\alpha^4 r_{\text{cut}}^4 + 2B_{\alpha,\text{cut}})^2\right)$$

and

$$\Delta F^{\text{short}}_{\text{c.d.}} \approx \sqrt{\frac{Q\mathcal{M}\chi^2_{\text{c.d.}}}{N_{\text{c}}}}$$

gives (2.75). ■

Remark 2.19. A slightly simpler estimate of the error $\Delta F^{\text{short}}_{\text{c.c.}}$, as stated in (2.74), is given by

$$\Delta F^{\text{short}}_{\text{c.c.}} \approx \frac{2Q}{\sqrt{\pi N_{\text{c}} V r_{\text{cut}}}}\,\mathrm{e}^{-\alpha^2 r_{\text{cut}}^2}, \tag{2.77}$$

see [KP92]. It is obtained from (2.74) by neglecting the second term within the brackets. □

Short range errors for the dipole particles

We treat the dipole particles in an analogue manner and split the short range forces via

$$\boldsymbol{F}^{\text{short}}(j) = \boldsymbol{F}^{\text{short}}_{\text{d.c.}}(j) + \boldsymbol{F}^{\text{short}}_{\text{d.d.}}(j),$$

where $j = N_{\text{c}}+1,\dots,N$. We obtain

$$\begin{aligned}\boldsymbol{F}^{\text{short}}_{\text{d.c.}}(j) &:= -\nabla_{\boldsymbol{x}_j}\boldsymbol{\mu}_j^\top\nabla_{\boldsymbol{x}_j}\sum_{\boldsymbol{n}\in\mathcal{S}}\sum_{i=1}^{N_{\text{c}}} q_i s_\alpha(r_{ij,\boldsymbol{n}}) \\ &= -\sum_{\boldsymbol{n}\in\mathcal{S}}\sum_{i=1}^{N_{\text{c}}} q_i s''_\alpha(r_{ij,\boldsymbol{n}})\boldsymbol{\mu}_j^\top\boldsymbol{x}_{ij,\boldsymbol{n}}\frac{\boldsymbol{x}_{ij,\boldsymbol{n}}}{r^2_{ij,\boldsymbol{n}}} + s'_\alpha(r_{ij,\boldsymbol{n}})\left(\frac{\boldsymbol{\mu}_j}{r_{ij,\boldsymbol{n}}} - \boldsymbol{\mu}_j^\top\boldsymbol{x}_{ij,\boldsymbol{n}}\frac{\boldsymbol{x}_{ij,\boldsymbol{n}}}{r^3_{ij,\boldsymbol{n}}}\right),\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{F}^{\text{short}}_{\text{d.d.}}(j) &:= -\nabla_{\boldsymbol{x}_j}\boldsymbol{\mu}_j^\top\nabla_{\boldsymbol{x}_j}{\sum_{\boldsymbol{n}\in\mathcal{S}}}'\sum_{i=N_{\text{c}}+1}^{N}\boldsymbol{\mu}_i^\top\nabla_{\boldsymbol{x}_i}s_\alpha(r_{ij,\boldsymbol{n}}) \\ &= -{\sum_{\boldsymbol{n}\in\mathcal{S}}}'\sum_{i=N_{\text{c}}+1}^{N} C_\alpha(r_{ij,\boldsymbol{n}})\left((\boldsymbol{\mu}_i^\top\boldsymbol{x}_{ij,\boldsymbol{n}})\boldsymbol{\mu}_j + (\boldsymbol{\mu}_j^\top\boldsymbol{x}_{ij,\boldsymbol{n}})\boldsymbol{\mu}_i + (\boldsymbol{\mu}_i^\top\boldsymbol{\mu}_j)\boldsymbol{x}_{ij,\boldsymbol{n}}\right) \\ &\quad + {\sum_{\boldsymbol{n}\in\mathcal{S}}}'\sum_{i=N_{\text{c}}+1}^{N} D_\alpha(r_{ij,\boldsymbol{n}})(\boldsymbol{\mu}_i^\top\boldsymbol{x}_{ij,\boldsymbol{n}})(\boldsymbol{\mu}_j^\top\boldsymbol{x}_{ij,\boldsymbol{n}})\boldsymbol{x}_{ij,\boldsymbol{n}},\end{aligned}$$

where the functions C_α and D_α are given by

$$C_\alpha(r) := \frac{s''_\alpha(r)}{r^2} - \frac{s'_\alpha(r)}{r^3} = \frac{4\alpha^3}{\sqrt{\pi}r^2}\,\mathrm{e}^{-\alpha^2r^2} + \frac{6\alpha}{\sqrt{\pi}r^4}\,\mathrm{e}^{-\alpha^2r^2} + \frac{3\mathrm{erfc}(\alpha r)}{r^5},$$
$$D_\alpha(r) := -\frac{s'''_\alpha(r)}{r^3} + 3\frac{s''_\alpha(r)}{r^4} - 3\frac{s'_\alpha(r)}{r^2} = \left(\frac{8\alpha^5}{\sqrt{\pi}r^2} + \frac{20\alpha^3}{\sqrt{\pi}r^4} + \frac{30\alpha}{\sqrt{\pi}r^6}\right)\mathrm{e}^{-\alpha^2r^2} + \frac{15\mathrm{erfc}(\alpha r)}{r^7},$$

see [WH01].

In the following we denote by $\Delta F_{\mathrm{d.c.}}^{\mathrm{short}}$ and $\Delta F_{\mathrm{d.d.}}^{\mathrm{short}}$ the corresponding rms force errors. We present estimations of these errors for uncorrelated systems in Theorem 2.20.

Theorem 2.20. *If the cutoff parameter $r_{\mathrm{cut}} > 0$ is chosen sufficiently large and the particles are sufficiently homogeneously distributed, the rms force errors in the short range parts of the dipoles can be approximated as follows, cf.* [*WH*01*, HNP*17].

$$\begin{aligned}\Delta F_{\mathrm{d.c.}}^{\mathrm{short}} &:= \sqrt{\frac{1}{N_{\mathrm{d}}}\sum_{j=N_{\mathrm{c}}+1}^{N}\|\boldsymbol{F}_{\mathrm{d.c.}}^{\mathrm{short}}(j) - \boldsymbol{F}_{\mathrm{d.c.}}^{\mathrm{short}}(j)_{\mathrm{t}}\|^2}\\ &\approx \frac{\mathrm{e}^{-\alpha^2r_{\mathrm{cut}}^2}}{\alpha^2r_{\mathrm{cut}}^3}\sqrt{\frac{2Q\mathcal{M}}{3N_{\mathrm{d}}Vr_{\mathrm{cut}}}\left(3B_{\alpha,\mathrm{cut}}^2 + 8\alpha^4r_{\mathrm{cut}}^4B_{\alpha,\mathrm{cut}} + 8\alpha^8r_{\mathrm{cut}}^8\right)}, \qquad (2.78)\\ \Delta F_{\mathrm{d.d.}}^{\mathrm{short}} &:= \sqrt{\frac{1}{N_{\mathrm{d}}}\sum_{j=N_{\mathrm{c}}+1}^{N}\|\boldsymbol{F}_{\mathrm{d.d.}}^{\mathrm{short}}(j) - \boldsymbol{F}_{\mathrm{d.d.}}^{\mathrm{short}}(j)_{\mathrm{t}}\|^2}\\ &\approx \frac{\mathrm{e}^{-\alpha^2r_{\mathrm{cut}}^2}\mathcal{M}}{\alpha^2r_{\mathrm{cut}}^4}\sqrt{\frac{1}{r_{\mathrm{cut}}N_{\mathrm{d}}V}\left(\frac{13}{6}C_{\alpha,\mathrm{cut}}^2 + \frac{2}{15}D_{\alpha,\mathrm{cut}}^2 + \frac{13}{15}C_{\alpha,\mathrm{cut}}D_{\alpha,\mathrm{cut}}\right)}, \qquad (2.79)\end{aligned}$$

where we define

$$\begin{aligned}B_{\alpha,\mathrm{cut}} &:= 2\alpha^2r_{\mathrm{cut}}^2 + 1,\\ C_{\alpha,\mathrm{cut}} &:= 4\alpha^4r_{\mathrm{cut}}^4 + 6\alpha^2r_{\mathrm{cut}}^2 + 3,\\ D_{\alpha,\mathrm{cut}} &:= 8\alpha^6r_{\mathrm{cut}}^6 + 20\alpha^4r_{\mathrm{cut}}^4 + 30\alpha^2r_{\mathrm{cut}}^2 + 15.\end{aligned}$$

Proof. The approximation of the rms error $\Delta F_{\mathrm{d.c.}}^{\mathrm{short}}$ is obtained analogously to $\Delta F_{\mathrm{c.d.}}^{\mathrm{short}}$, see Theorem 2.18. We obtain

$$\boldsymbol{F}_{\mathrm{d.c.}}^{\mathrm{short}}(j) - \boldsymbol{F}_{\mathrm{d.c.}}^{\mathrm{short}}(j)_{\mathrm{t}} = \|\boldsymbol{\mu}_j\|\sum_{i=1}^{N_{\mathrm{c}}} q_i\chi_{ij}^{\mathrm{d.c.}}$$

with $\chi_{ij}^{\mathrm{d.c.}} = -\chi_{ji}^{\mathrm{c.d.}}$, see (2.76).

In order to approximate $\Delta F_{\mathrm{d.d.}}^{\mathrm{short}}$ we start by considering the vector field

$$\begin{aligned}\boldsymbol{v}(\boldsymbol{r},\boldsymbol{\mu}_i,\boldsymbol{\mu}_j) &:= C(r)\left((\boldsymbol{\mu}_i^\top\boldsymbol{r})\boldsymbol{\mu}_j + (\boldsymbol{\mu}_j^\top\boldsymbol{r})\boldsymbol{\mu}_i + (\boldsymbol{\mu}_i^\top\boldsymbol{\mu}_j)\boldsymbol{r}\right) - D(r)(\boldsymbol{\mu}_i^\top\boldsymbol{r})(\boldsymbol{\mu}_j^\top\boldsymbol{r})\boldsymbol{r}\\ &= \|\boldsymbol{\mu}_i\|\|\boldsymbol{\mu}_j\|\left[C(r)\left((\tilde{\boldsymbol{\mu}}_i^\top\boldsymbol{r})\tilde{\boldsymbol{\mu}}_j + (\tilde{\boldsymbol{\mu}}_j^\top\boldsymbol{r})\tilde{\boldsymbol{\mu}}_i + (\tilde{\boldsymbol{\mu}}_i^\top\tilde{\boldsymbol{\mu}}_j)\boldsymbol{r}\right) - D(r)(\tilde{\boldsymbol{\mu}}_i^\top\boldsymbol{r})(\tilde{\boldsymbol{\mu}}_j^\top\boldsymbol{r})\boldsymbol{r}\right]. \qquad (2.80)\end{aligned}$$

The spherical coordinates are now chosen such that $\boldsymbol{\mu}_j$ points from the center to the north pole, i.e., $\tilde{\boldsymbol{\mu}}_j \mapsto (0,0,1)$, and that the azimuthal angle of $\boldsymbol{\mu}_i$ is $\phi_{\boldsymbol{\mu}_i} = 0$, i.e., $\tilde{\boldsymbol{\mu}}_i \mapsto (\sin\gamma, 0, \cos\gamma)$, where we set $\gamma := \sphericalangle(\boldsymbol{\mu}_i,\boldsymbol{\mu}_j)$.

Thus, we obtain for some $\boldsymbol{r} = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)$

$$\cos\sphericalangle(\boldsymbol{\mu}_i, \boldsymbol{r}) = \frac{\boldsymbol{\mu}_i^\top \boldsymbol{r}}{\|\boldsymbol{\mu}_i\| r} = \cos\theta\cos\gamma + \sin\theta\sin\gamma\cos\phi, \tag{2.81}$$

see [WH01, Eq. (27)].

We see that the term in brackets in (2.80) can be interpreted as a function of r, θ, ϕ and γ, i.e.,

$$C(r)\left((\tilde{\boldsymbol{\mu}}_i^\top \boldsymbol{r})\tilde{\boldsymbol{\mu}}_j + (\tilde{\boldsymbol{\mu}}_j^\top \boldsymbol{r})\tilde{\boldsymbol{\mu}}_i + (\tilde{\boldsymbol{\mu}}_i^\top \tilde{\boldsymbol{\mu}}_j)\boldsymbol{r}\right) - D(r)(\tilde{\boldsymbol{\mu}}_i^\top \boldsymbol{r})(\tilde{\boldsymbol{\mu}}_j^\top \boldsymbol{r})\boldsymbol{r} =: \tilde{\boldsymbol{v}}(r, \theta, \phi, \gamma).$$

Now we compute $\chi^2_{\text{d.d.}}$ as follows.

$$\chi^2_{\text{d.d.}} := \frac{1}{\pi}\int_0^\pi \frac{1}{V}\int_0^{2\pi}\int_0^\pi \int_{r_{\text{cut}}}^\infty \|\tilde{\boldsymbol{v}}(r, \theta, \phi, \gamma)\|^2 r^2 \sin(\theta)\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi\,\mathrm{d}\gamma,$$

i.e., the term $\chi^2_{\text{d.d.}}$ is as usual approximated by an integral in spherical coordinates and in addition we compute the mean values of all terms depending on the angle γ.

Finally, we obtain

$$\chi^2_{\text{d.d.}} = \frac{\mathrm{e}^{-2\alpha^2 r_{\text{cut}}^2}}{V\alpha^4 r_{\text{cut}}^9}\left(\frac{13}{6}C^2_{\alpha,\text{cut}} + \frac{2}{15}D^2_{\alpha,\text{cut}} + \frac{13}{15}C_{\alpha,\text{cut}}D_{\alpha,\text{cut}}\right),$$

i.e.,

$$\Delta F^{\text{short}}_{\text{d.d.}} \approx \sqrt{\frac{\chi^2_{\text{d.d.}}}{N_{\text{d}}}\mathcal{M}^2},$$

which gives (2.79), see also [WH01]. ■

2.3.3 Rms force errors in the Fourier space part

The Fourier coefficients $\hat{\psi}(\boldsymbol{k})$, as defined in (2.31), tend to zero exponentially fast as $\|\boldsymbol{L}^{-\top}\boldsymbol{k}\| \to \infty$. In order to compute the Fourier space potentials efficiently, we truncate the infinite sum, i.e., we replace the set $\mathbb{Z}^3$ by an appropriate finite subset $\mathcal{K} \subset \mathbb{Z}^3$.

Remark 2.21 (Truncation scheme)**.** Since $\hat{\psi}(\boldsymbol{k})$ is a function of $\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|$, it is reasonable to consider index sets of the form

$$\mathcal{K} := \{\boldsymbol{k} \in \mathbb{Z}^3 : \|\boldsymbol{L}^{-\top}\boldsymbol{k}\| \le c\},$$

where $c > 0$ is some cutoff radius.

But, in view of the Fourier methods applied later, see Chapters 3 and 4, we will for now mainly take into account index sets of the form

$$\mathcal{K} := \mathcal{J}_{\boldsymbol{M}} \cap \mathbb{Z}^3,$$

where we set

$$\mathcal{J}_{\boldsymbol{M}} := (-M_1/2, M_1/2) \times (-M_2/2, M_2/2) \times (-M_3/2, M_3/2)$$

for some $\boldsymbol{M} = (M_1, M_2, M_3) \in \mathbb{R}^3_+$. □

In the following we denote by $\hat{\boldsymbol{\ell}}_1, \hat{\boldsymbol{\ell}}_2, \hat{\boldsymbol{\ell}}_3 \in \mathbb{R}^3$ the dual lattice vectors, which are simply the columns of the matrix $\boldsymbol{L}^{-\top}$, i.e., we have $\hat{\boldsymbol{\ell}}_i^\top \boldsymbol{\ell}_j = \delta_{ij}$ as well as

$$\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 = \|k_1\hat{\boldsymbol{\ell}}_1 + k_2\hat{\boldsymbol{\ell}}_2 + k_3\hat{\boldsymbol{\ell}}_3\|^2.$$

Lemma 2.22. *We denote by $\beta > 0$ the cutoff parameter of the Fourier space part. If M_1, M_2 and M_3 are chosen such that*

$$\frac{M_1}{\|\boldsymbol{\ell}_1\|} = \frac{M_2}{\|\boldsymbol{\ell}_2\|} = \frac{M_3}{\|\boldsymbol{\ell}_3\|} = \beta, \tag{2.82}$$

we have

$$\left\{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\| < \tfrac{\beta}{2}\right\} \subset \left\{x_1\hat{\boldsymbol{\ell}}_1 + x_2\hat{\boldsymbol{\ell}}_2 + x_3\hat{\boldsymbol{\ell}}_3, \boldsymbol{x} \in \mathcal{J}_{\boldsymbol{M}}\right\}. \tag{2.83}$$

Proof. In the following we set

$$\hat{\boldsymbol{n}}_{kl} := \hat{\boldsymbol{\ell}}_k \times \hat{\boldsymbol{\ell}}_l.$$

In order to ensure (2.83) we choose $M_1, M_2, M_3 > 0$ such that

$$M_1\|\hat{\boldsymbol{\ell}}_1\| \cdot |\cos \sphericalangle(\hat{\boldsymbol{\ell}}_1, \hat{\boldsymbol{n}}_{23})| = M_2\|\hat{\boldsymbol{\ell}}_2\| \cdot |\cos \sphericalangle(\hat{\boldsymbol{\ell}}_2, \hat{\boldsymbol{n}}_{13})| = M_3\|\hat{\boldsymbol{\ell}}_3\| \cdot |\cos \sphericalangle(\hat{\boldsymbol{\ell}}_3, \hat{\boldsymbol{n}}_{12})| = \beta, \tag{2.84}$$

which means that the three heights of the parallelepiped

$$\{x_1\hat{\boldsymbol{\ell}}_1 + x_2\hat{\boldsymbol{\ell}}_2 + x_3\hat{\boldsymbol{\ell}}_3 : \boldsymbol{x} \in \mathcal{J}_{\boldsymbol{M}}\}$$

are all of length β.

The height of the parallelepiped with respect to the direction $\hat{\boldsymbol{n}}_{23}$, for instance, is given by dividing the volume by the size of the base area spanned by $\hat{\boldsymbol{\ell}}_2$ and $\hat{\boldsymbol{\ell}}_3$, which gives

$$\frac{|(M_1\hat{\boldsymbol{\ell}}_1)^\top(M_2\hat{\boldsymbol{\ell}}_2 \times M_3\hat{\boldsymbol{\ell}}_3)|}{|M_2\hat{\boldsymbol{\ell}}_2 \times M_3\hat{\boldsymbol{\ell}}_3|} = M_1\|\hat{\boldsymbol{\ell}}_1\| \cdot |\cos \sphericalangle(\hat{\boldsymbol{\ell}}_1, \hat{\boldsymbol{n}}_{23})|.$$

Now it is obvious that the ball of radius $\beta/2$ is included in the considered set. It remains to show that the conditions (2.84) are equivalent to (2.82).

As an example, we consider the vectors $\hat{\boldsymbol{n}}_{23}$ and $\boldsymbol{\ell}_1$, which are linearly dependent. Thus, we have $\cos \sphericalangle(\hat{\boldsymbol{\ell}}_1, \hat{\boldsymbol{n}}_{23}) = \cos \sphericalangle(\hat{\boldsymbol{\ell}}_1, \boldsymbol{\ell}_1)$ and obtain

$$M_1\|\hat{\boldsymbol{\ell}}_1\| \cdot |\cos \sphericalangle(\hat{\boldsymbol{\ell}}_1, \hat{\boldsymbol{n}}_{23})| = M_1\|\hat{\boldsymbol{\ell}}_1\| \cdot |\cos \sphericalangle(\hat{\boldsymbol{\ell}}_1, \boldsymbol{\ell}_1)| = M_1\|\hat{\boldsymbol{\ell}}_1\| \cdot \frac{|\hat{\boldsymbol{\ell}}_1^\top \boldsymbol{\ell}_1|}{\|\hat{\boldsymbol{\ell}}_1\| \cdot \|\boldsymbol{\ell}_1\|} = \frac{M_1}{\|\boldsymbol{\ell}_1\|},$$

since $\hat{\boldsymbol{\ell}}_1^\top \boldsymbol{\ell}_1 = 1$. ∎

Fourier space errors for the charges

For the charges, i.e., for $j = 1, \dots, N_\mathrm{c}$, we rewrite the Fourier space parts of the forces $\boldsymbol{F}(j) = -q_j\nabla_{\boldsymbol{x}_j}\phi(j)$ via

$$\boldsymbol{F}^\mathrm{F}(j) = \boldsymbol{F}^\mathrm{F}_\mathrm{c.c.}(j) + \boldsymbol{F}^\mathrm{F}_\mathrm{c.d.}(j),$$

where

$$\boldsymbol{F}^\mathrm{F}_\mathrm{c.c.}(j) := -\frac{q_j\nabla_{\boldsymbol{x}_j}}{\pi V}\sum_{\boldsymbol{k}\in\mathbb{Z}^3}\hat{\psi}(\boldsymbol{k})\sum_{i=1}^{N_\mathrm{c}} q_i\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}} = \frac{2\mathrm{i}q_j}{V}\sum_{i=1}^{N_\mathrm{c}} q_i\sum_{\boldsymbol{k}\in\mathbb{Z}^3}\hat{\psi}(\boldsymbol{k})(\boldsymbol{L}^{-\top}\boldsymbol{k})\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}}, \tag{2.85}$$

$$\begin{aligned}\boldsymbol{F}^\mathrm{F}_\mathrm{c.d.}(j) &:= -\frac{q_j\nabla_{\boldsymbol{x}_j}}{\pi V}\sum_{\boldsymbol{k}\in\mathbb{Z}^3}\hat{\psi}(\boldsymbol{k})\sum_{i=N_\mathrm{c}+1}^{N}\boldsymbol{\mu}_i^\top\nabla_{\boldsymbol{x}_i}\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}}\\ &= -\frac{4\pi q_j}{V}\sum_{i=N_\mathrm{c}+1}^{N}\|\boldsymbol{\mu}_i\|\sum_{\boldsymbol{k}\in\mathbb{Z}^3}\hat{\psi}(\boldsymbol{k})\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|\cos(\gamma_{i,\boldsymbol{k}})(\boldsymbol{L}^{-\top}\boldsymbol{k})\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}},\end{aligned}$$

in order to separate the interactions to the other charges from those to the dipole particles. Here, the angle between a lattice vector $\boldsymbol{L}^{-\top}\boldsymbol{k}$ and a dipole moment $\boldsymbol{\mu}_i$ is shortly denoted by $\gamma_{i,\boldsymbol{k}} := \sphericalangle(\boldsymbol{\mu}_i, \boldsymbol{L}^{-\top}\boldsymbol{k})$.

We denote the corresponding truncated versions by

$$\boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.c.}}(j)_{\mathrm{t}} := \frac{2\mathrm{i}q_j}{V}\sum_{i=1}^{N_{\mathrm{c}}} q_i \sum_{\boldsymbol{k}\in\mathbb{Z}^3\cap\mathcal{J}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})(\boldsymbol{L}^{-\top}\boldsymbol{k})\, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}}, \tag{2.86}$$

$$\boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.d.}}(j)_{\mathrm{t}} := -\frac{4\pi q_j}{V}\sum_{i=N_{\mathrm{c}}+1}^{N} \|\boldsymbol{\mu}_i\| \sum_{\boldsymbol{k}\in\mathbb{Z}^3\cap\mathcal{J}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|\cos(\gamma_{i,\boldsymbol{k}})(\boldsymbol{L}^{-\top}\boldsymbol{k})\, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}}. \tag{2.87}$$

Again, we denote by $\Delta F^{\mathrm{F}}_{\mathrm{c.c.}}$ and $\Delta F^{\mathrm{F}}_{\mathrm{c.d.}}$ the corresponding rms force errors and present estimates for uncorrelated systems in Theorem 2.23.

Theorem 2.23. *If the cutoff parameter $\boldsymbol{M} = (M_1, M_2, M_3) \in \mathbb{R}^3_+$ in* (2.86) *and* (2.87) *is chosen such that* (2.82) *with sufficiently large $\beta > 0$ is fulfilled, we have*

$$\Delta F^{\mathrm{F}}_{\mathrm{c.c.}} := \sqrt{\frac{1}{N_{\mathrm{c}}}\sum_{j=1}^{N_{\mathrm{c}}} \|\boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.c.}}(j) - \boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.c.}}(j)_{\mathrm{t}}\|^2} \approx \alpha Q\sqrt{\frac{8}{\pi\beta N_{\mathrm{c}}V}}\, \mathrm{e}^{-\pi^2\beta^2/4\alpha^2}, \tag{2.88}$$

$$\Delta F^{\mathrm{F}}_{\mathrm{c.d.}} := \sqrt{\frac{1}{N_{\mathrm{c}}}\sum_{j=1}^{N_{\mathrm{c}}} \|\boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.d.}}(j) - \boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.d.}}(j)_{\mathrm{t}}\|^2} \approx \alpha\sqrt{\frac{8\beta\pi Q\mathcal{M}}{3N_{\mathrm{c}}V}}\, \mathrm{e}^{-\pi^2\beta^2/4\alpha^2}, \tag{2.89}$$

provided that the particles are sufficiently homogeneously distributed, cf. [*KP*92, *HNP*17].

Proof. In the charge-charge case we obtain

$$\boldsymbol{\chi}^{\mathrm{c.c.}}_{ij} = \frac{2\mathrm{i}}{V}\sum_{\boldsymbol{k}\notin\mathbb{Z}^3\cap\mathcal{J}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})(\boldsymbol{L}^{-\top}\boldsymbol{k})\, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}}$$

and

$$\chi^2_{\mathrm{c.c.}} := \frac{1}{V^2}\int_{\boldsymbol{L}\mathbb{T}^3}\int_{\boldsymbol{L}\mathbb{T}^3} \|\boldsymbol{\chi}^{\mathrm{c.c.}}_{ij}\|^2 \mathrm{d}\boldsymbol{x}_i\,\mathrm{d}\boldsymbol{x}_j = \frac{4}{V^2}\sum_{\boldsymbol{k}\notin\mathbb{Z}^3\cap\mathcal{J}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2.$$

Now, we have

$$\left\{\boldsymbol{k}\in\mathbb{Z}^3 : \|\boldsymbol{L}^{-\top}\boldsymbol{k}\| < \beta/2\right\} \subset \left\{\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2, \boldsymbol{k}\in\mathbb{Z}^3\cap\mathcal{J}_{\boldsymbol{M}}\right\},$$

see Lemma 2.22. This is used to get an upper bound of the error sum, which can be estimated by an integral in spherical coordinates as follows.

$$\begin{aligned}
\chi^2_{\mathrm{c.c.}} &\approx \frac{4}{V^2\left|\det(\boldsymbol{L}^{-\top})\right|}\int_{\|\boldsymbol{k}\|\geq\beta/2} \frac{\mathrm{e}^{-2\pi^2\|\boldsymbol{k}\|^2/\alpha^2}}{\|\boldsymbol{k}\|^2}\mathrm{d}\boldsymbol{k} \\
&= \frac{4}{V}\int_0^{\pi}\sin\theta\int_0^{2\pi}\int_{\beta/2}^{\infty} \mathrm{e}^{-2\pi^2k^2/\alpha^2}\,\mathrm{d}k\,\mathrm{d}\phi\,\mathrm{d}\theta \\
&= \frac{16\pi}{V}\cdot\frac{\alpha}{2\sqrt{2\pi}}\mathrm{erfc}\left(\frac{\pi\beta}{\sqrt{2}\alpha}\right) \approx \frac{8\alpha^2}{\pi V\beta}\,\mathrm{e}^{-\pi^2\beta^2/2\alpha^2},
\end{aligned}$$

where we use the asymptotic expansion formula (2.21) for the final approximation. We remark that the asymptotic expansion formula only gives exact results if β is chosen not too small.

Via

$$\Delta F^{\mathrm{F}}_{\mathrm{c.c.}} \approx \sqrt{\frac{Q^2 \chi^2_{\mathrm{c.c.}}}{N_{\mathrm{c}}}}$$

we obtain (2.88). We remark that the result is slightly different from the one given in [KP92].

We proceed in an analogue manner in the charge-dipole case and obtain

$$\begin{aligned}\chi^2_{\mathrm{c.d.}} &\approx \frac{16\pi^2}{V^2|\det(\boldsymbol{L}^{-\top})|}\int_0^{\pi}\sin(\theta)\cos^2(\theta)\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\int_{\beta/2}^{\infty}\mathrm{e}^{-2\pi^2k^2/\alpha^2}k^2\mathrm{d}k\\ &= \frac{64\pi^3}{3V}\int_{\beta/2}^{\infty}\mathrm{e}^{-2\pi^2k^2/\alpha^2}k^2\mathrm{d}k \approx \frac{8\beta\alpha^2\pi q_j^2\mathcal{M}}{3V}\mathrm{e}^{-\pi^2\beta^2/2\alpha^2},\end{aligned}$$

where the asymptotic expansion formula (2.72) is used. Again, the approximation is only valid provided that β is chosen large enough. Finally, we obtain (2.89) based on

$$\Delta F^{\mathrm{F}}_{\mathrm{c.d.}} \approx \sqrt{\frac{Q\mathcal{M}\chi^2_{\mathrm{c.d.}}}{N_{\mathrm{c}}}}.$$

■

Fourier space errors for the dipole particles

For the dipoles, i.e., for $j = N_{\mathrm{c}}+1,\ldots,N$, we write the Fourier space parts of the forces $\boldsymbol{F}(j) = -\nabla_{\boldsymbol{x}_j}(\boldsymbol{\mu}_j^{\top}\nabla_{\boldsymbol{x}_j})\phi(j)$ via

$$\boldsymbol{F}^{\mathrm{F}}(j) = \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j) + \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j) \approx \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j)_{\mathrm{t}} + \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j)_{\mathrm{t}},$$

where

$$\begin{aligned}\boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j)_{\mathrm{t}} &:= -\frac{\nabla_{\boldsymbol{x}_j}(\boldsymbol{\mu}_j^{\top}\nabla_{\boldsymbol{x}_j})}{\pi V}\sum_{\boldsymbol{k}\in\mathbb{Z}^3\cap\mathcal{J}_{\boldsymbol{M}}}\hat{\psi}(\boldsymbol{k})\sum_{i=1}^{N_{\mathrm{c}}}q_i\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}}\\ &= \frac{4\pi\|\boldsymbol{\mu}_j\|}{V}\sum_{i=1}^{N_{\mathrm{c}}}q_i\sum_{\boldsymbol{k}\in\mathbb{Z}^3\cap\mathcal{J}_{\boldsymbol{M}}}\hat{\psi}(\boldsymbol{k})\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|\cos(\gamma_{j,\boldsymbol{k}})(\boldsymbol{L}^{-\top}\boldsymbol{k})\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}}\end{aligned} \tag{2.90}$$

and

$$\begin{aligned}\boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j)_{\mathrm{t}} &:= -\frac{\nabla_{\boldsymbol{x}_j}(\boldsymbol{\mu}_j^{\top}\nabla_{\boldsymbol{x}_j})}{\pi V}\sum_{\boldsymbol{k}\in\mathbb{Z}^3\cap\mathcal{J}_{\boldsymbol{M}}}\hat{\psi}(\boldsymbol{k})\sum_{i=N_{\mathrm{c}}+1}^{N}\boldsymbol{\mu}_i^{\top}\nabla_{\boldsymbol{x}_i}\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}}\\ &:= \frac{8\pi^2\mathrm{i}\|\boldsymbol{\mu}_j\|}{V}\sum_{i=N_{\mathrm{c}}+1}^{N}\|\boldsymbol{\mu}_i\|\sum_{\boldsymbol{k}\in\mathbb{Z}^3\cap\mathcal{J}_{\boldsymbol{M}}}\hat{\psi}(\boldsymbol{k})\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2\cos(\gamma_{j,\boldsymbol{k}})\cos(\gamma_{i,\boldsymbol{k}})(\boldsymbol{L}^{-\top}\boldsymbol{k})\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}}.\end{aligned} \tag{2.91}$$

We denote by $\Delta F^{\mathrm{F}}_{\mathrm{d.c.}}$ and $\Delta F^{\mathrm{F}}_{\mathrm{d.d.}}$ the corresponding rms force errors and present estimates for uncorrelated systems in Theorem 2.24.

Theorem 2.24. *If the cutoff parameter* $\boldsymbol{M} = (M_1, M_2, M_3) \in \mathbb{R}^3_+$ *in* (2.90) *and* (2.91) *is chosen such that* (2.82) *with sufficiently large* $\beta > 0$ *is fulfilled, we have*

$$\begin{aligned}
\Delta F^{\mathrm{F}}_{\mathrm{d.c.}} &:= \sqrt{\frac{1}{N_{\mathrm{d}}} \sum_{j=N_{\mathrm{c}}+1}^{N} \|\boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j) - \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j)_{\mathrm{t}}\|^2} \approx \alpha \sqrt{\frac{8\beta\pi Q\mathcal{M}}{3N_{\mathrm{d}}V}}\, \mathrm{e}^{-\pi^2\beta^2/4\alpha^2}, \\
\Delta F^{\mathrm{F}}_{\mathrm{d.d.}} &:= \sqrt{\frac{1}{N_{\mathrm{d}}} \sum_{j=N_{\mathrm{c}}+1}^{N} \|\boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j) - \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j)_{\mathrm{t}}\|^2} \approx 4\alpha\mathcal{M} \sqrt{\frac{\pi^3\beta^3}{15N_{\mathrm{d}}V}}\, \mathrm{e}^{-\pi^2\beta^2/4\alpha^2},
\end{aligned} \tag{2.92}$$

provided that the particles are sufficiently homogeneously distributed, cf. [*WH*01*, HNP*17].

Proof. The same steps which led to the approximation of $\Delta F^{\mathrm{F}}_{\mathrm{c.d.}}$, see proof of Theorem 2.23, yield

$$\Delta F^{\mathrm{F}}_{\mathrm{d.c.}} \approx \sqrt{\frac{Q\mathcal{M}\chi^2_{\mathrm{d.c.}}}{N_{\mathrm{d}}}} \approx \alpha \sqrt{\frac{8\beta\pi Q\mathcal{M}}{3N_{\mathrm{c}}V}}\, \mathrm{e}^{-\pi^2\beta^2/4\alpha^2}.$$

Again, the spherical coordinates are chosen such that $\theta = \gamma_{j,\boldsymbol{k}}$, as explained above. The same is done in the dipole-dipole setting, which has already been studied in [WH01]. In case of dipole-dipole interactions we additionally assume $\boldsymbol{\mu}_i = \|\boldsymbol{\mu}_i\| \cdot (\sin\gamma, 0, \cos\gamma)$ with $\gamma := \sphericalangle(\boldsymbol{\mu}_i, \boldsymbol{\mu}_j)$. Furthermore, $\cos\gamma_{i,\boldsymbol{k}}$ is rewritten via (2.81), i.e.,

$$\cos\gamma_{i,\boldsymbol{k}} = \frac{\boldsymbol{\mu}_i^\top \boldsymbol{k}}{\|\boldsymbol{\mu}_i\| k} = \cos\theta\cos\gamma + \sin\theta\sin\gamma\cos\phi,$$

where $\boldsymbol{k} = (k\sin\theta\cos\phi, k\sin\theta\sin\phi, k\cos\theta)$. In the obtained expression, the sine and cosine terms depending on the angle γ are again simply replaced by the corresponding mean values, i.e., we set $\chi^2_{\mathrm{d.d.}} :=$

$$\frac{1}{\pi}\int_0^\pi \frac{64\pi^4}{V} \int_0^\pi \int_0^{2\pi} \int_{\beta/2}^\infty \mathrm{e}^{-2\pi^2k^2/\alpha^2} k^4 \sin\theta\cos^2\theta\,(\cos\theta\cos\gamma + \sin\theta\sin\phi\sin\gamma)^2 \,\mathrm{d}k\,\mathrm{d}\phi\,\mathrm{d}\theta\,\mathrm{d}\gamma.$$

This gives, cf. [WH01],

$$\chi^2_{\mathrm{d.d.}} \approx 16\alpha^2 \frac{\pi^3\beta^3}{15V}\, \mathrm{e}^{-\pi^2\beta^2/2\alpha^2} \iff \Delta F^{\mathrm{F}}_{\mathrm{d.d.}} \approx 4\alpha\mathcal{M}\sqrt{\frac{\pi^3\beta^3}{15N_{\mathrm{d}}V}}\, \mathrm{e}^{-\pi^2\beta^2/4\alpha^2}.$$

As already remarked in the proof of Theorem 2.23, the estimates are based on the asymptotic expansion formula, which is only applicable if the cutoff parameter β is chosen not too small. ■

2.3.4 Parameter choice

Given a particle system as well as a required accuracy ε we may ask for possible values for the splitting parameter α, the near field cutoff radius r_{cut} as well as the far field cutoff parameter β. As explained in the following remark, the corresponding errors are constant among systems having the same particle densities.

Remark 2.25. Based on the estimates as presented in Theorems 2.18, 2.20, 2.23 and 2.24, it can be seen easily that we expect approximately the same truncation errors among systems having the same particle densities

$$\frac{Q}{V}, \frac{N_\mathrm{c}}{V}, \frac{\mathcal{M}}{V} \text{ and } \frac{N_\mathrm{d}}{V},$$

provided that the same parameters α, r_cut and β are applied. For larger systems the number of Fourier coefficients may simply be scaled up according to the dimensions of the primary box via (2.82). Furthermore, the shape of the primary box is not relevant for the presented estimates. □

Based on the Theorems 2.18, 2.20, 2.23 and 2.24 we may choose the Ewald summation parameters α, r_cut and β in order to reach the required accuracy ε. Roughly speaking, we will ask for two conditions to be fulfilled, namely that the accuracy ε is reached in the near field as well as regarding the approximation in Fourier space. Since three parameters have to be tuned, we still have one degree of freedom. In the following we discuss how the parameters α and β may be determined, provided that a certain near field cutoff radius r_cut is given.

We apply the following common approach.

1. Specify a certain accuracy ε and a near field cutoff radius $r_\mathrm{cut} > 0$.

2. Choose the splitting parameter α such that

$$\max\{\Delta F_\mathrm{c.c.}^\mathrm{short}, \Delta F_\mathrm{c.d.}^\mathrm{short}, \Delta F_\mathrm{d.c.}^\mathrm{short}, \Delta F_\mathrm{d.d.}^\mathrm{short}\} \approx \varepsilon.$$

3. Choose the far field cutoff $\beta > 0$ such that

$$\max\{\Delta F_\mathrm{c.c.}^\mathrm{F}, \Delta F_\mathrm{c.d.}^\mathrm{F}, \Delta F_\mathrm{d.c.}^\mathrm{F}, \Delta F_\mathrm{d.d.}^\mathrm{F}\} \approx \varepsilon. \tag{2.93}$$

Based on the error estimates (2.74), (2.75), (2.78) and (2.79), the appropriate splitting parameter α, see step 2, can be found based on a simple binary search algorithm. In case of charge-charge systems we obtain

$$\frac{2Q}{\sqrt{\pi N_\mathrm{c} V r_\mathrm{cut}}} \mathrm{e}^{-\alpha^2 r_\mathrm{cut}^2} \approx \varepsilon \iff \alpha \approx \frac{1}{r_\mathrm{cut}} \sqrt{\ln\left(\frac{2Q}{\varepsilon\sqrt{\pi N_\mathrm{c} V r_\mathrm{cut}}}\right)},$$

based on (2.77). The obtained α is real valued, provided that $\varepsilon \leq (\pi N_\mathrm{c} V r_\mathrm{cut})^{-1/2} 2Q$. Otherwise, the estimate (2.77) is less than ε for all $\alpha \geq 0$.

In order to determine a sufficiently large cutoff parameter $\beta > 0$, see step 3, we make use of the well known Lambert-W function [CGH+96]. The function is implicitly defined via $\mathrm{W}(x)\,\mathrm{e}^{\mathrm{W}(x)} = x$, i.e., it is the inverse of $f(x) := x\,\mathrm{e}^x$ and is therefore not uniquely defined for $x < 0$. We will make use of the two branches W_0 and W_{-1}, as depicted in Figure 2.6.

In case of charge-charge interactions we determine β based on (2.88) by making use of the principal branch W_0. We have

$$\begin{aligned}\Delta F_\mathrm{c.c.}^\mathrm{F} \approx \varepsilon &\iff \alpha Q \sqrt{\frac{8}{\pi\beta N_\mathrm{c} V}}\, \mathrm{e}^{-\pi^2\beta^2/4\alpha^2} \approx \varepsilon \\ &\iff \frac{8^2\alpha^4 Q^4}{\pi^2\beta^2 {N_\mathrm{c}}^2 V^2}\, \mathrm{e}^{-\pi^2\beta^2/\alpha^2} \approx \varepsilon^4 \iff \frac{\pi^2\beta^2}{\alpha^2}\, \mathrm{e}^{\pi^2\beta^2/\alpha^2} \approx \frac{8^2\alpha^2 Q^4}{{N_\mathrm{c}}^2 V^2 \varepsilon^4}.\end{aligned}$$

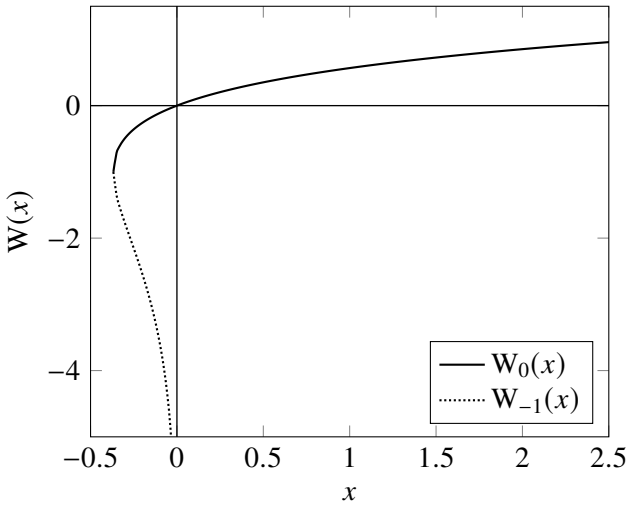


Figure 2.6: The two branches of the Lambert-W function for real arguments $x \geq \mathrm{e}^{-1}$.

Thus, we set

$$\beta_{\text{c.c.}} := \frac{\alpha}{\pi}\sqrt{\mathrm{W}_0\left(\frac{8^2\alpha^2Q^4}{{N_\mathrm{c}}^2V^2\varepsilon^4}\right)}.$$

Note that $\mathrm{W}_0(x) \geq 0$ for all $x \geq 0$ and thus the obtained expression for β is well defined.

In the following considerations we need the Lambert-W function for negative arguments. We obtain

$$\Delta F^{\mathrm{F}}_{\text{c.d.}} \approx \varepsilon \iff \frac{8^2\alpha^6Q^2\mathcal{M}^2}{3^2{N_\mathrm{c}}^2V^2} \cdot \frac{\pi^2\beta^2}{\alpha^2}\,\mathrm{e}^{-\pi^2\beta^2/\alpha^2} \approx \varepsilon^4 \iff -\frac{\pi^2\beta^2}{\alpha^2}\,\mathrm{e}^{-\pi^2\beta^2/\alpha^2} \approx -\frac{3^2\varepsilon^4{N_\mathrm{c}}^2V^2}{8^2\alpha^6Q^2\mathcal{M}^2}$$

and choose

$$\beta_{\text{c.d.}} := \frac{\alpha}{\pi}\sqrt{-\mathrm{W}_{-1}\left(-\frac{3^2\varepsilon^4{N_\mathrm{c}}^2V^2}{8^2\alpha^6Q^2\mathcal{M}^2}\right)}.$$

The final expression is again well defined, provided that the accuracy ε adopts reasonable values. Since $\alpha^{-2}\pi^2\beta^2\,\mathrm{e}^{-\pi^2\beta^2/\alpha^2} \leq \mathrm{e}^{-1}$ for all $\beta \in \mathbb{R}$ we have to choose ε such that

$$\varepsilon^4 \leq \frac{8^2\alpha^6Q^2\mathcal{M}^2}{3^2{N_\mathrm{c}}^2V^2\,\mathrm{e}^1},$$

which ensures that the argument of the Lambert-W function lies in the interval $[-\mathrm{e}^{-1}, 0]$. Since $\mathrm{W}(x) \leq 0$ for $x \in [-\mathrm{e}^{-1}, 0]$ the obtained value β is well defined. Note that the error formula (2.89) is only valid for large enough β. Especially, the error should decrease with growing β, which is only fulfilled if $\beta \geq \pi^{-1}\alpha$. Since $\mathrm{W}_0(x) \in [-1, 0]$ and $\mathrm{W}_{-1}(x) \leq -1$ for $x \in [-\mathrm{e}^{-1}, 0]$ only the usage of the function W_{-1} is reasonable. The same argument applies to the following two cases. In (2.92) only $\beta \geq \pi^{-1/2}\alpha\sqrt{3}$ is reasonable.

Analogously to $\beta_{\text{c.d.}}$, we compute

$$\beta_{\text{d.c.}} := \frac{\alpha}{\pi}\sqrt{-\mathrm{W}_{-1}\left(-\frac{3^2\varepsilon^4 {N_\text{d}}^2 V^2}{8^2\alpha^6 Q^2 \mathcal{M}^2}\right)}$$

in order to obtain $\Delta F^{\text{F}}_{\text{d.c.}} \approx \varepsilon$.

For the dipole-dipole part we get

$$\Delta F^{\text{F}}_{\text{d.c.}} \approx \varepsilon \iff \sqrt[3]{\frac{2^8 3\alpha^{10}\mathcal{M}^4}{5^2 {N_\text{d}}^2 V^2}} \cdot \frac{\pi^2\beta^2}{3\alpha^2}\, \mathrm{e}^{-\pi^2\beta^2/3\alpha^2} \approx \sqrt[3]{\varepsilon^4}$$

and thus

$$\beta_{\text{d.d.}} := \frac{\alpha}{\pi}\sqrt{-3\,\mathrm{W}_{-1}\left(-\sqrt[3]{\frac{5^2 {N_\text{d}}^2 V^2 \varepsilon^4}{2^8 3\alpha^{10}\mathcal{M}^4}}\right)}.$$

In order to end up with (2.93) we may simply use the maximum cutoff parameter

$$\beta_{\text{max}} := \max\{\beta_{\text{c.c.}}, \beta_{\text{c.d.}}, \beta_{\text{d.c.}}, \beta_{\text{d.d.}}\}$$

in order to approximate the Fourier space parts of the forces based on (2.86), (2.87), (2.90) and (2.91), where $\boldsymbol{M}$ is defined via (2.82) with $\beta := \beta_{\text{max}}$.

Example 2.26. We assume a particle system with $N_\text{c} = 300$ charges $q_j = (-1)^j$, $j = 1, \ldots, N_\text{c}$, and $N_\text{d} = 300$ dipoles with $\|\boldsymbol{\mu}_j\| = 1$, $j = N_\text{c}+1, \ldots, N$. Thus, we obtain $Q = \mathcal{M} = 300$. Furthermore, we suppose that the particles are located in an orthorhombic box with $\boldsymbol{\ell}_1 = (20, 0, 0)^\top$, $\boldsymbol{\ell}_2 = (0, 10, 0)^\top$ and $\boldsymbol{\ell}_3 = (0, 0, 10)^\top$, i.e. $V = 2 \cdot 10^3$.

If we choose the accuracy $\varepsilon := 10^{-16}$ and the near field cutoff $r_\text{cut} := 5$, we obtain

$$\alpha \approx 1.27$$

based on the estimates regarding the short range part, see step 2. Consequently, step 3 yields

$$\beta_{\text{max}} \approx 5.14.$$

In Figure 2.7 we plot the estimated rms force errors $\Delta F^{\text{short}}_{\text{type}}$ and $\Delta F^{\text{F}}_{\text{type}}$ for type $\in \{\text{c.c.}, \text{c.d.}, \text{d.c.}, \text{d.d.}\}$ with respect to the splitting parameter α, where we set $r_\text{cut} := 5$ and $\beta := \beta_{\text{max}} \approx 5.14$. The results validate the reasonability of the tuned parameters α and β. □

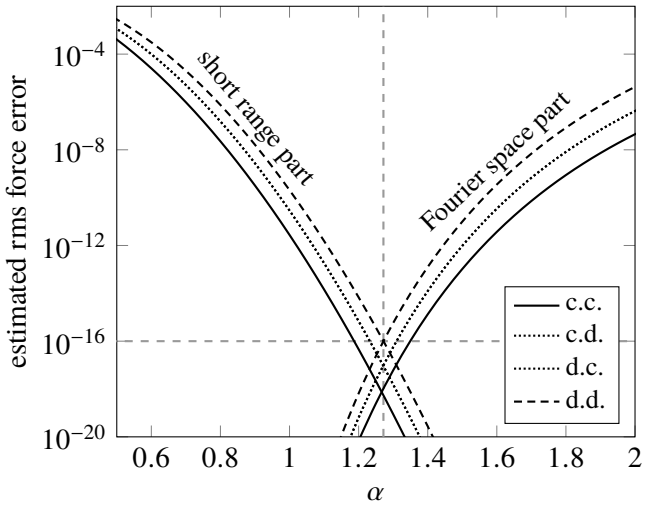


Figure 2.7: Estimated rms force errors for different splitting parameters α, where we set $r_{\text{cut}} := 5$ and $\beta \approx 5.14$. System: $N_{\text{c}} = N_{\text{d}} = Q = \mathcal{M} = 300$, $V = 2 \cdot 10^3$.

3 Nonequispaced fast Fourier transforms in three variables

The well-known fast Fourier transform (FFT) is one of the most important algorithms for a broad variety of applications in the areas of engineering, natural sciences, scientific computing and signal processing. It enables the evaluation of trigonometric sums at equispaced nodes with an outstanding arithmetic complexity of $O(N \log N)$, where N denotes the number of given nodes. The efficient computation is based on a divide-and-conquer strategy, which dates back to the famous algorithm of Cooley and Tukey [CT65] proposed in 1965.

In a wide range of applications a sampling at nonequispaced knots is required, which led to the development of the nonuniform FFT or fast Fourier transform for nonequispaced data (NFFT). In the following we at first introduce the necessary notations as well as the basic concept of FFT. Secondly, we describe the fast evaluation of trigonometric sums at nonequispaced knots in detail. Afterward, we discuss error estimates and present numerical examples.

3.1 Preliminaries and Definitions

In the following we consider 1-periodic functions, where again the one-dimensional torus is denoted by $\mathbb{T} := \mathbb{R}/\mathbb{Z} \simeq [-1/2, 1/2)$. Each function $f \in \mathcal{L}_2(\mathbb{T}^3)$ can be represented by its Fourier series

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^3} c_{\boldsymbol{k}}(f)\, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{x}},$$

where the infinite sum converges with respect to the $\mathcal{L}_2$-norm

$$\|f\|_{\mathcal{L}_2} := \left(\int_{[-1/2, 1/2)^3} |f(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x} \right)^{1/2}.$$

The Fourier coefficients of f are given by

$$c_{\boldsymbol{k}}(f) := \int_{[1/2, 1/2)^3} f(\boldsymbol{x})\, \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{x}} \mathrm{d}\boldsymbol{x}, \quad \boldsymbol{k} \in \mathbb{Z}^3. \tag{3.1}$$

In the following we define for some $\boldsymbol{M} = (M_1, M_2, M_3)^\top \in 2\mathbb{N}^3$ the three dimensional index set

$$\mathcal{I}_{\boldsymbol{M}} := \{-M_1/2, \ldots, M_1/2 - 1\} \times \ldots \times \{-M_3/2, \ldots, M_3/2 - 1\}. \tag{3.2}$$

Furthermore, we denote the component wise product of two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$ via

$$\boldsymbol{x} \odot \boldsymbol{y} := (x_1 y_1, x_2 y_2, x_3 y_3)^\top \in \mathbb{R}^3$$

and for a vector $\boldsymbol{x} \in \mathbb{R}^3$ with non vanishing components we set

$$\boldsymbol{x}^{-1} := (x_1^{-1}, x_2^{-1}, x_3^{-1})^\top \in \mathbb{R}^3.$$

As usual, the Euclidean norm is generalized for complex valued vectors $\boldsymbol{x} \in \mathbb{C}^3$ via

$$\|\boldsymbol{x}\| := \sqrt{\overline{\boldsymbol{x}}^\top \boldsymbol{x}} = \sqrt{|x_1|^2 + |x_2|^2 + |x_3|^2}.$$

Discrete Fourier transform

If the periodic function f is only known at equispaced samples $f_{\boldsymbol{j}} := f(\boldsymbol{j} \odot \boldsymbol{M}^{-1})$, where $\boldsymbol{j} \in \mathcal{I}_{\boldsymbol{M}}$, we have

$$c_{\boldsymbol{k}}(f) \approx \hat{f}_{\boldsymbol{k}} := \frac{1}{|\mathcal{I}_{\boldsymbol{M}}|} \sum_{\boldsymbol{j} \in \mathcal{I}_{\boldsymbol{M}}} f_{\boldsymbol{j}} \, \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^{\top} (\boldsymbol{j} \odot \boldsymbol{M}^{-1})} \quad \text{for all} \quad \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}, \tag{3.3}$$

i.e., a good approximation of the Fourier coefficients over the multi-index set $\mathcal{I}_{\boldsymbol{M}}$ is obtained by simply applying the trapezoidal quadrature rule.

The mapping

$$\{f_{\boldsymbol{j}}\}_{\boldsymbol{j} \in \mathcal{I}_{\boldsymbol{M}}} \mapsto \{\hat{f}_{\boldsymbol{k}}\}_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}}$$

is widely known as discrete Fourier transform (DFT). The DFT is invertible and the function values $f_{\boldsymbol{j}}$ are reconstructed via

$$f_{\boldsymbol{j}} = \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{f}_{\boldsymbol{k}} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^{\top} (\boldsymbol{j} \odot \boldsymbol{M}^{-1})}, \quad \boldsymbol{j} \in \mathcal{I}_{\boldsymbol{M}}, \tag{3.4}$$

as is well known. Note that a naive computation of the sums (3.3) and (3.4) requires $O(|\mathcal{I}_{\boldsymbol{M}}|^2)$ arithmetic operations, which is not satisfying in terms of numerical performance. A significantly more efficient computation is possible via a divide-and-conquer strategy utilizing the underlying trigonometric structure. Roughly speaking, it can be shown that the computation of a one-dimensional DFT of length $M \in 2\mathbb{N}$ requires the computation of two DFTs of length $M/2$, and so on. Of course, the same applies to the inverse DFT. Consequently, the resulting algorithms, which are widely known as the FFT and the inverse FFT (iFFT), enable the exact computation of the sums (3.3) and (3.4), respectively, with only $O(|\mathcal{I}_{\boldsymbol{M}}| \log |\mathcal{I}_{\boldsymbol{M}}|)$ arithmetic operations.

3.2 Fast evaluation of trigonometric sums at nonequispaced knots

For given nonequispaced data, which is the case in several applications, the evaluation of a trigonometric polynomial is not possible based on the classical FFT. An efficient method realizing a fast approximation for nonequispaced data is known as NFFT, see [DR93, Bey95, Ste98, DS99, War98, PST01, GL04, KKP09], for instance. In the following we give an introduction to the NFFT method as well as to some of its modifications. All the considered variants of NFFT take a central role in the algorithms presented in Chapter 4.

3.2.1 NFFT and adjoint NFFT

We are interested in the efficient evaluation of the trigonometric polynomial

$$f(\boldsymbol{x}) := \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{f}_{\boldsymbol{k}} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{x}} \tag{3.5}$$

at N arbitrary nodes $\boldsymbol{x}_j \in \mathbb{T}^3$, which corresponds to the computation of the sums

$$f_j := f(\boldsymbol{x}_j) = \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{f}_{\boldsymbol{k}} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{x}_j}, \quad j = 1, \ldots, N. \tag{3.6}$$

The direct computation requires $O(N|\mathcal{I}_{\boldsymbol{M}}|)$ arithmetic operations and is known as NDFT. In the special case that $\{\boldsymbol{x}_j, j = 1, \ldots, N\} = \{\boldsymbol{j} \odot \boldsymbol{M}^{-1}, \boldsymbol{j} \in \mathcal{I}_{\boldsymbol{M}}\}$ a fast evaluation is possible via the inverse FFT, see (3.4). The NFFT enables an efficient evaluation in the case of arbitrary knots $\boldsymbol{x}_j$.

The basic idea of the NFFT is to apply the ordinary inverse FFT on the given equispaced mesh in Fourier space. Afterward, the function values $f(\boldsymbol{x}_j)$ are recovered based on the computed equispaced data, which is realized via a so called window function, see [Ste98].

More precisely, the function f is finally approximated by a sum of translates of a one-periodic function $\tilde{\varphi}$, which reads as

$$f_{\approx}(\boldsymbol{x}) := \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{\sigma} \odot \boldsymbol{M}}} g_{\boldsymbol{l}} \tilde{\varphi}\left(\boldsymbol{x} - \boldsymbol{l} \odot (\boldsymbol{\sigma} \odot \boldsymbol{M})^{-1}\right). \tag{3.7}$$

In other words, the approximate function values are obtained by computing a discrete convolution of a given window function with some coefficients $g_{\boldsymbol{l}}$, which are by now unknown. Since the number of given Fourier coefficients is $|\mathcal{I}_{\boldsymbol{M}}|$, it seems at first glance reasonable to choose a regular grid of the same size. In order to make the approximation (3.7) more accurate, we may choose an even larger grid of size $\boldsymbol{\sigma} \odot \boldsymbol{M}$, where we denote by $\boldsymbol{\sigma} \geq 1$ (component wise) the oversampling factor. Theoretically, even $\boldsymbol{\sigma} < 1$ is possible.

In the following we denote the oversampled grid size shortly by $\boldsymbol{M}_{\mathrm{o}} := \boldsymbol{\sigma} \odot \boldsymbol{M}$. The function $\tilde{\varphi}$ is the periodization of a window function φ, which is constructed via a tensor product scheme involving univariate functions $\varphi_j : \mathbb{T} \to \mathbb{R}$ concerning the single dimensions $j \in \{1, 2, 3\}$. More precisely, we set

$$\tilde{\varphi}(\boldsymbol{x}) := \sum_{\boldsymbol{r} \in \mathbb{Z}^3} \varphi(\boldsymbol{x} + \boldsymbol{r}), \text{ where } \varphi(\boldsymbol{y}) = \prod_{j=1}^{3} \varphi_j(y_j) \text{ for } \boldsymbol{y} = (y_1, y_2, y_3)^{\top} \in \mathbb{R}^3. \tag{3.8}$$

A transformation of $f_{\approx}$ into Fourier space yields

$$f_{\approx}(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}} \sum_{\boldsymbol{r} \in \mathbb{Z}^3} \hat{g}_{\boldsymbol{k}} c_{\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi})\, \mathrm{e}^{2\pi \mathrm{i} (\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}})^{\top} \boldsymbol{x}}, \tag{3.9}$$

where we denote by $c_{\boldsymbol{k}}(\tilde{\varphi})$, $\boldsymbol{k} \in \mathbb{Z}^3$, the Fourier coefficients of the periodized window function, see (3.1), and by $\hat{g}_{\boldsymbol{k}}$ the discrete Fourier coefficients of $g_{\boldsymbol{l}}$, $\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}$, see (3.3). Note that we have $c_{\boldsymbol{k}}(\tilde{\varphi}) = \hat{\varphi}(\boldsymbol{k})$ for $\boldsymbol{k} \in \mathbb{Z}^3$, which follows from the Poisson summation formula (2.25).

For the following considerations we assume that we have $c_{\boldsymbol{k}}(\tilde{\varphi}) \in \mathbb{R}$. The idea is now to choose the coefficients $\hat{g}_{\boldsymbol{k}}$ appropriately. Then, the coefficients $g_{\boldsymbol{l}}$ in (3.7) can be recovered by the inverse FFT, see (3.4), and the evaluation of (3.7) gives the approximate function values $f_{\approx}(\boldsymbol{x}_j) \approx f(\boldsymbol{x}_j)$.

The evaluation of the sums (3.7) is computationally demanding unless φ is compactly supported on a comparably small domain or at least sufficiently small outside of it. In the latter case we replace the window function φ by a truncated version

$$\varphi_{\mathrm{t}}(\boldsymbol{x}) := \varphi(\boldsymbol{x}) \cdot \prod_{j=1}^{3} \chi_{[-\frac{m}{\sigma_j M_j}, \frac{m}{\sigma_j M_j}]}(x_j) = \begin{cases} \varphi(\boldsymbol{x}) & : \boldsymbol{x} \in \bigotimes_{j=1}^{3} [-\frac{m}{\sigma_j M_j}, \frac{m}{\sigma_j M_j}], \\ 0 & : \text{else}, \end{cases} \tag{3.10}$$

and approximate f by

$$f_{\approx}(\boldsymbol{x}) := \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}} g_{\boldsymbol{l}} \tilde{\varphi}_{\mathrm{t}}\left(\boldsymbol{x} - \boldsymbol{l} \odot \boldsymbol{M}_{\mathrm{o}}^{-1}\right) = \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}, m}(\boldsymbol{x}_j)} g_{\boldsymbol{l}} \tilde{\varphi}_{\mathrm{t}}\left(\boldsymbol{x}_j - \boldsymbol{l} \odot \boldsymbol{M}_{\mathrm{o}}^{-1}\right), \tag{3.11}$$

where we define the sets

$$\mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}, m}(\boldsymbol{x}) := \left\{ \boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}} : \exists \boldsymbol{z} \in \mathbb{Z}^3 \text{ with } \boldsymbol{x} - \boldsymbol{l} \odot \boldsymbol{M}_{\mathrm{o}}^{-1} + \boldsymbol{z} \in \bigotimes_{j=1}^{3} [-\tfrac{m}{\sigma_j M_j}, \tfrac{m}{\sigma_j M_j}] \right\}.$$

In other words, we simply replace $\tilde{\varphi}$ by $\tilde{\varphi}_\mathrm{t}$ in (3.7). Note that in the second variant we only have $(2m+1)^3 \ll |\mathcal{I}_{\boldsymbol{M}_\mathrm{o}}|$ non zero summands, which drastically reduces the required computational effort. Since the parameter $m \in \mathbb{N}$ determines the size of the support $\operatorname{supp}(\varphi_\mathrm{t})$ we will refer to m as the support parameter.

In order to determine the coefficients $g_{\boldsymbol{l}}$ in (3.7) we consider the difference between the trigonometric polynomial f, see (3.5), and its approximation $f_\approx$, see (3.9), where we replace $\tilde{\varphi}$ by $\tilde{\varphi}_\mathrm{t}$ in the case that φ is not compactly supported. We obtain $f(\boldsymbol{x}) - f_\approx(\boldsymbol{x}) =$

$$\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{x}}\hat{f}_{\boldsymbol{k}} - \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}_\mathrm{o}}} \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{x}}\left(\hat{g}_{\boldsymbol{k}}\,c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t}) + \sum_{\boldsymbol{r}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\}} \hat{g}_{\boldsymbol{k}}\,c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_\mathrm{o}}(\tilde{\varphi}_\mathrm{t})\,\mathrm{e}^{2\pi\mathrm{i}(\boldsymbol{r}\odot\boldsymbol{M}_\mathrm{o})^\top\boldsymbol{x}}\right). \tag{3.12}$$

We see that the first $\mathcal{I}_{\boldsymbol{M}_\mathrm{o}}$ Fourier coefficients vanish, i.e, only the aliasing terms are left, if we set

$$\hat{g}_{\boldsymbol{k}} := \begin{cases} \hat{d}_{\boldsymbol{k}}\hat{f}_{\boldsymbol{k}} & : \boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}, \\ 0 & : \text{else}, \end{cases} \tag{3.13}$$

where we define the deconvolution coefficients

$$\hat{d}_{\boldsymbol{k}} := \frac{1}{c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t})}, \tag{3.14}$$

cf. [PST01] for instance. Of course, the choice of the coefficients $\hat{g}_{\boldsymbol{k}}$ or $\hat{d}_{\boldsymbol{k}}$ and can be further optimized.

If we optimize the deconvolution coefficients with respect to the $\mathcal{L}_2$-norm, we obtain the following result, see also [DS99, Appendix A] or [Jac09], for instance.

Lemma 3.1. *For f and $f_\approx$ as defined via* (3.5) *and* (3.11)*, respectively, we have*

$$\operatorname*{arg\,min}_{\hat{d}_{\boldsymbol{k}}\in\mathbb{R},\,\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \|f - f_\approx\|_{\mathcal{L}_2} = \left\{ \frac{c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t})}{\sum\limits_{\boldsymbol{r}\in\mathbb{Z}^3} c^2_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_\mathrm{o}}(\tilde{\varphi}_\mathrm{t})}, \boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}} \right\}. \tag{3.15}$$

Proof. We easily compute

$$\|f - f_\approx\|^2_{\mathcal{L}_2} = \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \left(1 - \hat{d}_{\boldsymbol{k}}\,c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t})\right)^2 + \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \sum_{\boldsymbol{r}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}} \hat{d}^2_{\boldsymbol{k}}\,c^2_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_\mathrm{o}}(\tilde{\varphi}_\mathrm{t}) \tag{3.16}$$

by using (3.12), (3.13) as well as Parseval's identity. Now, the optimal coefficients $\hat{d}_{\boldsymbol{k}}$ are simply determined by differentiating (3.16) with respect to $\hat{d}_{\boldsymbol{k}}$ and setting the result equal to zero for each $\boldsymbol{k}$. We obtain

$$\begin{aligned} 0 &= -2c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t})(1 - \hat{d}_{\boldsymbol{k}}\,c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t})) + 2\hat{d}_{\boldsymbol{k}} \sum_{\boldsymbol{r}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}} c^2_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_\mathrm{o}}(\tilde{\varphi}_\mathrm{t}) \\ &= -2c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t}) + 2\hat{d}_{\boldsymbol{k}} \sum_{\boldsymbol{r}\in\mathbb{Z}^3} c^2_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_\mathrm{o}}(\tilde{\varphi}_\mathrm{t}). \\ \iff \hat{d}_{\boldsymbol{k}} &= \frac{c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t})}{\sum\limits_{\boldsymbol{r}\in\mathbb{Z}^3} c^2_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_\mathrm{o}}(\tilde{\varphi}_\mathrm{t})}, \end{aligned} \tag{3.17}$$

which gives (3.15). ∎

In summary, we obtain

$$\|f - f_{\approx}\|_{\mathcal{L}_2}^2 = \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \sum_{\boldsymbol{r} \in \mathbb{Z}^3 \setminus \{\boldsymbol{0}\}} \frac{c_{\boldsymbol{k}+\boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}})}{c_{\boldsymbol{k}}^2(\tilde{\varphi}_{\mathrm{t}})} \qquad \text{for } \hat{d}_{\boldsymbol{k}} := \frac{1}{c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})}, \tag{3.18}$$

$$\|f - f_{\approx}\|_{\mathcal{L}_2}^2 = \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \frac{\sum\limits_{\boldsymbol{r} \in \mathbb{Z}^3 \setminus \{\boldsymbol{0}\}} c_{\boldsymbol{k}+\boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}})}{\sum\limits_{\boldsymbol{r} \in \mathbb{Z}^3} c_{\boldsymbol{k}+\boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}})} \qquad \text{for } \hat{d}_{\boldsymbol{k}} := \frac{c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})}{\sum\limits_{\boldsymbol{r} \in \mathbb{Z}^3} c_{\boldsymbol{k}+\boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}})}. \tag{3.19}$$

Usually, the Fourier transform of the window function is supposed to be well localized, i.e., we will typically have $|c_{\boldsymbol{k}+\boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})| \ll |c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})|$. Thus, we obtain

$$\frac{c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})}{\sum\limits_{\boldsymbol{r} \in \mathbb{Z}^3} c_{\boldsymbol{k}+\boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}})} \approx \frac{c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})}{c_{\boldsymbol{k}}^2(\tilde{\varphi}_{\mathrm{t}})} = \frac{1}{c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})},$$

i.e., the optimal deconvolution approach is more or less equivalent to the standard approach, see (3.14), provided that the Fourier coefficients of the window function decay fast enough.

The NFFT algorithm has the arithmetic complexity $O(N + |\mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}| \log |\mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}|)$ and can be summarized as follows.

Algorithm 3.1 (NFFT)**.**
Input: Nodes $\boldsymbol{x}_j \in \mathbb{T}^3$ ($j = 1, \ldots, N$), coefficients $\hat{f}_{\boldsymbol{k}} \in \mathbb{C}$ ($\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}$), oversampling factor $\boldsymbol{\sigma} \in \mathbb{R}^3$, $\boldsymbol{\sigma} \geq \boldsymbol{1}$, window function φ and support parameter $m \in \mathbb{N}$.

i) (De-)convolution in Fourier space:
Define the factors $\hat{d}_{\boldsymbol{k}} \in \mathbb{C}$ for all $\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}$, e.g., as given in (3.14) or (3.17), and set

$$\hat{g}_{\boldsymbol{k}} := \begin{cases} \hat{d}_{\boldsymbol{k}} \hat{f}_{\boldsymbol{k}} & : \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}, \\ 0 & : \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}} \setminus \mathcal{I}_{\boldsymbol{M}}. \end{cases}$$

Complexity: $O(|\mathcal{I}_{\boldsymbol{M}}|)$.

ii) Apply the inverse FFT in order to compute the coefficients

$$g_{\boldsymbol{l}} := \frac{1}{|\mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}|} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}} \hat{g}_{\boldsymbol{k}} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^{\top} (\boldsymbol{l} \odot \boldsymbol{M}_{\mathrm{o}}^{-1})}$$

for all $\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}$. Complexity: $O(|\mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}| \log |\mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}|)$.

iii) Convolution in spatial domain: Compute

$$f(\boldsymbol{x}_j) \approx f_{\approx}(\boldsymbol{x}_j) := \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}} g_{\boldsymbol{l}} \tilde{\varphi}_{\mathrm{t}} \left(\boldsymbol{x}_j - \boldsymbol{l} \odot \boldsymbol{M}_{\mathrm{o}}^{-1} \right)$$

for all $j = 1, \ldots, N$. Complexity: $O(m^3 N)$.

Output: $f_{\approx}(\boldsymbol{x}_j) \approx f(\boldsymbol{x}_j)$ for $j = 1, \ldots, N$. □

Using a matrix-vector notation we can rewrite (3.6) via

$$\boldsymbol{f} = \boldsymbol{A}\hat{\boldsymbol{f}},$$

where we define

$$\boldsymbol{f} := (f_j)_{j=1}^{N} \in \mathbb{C}^N, \quad \hat{\boldsymbol{f}} := (\hat{f}_{\boldsymbol{k}})_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \in \mathbb{C}^{|\mathcal{I}_{\boldsymbol{M}}|}, \quad \boldsymbol{A} = (\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{x}_j})_{j=1,\ldots,N;\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \in \mathbb{C}^{N\times|\mathcal{I}_{\boldsymbol{M}}|}$$

and assume the multi-indices $\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}$ to be sorted in a certain manner. The NFFT algorithm can be reformulated via

$$\boldsymbol{f}_{\approx} := \boldsymbol{B}\boldsymbol{F}\boldsymbol{D}\hat{\boldsymbol{f}}, \tag{3.20}$$

where $\boldsymbol{D} \in \mathbb{C}^{|\mathcal{I}_{\boldsymbol{M}_\mathrm{o}}|\times|\mathcal{I}_{\boldsymbol{M}}|}$ is a block diagonal matrix with entries $\hat{d}_{\boldsymbol{k}}$ and 0, the matrix $\boldsymbol{F} \in \mathbb{C}^{|\mathcal{I}_{\boldsymbol{M}_\mathrm{o}}|\times|\mathcal{I}_{\boldsymbol{M}_\mathrm{o}}|}$ is representing the inverse FFT in three dimensions and $\boldsymbol{B} \in \mathbb{R}^{N\times|\mathcal{I}_{\boldsymbol{M}_\mathrm{o}}|}$ is a sparse, real valued matrix with entries $\tilde{\varphi}_\mathrm{t}(\boldsymbol{x}_j - \boldsymbol{l} \odot \boldsymbol{M}_\mathrm{o}^{-1})$.

The direct evaluation of sums of the form

$$h(\boldsymbol{k}) = \sum_{j=1}^{N} f_j \,\mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{x}_j}, \quad \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}, \tag{3.21}$$

with given coefficients $f_j \in \mathbb{C}$ requires $O(N|\mathcal{I}_{\boldsymbol{M}}|)$ arithmetic operations and is known as the adjoint NDFT. It gets its name from the fact that the problem can be reformulated in matrix-vector notation via

$$\boldsymbol{h} = \boldsymbol{A}^{*}\boldsymbol{f},$$

where $\boldsymbol{A}^{*}$ is the Hermitian transpose of $\boldsymbol{A}$ or rather the adjoint of the assigned operator. Thus, the derivation of a fast algorithm, which we refer to as the adjoint NFFT, is straight forward and derived directly from the matrix representation

$$\boldsymbol{h}_{\approx} := \boldsymbol{D}^{*}\boldsymbol{F}^{*}\boldsymbol{B}^{*}\boldsymbol{f}. \tag{3.22}$$

The algorithm can be summarized as follows.

Algorithm 3.2 (adjoint NFFT)**.**
Input: Nodes $\boldsymbol{x}_j \in \mathbb{T}^3$ and coefficients $f_j \in \mathbb{C}$ $(j = 1, \ldots, N)$, mesh size $\boldsymbol{M} \in 2\mathbb{N}^3$, oversampling factor $\boldsymbol{\sigma} \in \mathbb{R}^3$, $\boldsymbol{\sigma} \geq \boldsymbol{1}$, window function φ and support parameter $m \in \mathbb{N}$.

i) Convolution in spatial domain: Compute

$$g_{\boldsymbol{l}} := \sum_{j=1}^{N} f_j \tilde{\varphi}_\mathrm{t}\left(\boldsymbol{x}_j - \boldsymbol{l} \odot \boldsymbol{M}_\mathrm{o}^{-1}\right)$$

for all $\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}_\mathrm{o}}$. Complexity: $O(m^3 N)$.

ii) Apply the FFT in order to compute the coefficients

$$\hat{g}_{\boldsymbol{k}} := \frac{1}{|\mathcal{I}_{\boldsymbol{M}_\mathrm{o}}|} \sum_{\boldsymbol{l}\in\mathcal{I}_{\boldsymbol{M}_\mathrm{o}}} g_{\boldsymbol{l}} \,\mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{k}^{\top}(\boldsymbol{l}\odot\boldsymbol{M}_\mathrm{o}^{-1})}$$

for all $\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}_\mathrm{o}}$. Complexity: $O(|\mathcal{I}_{\boldsymbol{M}_\mathrm{o}}| \log |\mathcal{I}_{\boldsymbol{M}_\mathrm{o}}|)$.

iii) (De-)convolution in Fourier space:
Define the factors $\hat{d}_{\boldsymbol{k}} \in \mathbb{C}$ for all $\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}$, e.g., as given in (3.14) or (3.17), and set

$$h_{\approx}(\boldsymbol{k}) := \hat{d}_{\boldsymbol{k}} \hat{g}_{\boldsymbol{k}}, \quad \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}.$$

Complexity: $\mathcal{O}(|\mathcal{I}_{\boldsymbol{M}}|)$.

Output: $h_{\approx}(\boldsymbol{k}) \approx h(\boldsymbol{k})$ for $\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}$. □

By applying

$$\tilde{\varphi}_{\mathrm{t}}(\boldsymbol{x}) = \sum_{\boldsymbol{r} \in \mathbb{Z}^3} c_{\boldsymbol{r}}(\tilde{\varphi}_{\mathrm{t}}) \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{r}^{\top} \boldsymbol{x}} = \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}} \sum_{\boldsymbol{r} \in \mathbb{Z}^3} c_{\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}}) \, \mathrm{e}^{2\pi \mathrm{i} (\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}})^{\top} \boldsymbol{x}}$$

we easily compute

$$\begin{aligned} h_{\approx}(\boldsymbol{k}) &= \frac{\hat{d}_{\boldsymbol{k}}}{|\mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}|} \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}} \left(\sum_{j=1}^{N} f_j \tilde{\varphi}_{\mathrm{t}} \left(\boldsymbol{x}_j - \boldsymbol{l} \odot \boldsymbol{M}_{\mathrm{o}}^{-1} \right) \right) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{l}^{\top} (\boldsymbol{l} \odot \boldsymbol{M}_{\mathrm{o}}^{-1})} \\ &= \sum_{\boldsymbol{r} \in \mathbb{Z}^3} \sum_{j=1}^{N} f_j \hat{d}_{\boldsymbol{k}} \, c_{\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}}(\varphi_{\mathrm{t}}) \, \mathrm{e}^{-2\pi \mathrm{i} (\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}})^{\top} \boldsymbol{x}_j}. \end{aligned} \tag{3.23}$$

Remark 3.2 (Operator notation)**.** For some $\boldsymbol{x} \in \mathbb{T}^3$ we define the operator $\mathcal{T}_{\boldsymbol{x}} : \mathbb{C}^{|\mathcal{I}_{\boldsymbol{M}}|} \to \mathbb{C}$ by

$$\mathcal{T}_{\boldsymbol{x}} : (\hat{f}_{\boldsymbol{k}})_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \mapsto \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{f}_{\boldsymbol{k}} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{x}}.$$

The adjoint operator $\mathcal{T}_{\boldsymbol{x}}^{*} : \mathbb{C} \to \mathbb{C}^{|\mathcal{I}_{\boldsymbol{M}}|}$ is given by

$$\mathcal{T}_{\boldsymbol{x}}^{*} : y \mapsto \left(y \, \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{x}} \right)_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}}.$$

Based on (3.6) and (3.20) we can write

$$f_j = \mathcal{T}_{\boldsymbol{x}_j} \hat{\boldsymbol{f}} \approx \mathcal{B}_{\boldsymbol{x}_j} \mathcal{F} \mathcal{D} \hat{\boldsymbol{f}}, \quad j = 1, \ldots, N,$$

as well as

$$h(\boldsymbol{k}) = \sum_{j=1}^{N} \mathcal{T}_{\boldsymbol{x}_j}^{*} f_j \approx \sum_{j=1}^{N} \mathcal{D}^{*} \mathcal{F}^{*} \mathcal{B}_{\boldsymbol{x}_j}^{*} f_j, \quad \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}},$$

see (3.21) and (3.22).

The deconvolution coefficients (3.17) are optimal for the adjoint problem in the following sense.

$$\operatorname*{arg\,min}_{\hat{d}_{\boldsymbol{k}} \in \mathbb{R}, \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \int_{[-1/2, 1/2)^3} \| \mathcal{T}_{\boldsymbol{x}}^{*} y - \mathcal{D}^{*} \mathcal{F}^{*} \mathcal{B}_{\boldsymbol{x}}^{*} y \|^2 \, \mathrm{d}\boldsymbol{x} = \left\{ \frac{c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})}{\sum\limits_{\boldsymbol{r} \in \mathbb{Z}^3} c_{\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}}^2 (\tilde{\varphi}_{\mathrm{t}})}, \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}} \right\}$$

for each $y \in \mathbb{C}$.

The result can be shown as follows. In the case

$$h(\boldsymbol{k}) = y \, \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{x}}$$

we obtain by using (3.23)

$$\int_{[-1/2,1/2)^3} \|\boldsymbol{h} - \boldsymbol{h}_{\approx}\|^2 \,\mathrm{d}\boldsymbol{x} = \int_{[-1/2,1/2)^3} \|\mathcal{T}_{\boldsymbol{x}}^* y - \mathcal{D}^*\mathcal{F}^*\mathcal{B}_{\boldsymbol{x}}^* y\|^2 \,\mathrm{d}\boldsymbol{x}$$
$$= |y|^2 \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \left(1 - \hat{d}_{\boldsymbol{k}}\, c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})\right)^2 + |y|^2 \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \sum_{\boldsymbol{r}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\}} \hat{d}_{\boldsymbol{k}}^2\, c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}}).$$

Now, the optimal deconvolution coefficients are easily determined via minimizing each summand with respect to $\hat{d}_{\boldsymbol{k}}$, cf. proof of Lemma 3.1. □

3.2.2 Gradient, adjoint Gradient and Hessian NFFT

Gradient NFFT

We consider a trigonometric polynomial f as given in (3.5) and aim to evaluate its gradient ∇f for the N given nodes $\boldsymbol{x}_j \in \mathbb{T}^3$. The fast algorithm for the efficient approximation of the gradients is known as gradient NFFT, see [Pip15].

The gradients can be computed by applying the differentiation operator directly in Fourier space, which gives

$$\nabla f(\boldsymbol{x}_j) = 2\pi\mathrm{i} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \boldsymbol{k}\, \hat{f}_{\boldsymbol{k}}\, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{x}_j} \in \mathbb{C}^3, \tag{3.24}$$

$j = 1,\dots,N$. An efficient computation is now possible via the NFFT, see Algorithm 3.1, applied in each of the single dimensions. In other words, we formally replace the scalar coefficients $\hat{f}_{\boldsymbol{k}}$ by the vector values coefficients $2\pi\mathrm{i}\boldsymbol{k}\,\hat{f}_{\boldsymbol{k}}$ in Algorithm 3.1, which gives

$$\nabla f(\boldsymbol{x}_j) \approx \nabla f_{\mathrm{ik}}(\boldsymbol{x}_j) := 2\pi\mathrm{i} \sum_{\boldsymbol{r}\in\mathbb{Z}^3} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \boldsymbol{k}\, \hat{d}_{\boldsymbol{k}}\, \hat{f}_{\boldsymbol{k}}\, c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\varphi_{\mathrm{t}})\, \mathrm{e}^{2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^\top\boldsymbol{x}_j}. \tag{3.25}$$

This approach is widely known as $\mathrm{i}\boldsymbol{k}$-differentiation. Note that we need to compute three inverse 3d-FFTs in the second step since all sums are vector valued. The corresponding approximation error measured in the $\mathcal{L}_2$ norm is given by

$$\|\nabla f - \nabla f_{\mathrm{ik}}\|_{\mathcal{L}_2}^2 := \int_{[-1/2,1/2)^3} \|\nabla f(\boldsymbol{x}) - \nabla f_{\mathrm{ik}}(\boldsymbol{x})\|^2 \,\mathrm{d}\boldsymbol{x}$$
$$= 4\pi^2 \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \|\boldsymbol{k}\|^2 \left(1 - \hat{d}_{\boldsymbol{k}}\, c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})\right)^2 + 4\pi^2 \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \|\boldsymbol{k}\|^2 \sum_{\boldsymbol{r}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}} \hat{d}_{\boldsymbol{k}}^2\, c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}}).$$

Differentiation with respect to $\hat{d}_{\boldsymbol{k}}$ shows that the deconvolution coefficients given in (3.17) are also optimal for the $\mathrm{i}\boldsymbol{k}$ gradient NFFT.

In a second variant the differentiation operator is applied to the window function $\tilde{\varphi}$ in the third step of the algorithm. In other words, we formally replace $\tilde{\varphi}_{\mathrm{t}}$ by its gradient $\nabla\tilde{\varphi}_{\mathrm{t}}$ in Algorithm 3.1 and obtain

$$\nabla f_{\mathrm{ad}}(\boldsymbol{x}_j) := \sum_{\boldsymbol{l}\in\mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}} \left(\frac{1}{|\mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}|} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{d}_{\boldsymbol{k}}\, \hat{f}_{\boldsymbol{k}}\, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top(\boldsymbol{l}\odot\boldsymbol{M}_{\mathrm{o}}^{-1})} \right) \nabla\tilde{\varphi}_{\mathrm{t}}(\boldsymbol{x}_j - \boldsymbol{l}\odot\boldsymbol{M}_{\mathrm{o}}^{-1}).$$

The first and the second step of the described procedure are the same as in Algorithm 3.1, i.e., only scalar sums have to be evaluated. The step from one to three dimensions has completely been shifted to the third step. Thus, only one inverse 3d-FFT has to be computed in the second step, while

three inverse 3d-FFTs are required for the i$\boldsymbol{k}$-differentiation approach. Replacing the gradient of the truncated window function by its Fourier representation

$$\nabla\tilde{\varphi}_{\mathrm{t}}(\boldsymbol{x}) = 2\pi\mathrm{i}\sum_{\boldsymbol{r}\in\mathbb{Z}^3}\boldsymbol{r}\,c_{\boldsymbol{r}}(\tilde{\varphi}_{\mathrm{t}})\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{r}^{\top}\boldsymbol{x}} \tag{3.26}$$

yields

$$\nabla f_{\mathrm{ad}}(\boldsymbol{x}_j) = 2\pi\mathrm{i}\sum_{\boldsymbol{r}\in\mathbb{Z}^3}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})\hat{d}_{\boldsymbol{k}}\hat{f}_{\boldsymbol{k}}\,c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\varphi_{\mathrm{t}})\,\mathrm{e}^{2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{x}_j}. \tag{3.27}$$

The approximation error measured in the $\mathcal{L}_2$-norm reads as

$$\begin{aligned}\|\nabla f - \nabla f_{\mathrm{ad}}\|_{\mathcal{L}_2}^2 &:= \int_{[-1/2,1/2)^3}\|\nabla f(\boldsymbol{x}) - \nabla f_{\mathrm{ad}}(\boldsymbol{x})\|^2\,\mathrm{d}\boldsymbol{x}\\ &= 4\pi^2\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}|\hat{f}_{\boldsymbol{k}}|^2\|\boldsymbol{k}\|^2\left(1-\hat{d}_{\boldsymbol{k}}c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})\right)^2\\ &\quad + 4\pi^2\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}|\hat{f}_{\boldsymbol{k}}|^2\sum_{\boldsymbol{r}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}}\|\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}\|^2\hat{d}_{\boldsymbol{k}}^2c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}})\end{aligned}$$

and becomes minimal for the optimal deconvolution coefficients

$$\hat{d}_{\boldsymbol{k}} = \frac{\|\boldsymbol{k}\|^2 c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})}{\displaystyle\sum_{\boldsymbol{r}\in\mathbb{Z}^3}\|\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}\|^2 c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}})}. \tag{3.28}$$

Note that $\|\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}\| \geq \|\boldsymbol{k}\|$ for all $\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}$ and $\boldsymbol{r}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}$ and thus we expect somewhat larger approximation errors than in case of i$\boldsymbol{k}$-differentiation.

Remark 3.3. We remark that the obtained error estimate in case of analytic differentiation is only valid if the truncated window function $\tilde{\varphi}_{\mathrm{t}}$ is differentiable on $\mathbb{T}^3$. Otherwise, the representation (3.26) is not valid. □

Hessian NFFT

Very similarly to the gradient NFFT, we are also able to compute the Hessians

$$(\nabla\nabla^{\top})f(\boldsymbol{x}_j) =: (\mathcal{H}f)(\boldsymbol{x}_j) \in \mathbb{C}^{3\times 3} \tag{3.29}$$

of the trigonometric polynomial (3.5) at arbitrary nodes $\boldsymbol{x}_j\in\mathbb{T}^3$, $j=1,\ldots,N$. Differentiation in Fourier space yields

$$(\mathcal{H}f)(\boldsymbol{x}_j) = -4\pi^2\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}\boldsymbol{k}\boldsymbol{k}^{\top}\hat{f}_{\boldsymbol{k}}\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{x}_j}.$$

The efficient approximation of the matrices is possible via the NFFT, see Algorithm 3.1, where we formally replace the scalar coefficients $\hat{f}_{\boldsymbol{k}}\in\mathbb{C}$ by the matrix valued coefficients $-4\pi^2\boldsymbol{k}\boldsymbol{k}^{\top}\hat{f}_{\boldsymbol{k}}\in\mathbb{C}^{3\times 3}$. The described i$\boldsymbol{k}$-differentiation approach results in an approximation of the Hessian matrices by

$$\mathcal{H}f_{\mathrm{ik}}(\boldsymbol{x}_j) := -4\pi^2\sum_{\boldsymbol{r}\in\mathbb{Z}^3}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}}\boldsymbol{k}\boldsymbol{k}^{\top}\hat{d}_{\boldsymbol{k}}\hat{f}_{\boldsymbol{k}}\,c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\varphi_{\mathrm{t}})\,\mathrm{e}^{2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{x}_j}. \tag{3.30}$$

We obtain

$$\begin{aligned}\|\mathcal{H}f - \mathcal{H}f_{\mathrm{ik}}\|_{\mathcal{L}_2}^2 &:= \int_{[-1/2,1/2)^3} \|(\mathcal{H}f)(\boldsymbol{x}) - (\mathcal{H}f_{\mathrm{ik}})(\boldsymbol{x})\|_{\mathrm{F}}^2 \,\mathrm{d}\boldsymbol{x} \\ &= 16\pi^4 \sum_{\boldsymbol{k}\in I_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \|\boldsymbol{k}\boldsymbol{k}^{\top}\|_{\mathrm{F}}^2 \left(1 - \hat{d}_{\boldsymbol{k}} c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})\right)^2 \\ &\quad + 16\pi^4 \sum_{\boldsymbol{k}\in I_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \|\boldsymbol{k}\boldsymbol{k}^{\top}\|_{\mathrm{F}}^2 \sum_{\boldsymbol{r}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}} \hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}}),\end{aligned}$$

where we denote by

$$\|\boldsymbol{A}\|_{\mathrm{F}} := \sqrt{\sum_{i,j=1}^{3} |a_{ij}|^2}, \quad \boldsymbol{A} = (a_{ij})_{i,j=1}^{3} \in \mathbb{C}^{3\times 3},$$

the well known Frobenius norm. The corresponding optimal deconvolution coefficients are again given by (3.17).

For the analytic differentiation approach we formally replace the window function φ_{t} by its Hessian $\mathcal{H}\varphi_{\mathrm{t}}$ in the third step of Algorithm 3.1. The computed approximations are of the form

$$\mathcal{H}f_{\mathrm{ad}}(\boldsymbol{x}) := -4\pi^2 \sum_{\boldsymbol{r}\in\mathbb{Z}^3} \sum_{\boldsymbol{k}\in I_{\boldsymbol{M}_{\mathrm{o}}}} (\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top} \hat{d}_{\boldsymbol{k}} \hat{f}_{\boldsymbol{k}} c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\varphi_{\mathrm{t}})\, \mathrm{e}^{2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{x}_j},$$

which is obtained analogously to (3.27). The corresponding approximation error measured in the $\mathcal{L}_2$-norm

$$\begin{aligned}\|\mathcal{H}f - \mathcal{H}f_{\mathrm{ad}}\|_{\mathcal{L}_2}^2 &:= \int_{[-1/2,1/2)^3} \|(\mathcal{H}f)(\boldsymbol{x}) - (\mathcal{H}f_{\mathrm{ad}})(\boldsymbol{x})\|_{\mathrm{F}}^2 \,\mathrm{d}\boldsymbol{x} \\ &= 16\pi^4 \sum_{\boldsymbol{k}\in I_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \|\boldsymbol{k}\boldsymbol{k}^{\top}\|_{\mathrm{F}}^2 \left(1 - \hat{d}_{\boldsymbol{k}} c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})\right)^2 \\ &\quad + 16\pi^4 \sum_{\boldsymbol{k}\in I_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \sum_{\boldsymbol{r}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}} \|(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\|_{\mathrm{F}}^2 \hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}})\end{aligned}$$

adopts its minimum for the optimal deconvolution coefficients

$$\hat{d}_{\boldsymbol{k}} = \frac{\|\boldsymbol{k}\boldsymbol{k}^{\top}\|_{\mathrm{F}}^2 c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})}{\displaystyle\sum_{\boldsymbol{r}\in\mathbb{Z}^3} \|(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\|_{\mathrm{F}}^2 c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}})}.$$

We remark again that the analytic differentiation approach is computationally less demanding, since we have to compute only one inverse 3d-FFT in the second step of the algorithm, whereas, if symmetry properties are utilized, six different inverse 3d-FFTs are required for the i$\boldsymbol{k}$-differentiation approach.

Adjoint gradient NFFT

For the given vector valued coefficients $\boldsymbol{f}_j \in \mathbb{C}^3$, $j = 1,\ldots,N$, we consider the sums

$$h(\boldsymbol{k}) := \sum_{j=1}^{N} \boldsymbol{f}_j^{\top} \nabla_{\boldsymbol{x}_j} \mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{x}_j} := \sum_{j=1}^{N} \boldsymbol{f}_j^{\top} \nabla_{\boldsymbol{x}} \mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{x}}\Big|_{\boldsymbol{x}=\boldsymbol{x}_j} \in \mathbb{C}. \tag{3.31}$$

We refer to the efficient approximation of these sums as the adjoint gradient NFFT. The fast evaluation can be realized as follows. If the differentiation operators are directly applied in Fourier space, we obtain

$$h(\boldsymbol{k}) = -2\pi\mathrm{i}\sum_{j=1}^{N} \boldsymbol{f}_j^{\top}\boldsymbol{k}\ \mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{x}_j}.$$

Following the i$\boldsymbol{k}$-differentiation approach, we formally replace the scalar coefficients $f_j \in \mathbb{C}$ by the vectors $\boldsymbol{f}_j \in \mathbb{C}^3$ in Algorithm 3.2. The obtained vectors are multiplied by $-2\pi\mathrm{i}\boldsymbol{k}$ afterward, i.e., for all $\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}$ we finally set

$$h_{\mathrm{ik}}(\boldsymbol{k}) := -2\pi\mathrm{i}\hat{d}_{\boldsymbol{k}}\hat{\boldsymbol{g}}_{\boldsymbol{k}}^{\top}\boldsymbol{k} = -2\pi\mathrm{i}\hat{d}_{\boldsymbol{k}}\boldsymbol{k}^{\top}\sum_{\boldsymbol{r}\in\mathbb{Z}^3}\sum_{j=1}^{N}\boldsymbol{f}_j c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\,\mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{x}}, \tag{3.32}$$

where the computed Fourier coefficients are now vector valued, i.e., $\hat{\boldsymbol{g}}_{\boldsymbol{k}} \in \mathbb{C}^3$.

For the analytic differentiation approach we formally replace $f_j\tilde{\varphi}_{\mathrm{t}}$ by $\boldsymbol{f}_j^{\top}\nabla\tilde{\varphi}_{\mathrm{t}}$ in the first step of Algorithm 3.2 and obtain the approximations

$$h_{\mathrm{ad}}(\boldsymbol{k}) := -2\pi\mathrm{i}\hat{d}_{\boldsymbol{k}}\sum_{\boldsymbol{r}\in\mathbb{Z}^3}\sum_{j=1}^{N}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{f}_j c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\,\mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{x}}. \tag{3.33}$$

Remark 3.4 (Operator notation)**.** We consider the operator $\mathcal{T}_{\boldsymbol{x}} : \mathbb{C}^{|\mathcal{I}_{\boldsymbol{M}}|} \to \mathbb{C}^3$ defined via

$$\mathcal{T}_{\boldsymbol{x}} : (\hat{f}_{\boldsymbol{k}})_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \mapsto 2\pi\mathrm{i}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}\boldsymbol{k}\hat{f}_{\boldsymbol{k}}\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{x}}.$$

The adjoint operator $\mathcal{T}_{\boldsymbol{x}}^{*} : \mathbb{C}^3 \to \mathbb{C}^{|\mathcal{I}_{\boldsymbol{M}}|}$ is given by

$$\mathcal{T}_{\boldsymbol{x}}^{*} : \boldsymbol{y} \mapsto \left(2\pi\mathrm{i}\boldsymbol{y}^{\top}\boldsymbol{k}\ \mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{x}}\right)_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}.$$

We approximate this operator via

$$\mathcal{T}_{\boldsymbol{x}}^{*} \approx \mathcal{T}_{\boldsymbol{x},\mathrm{ik}}^{*} : \boldsymbol{y} \mapsto \left(-2\pi\mathrm{i}\hat{d}_{\boldsymbol{k}}\boldsymbol{y}^{\top}\boldsymbol{k}\sum_{\boldsymbol{r}\in\mathbb{Z}^3}c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\,\mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{x}}\right)_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}$$

or

$$\mathcal{T}_{\boldsymbol{x}}^{*} \approx \mathcal{T}_{\boldsymbol{x},\mathrm{ad}}^{*} : \boldsymbol{y} \mapsto \left(-2\pi\mathrm{i}\hat{d}_{\boldsymbol{k}}\sum_{\boldsymbol{r}\in\mathbb{Z}^3}\boldsymbol{y}^{\top}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\,\mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{x}}\right)_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}},$$

respectively. In case of i$\boldsymbol{k}$-differentiation we easily compute

$$\begin{aligned}\int_{[-1/2,1/2)^3}\|\mathcal{T}_{\boldsymbol{x}}^{*}\boldsymbol{y} - \mathcal{T}_{\boldsymbol{x},\mathrm{ik}}^{*}\boldsymbol{y}\|^2\mathrm{d}\boldsymbol{x} &= 4\pi^2\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}|\boldsymbol{y}^{\top}\boldsymbol{k}|^2(1-\hat{d}_{\boldsymbol{k}}c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}}))^2\\ &\quad + 4\pi^2\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}\sum_{\boldsymbol{r}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}}|\boldsymbol{y}^{\top}\boldsymbol{k}|^2\hat{d}_{\boldsymbol{k}}^2c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}}),\end{aligned}$$

for which again the optimal deconvolution coefficients are given by (3.17).

The approximation error for the analytic differentiation reads as

$$\int_{\mathbb{T}^3} \|\mathcal{T}_{\boldsymbol{x}}^* - \mathcal{T}_{\boldsymbol{x},\mathrm{ad}}^*\|^2 \mathrm{d}\boldsymbol{x} = 4\pi^2 \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} |\boldsymbol{y}^\top \boldsymbol{k}|^2 (1 - \hat{d}_{\boldsymbol{k}} c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t}))^2 + 4\pi^2 \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \sum_{\boldsymbol{r}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}} |\boldsymbol{y}^\top (\boldsymbol{k} + \boldsymbol{r}\odot \boldsymbol{M}_\mathrm{o})|^2 \hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_\mathrm{o}}^2(\tilde{\varphi}_\mathrm{t}).$$

Consequently, the optimal deconvolution coefficients, which may again be determined by differentiating with respect to $\hat{d}_{\boldsymbol{k}}$ for each $\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}$, depend on $\boldsymbol{y} \in \mathbb{C}^3$. Considering the worst case error

$$\int_{\mathbb{T}^3} \|\mathcal{T}_{\boldsymbol{x}}^* - \mathcal{T}_{\boldsymbol{x},\mathrm{ad}}^*\|^2 \mathrm{d}\boldsymbol{x} \le 4\pi^2 \|\boldsymbol{y}\|^2 \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \|\boldsymbol{k}\|^2 (1 - \hat{d}_{\boldsymbol{k}} c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t}))^2 + 4\pi^2 \|\boldsymbol{y}\|^2 \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \sum_{\boldsymbol{r}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}} \|\boldsymbol{k} + \boldsymbol{r}\odot \boldsymbol{M}_\mathrm{o}\|^2 \hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_\mathrm{o}}^2(\tilde{\varphi}_\mathrm{t}),$$

the optimal deconvolution coefficients are the same as found for the gradient NFFT with analytic differentiation, see (3.28). □

3.3 Window functions and error estimates

The basic concept of the NFFT algorithm was introduced in Section 3.2.1. A key element of the method is the so called window function, which is used to map the arbitrarily spaced knots onto a regular grid. In three dimensions we use a tensor product scheme, see (3.8), to define the window function. Thus, the Fourier coefficients of the periodization $\tilde{\varphi}$ have the form

$$c_{\boldsymbol{k}}(\tilde{\varphi}) = \prod_{j=1}^{3} c_{k_j}(\tilde{\varphi}_j), \quad \boldsymbol{k} \in \mathbb{Z}^3. \tag{3.34}$$

Also the derivatives are obtained easily based on the one-dimensional case via

$$\left(\frac{\partial}{\partial x_1}\right)^{n_1} \left(\frac{\partial}{\partial x_2}\right)^{n_2} \left(\frac{\partial}{\partial x_3}\right)^{n_3} \varphi(\boldsymbol{x}) = \prod_{j=1}^{3} \left(\frac{\partial}{\partial x_j}\right)^{n_j} \varphi(x_j).$$

The evaluation of the underlying sums is only able in an efficient way if the window function is well localized in spatial as well as in Fourier domain. In case that the window function is not compactly supported we use a truncated version with

$$\operatorname{supp}(\varphi_\mathrm{t}) = \bigotimes_{j=1}^{3} [-{}^{m}\!/_{\sigma_j M_j}, {}^{m}\!/_{\sigma_j M_j}], \quad m \ll M_j \tag{3.35}$$

instead, see (3.10). In the following we introduce possible window functions. In order to estimate the approximation errors as introduced in (3.18) and (3.19), we consider the one-dimensional case in the following and derive estimates of the error sums

$$\sum_{r\in\mathbb{Z}\setminus\{0\}} \frac{c_{k+r\sigma_j M_j}^2(\tilde{\varphi}_{\mathrm{t},j})}{c_k^2(\tilde{\varphi}_{\mathrm{t},j})} \tag{3.36}$$

for the different window functions. The multivariate case is revisited in Section 3.3.2.

B-spline

We define the centered cardinal B-splines via

$$B_1(x) = \begin{cases} 1 & : x \in [-1/2, 1/2), \\ 0 & : \text{else}, \end{cases}$$

and for $n \in \mathbb{N}$, $n > 1$,

$$B_n(x) := B_{n-1}(x) * B_1(x) = \int_{-1/2}^{1/2} B_{n-1}(x-t)\,\mathrm{d}t.$$

The number $n \in \mathbb{N}$ is called order of the B-spline and determines the size of the support, which is $\operatorname{supp}(B_n) = [-n/2, n/2]$. Within the three dimensional NFFT the B-spline window function is commonly defined via

$$\varphi(\boldsymbol{x}) := \prod_{j=1}^{3} B_{2m}(\sigma_j M_j x_j), \tag{3.37}$$

cf. [Bey95, Ste98, PS03]. Note that the introduced B-spline window is inherently compactly supported, i.e., $\varphi_{\mathrm{t}} = \varphi$, and that the parameter $m \in \mathbb{N}$ equals the NFFT support parameter, see (3.10) and (3.35).

The Fourier coefficients of the periodic function $\tilde{\varphi}$ with respect to the single dimensions $j \in \{1, 2, 3\}$ are given by

$$c_k(\tilde{\varphi}_j) = \frac{1}{\sigma_j M_j} \operatorname{sinc}^{2m}\left(\frac{\pi k}{\sigma_j M_j}\right), \quad k \in \mathbb{Z},$$

where we denote by

$$\operatorname{sinc}(x) := \begin{cases} \frac{\sin x}{x} & : x \neq 0, \\ 1 & : x = 0, \end{cases}$$

the well-known cardinal sine function. The Fourier coefficients of the window function in three dimensions are obtained via the tensor product approach based on the one-dimensional case, see (3.34).

The one-dimensional error sums (3.36) can be estimated as follows.

Theorem 3.5. *For the B-spline window as defined in* (3.37) *we have*

$$\sum_{r \in \mathbb{Z}\setminus\{0\}} \frac{c^2_{k+r\sigma_j M_j}(\tilde{\varphi}_j)}{c^2_k(\tilde{\varphi}_j)} = \sum_{r \in \mathbb{Z}\setminus\{0\}} \left(\frac{\frac{k}{\sigma_j M_j}}{\frac{k}{\sigma_j M_j} + r}\right)^{4m} \leq \frac{8m}{4m-1} \left(\frac{\frac{|k|}{\sigma_j M_j}}{\frac{|k|}{\sigma_j M_j} - 1}\right)^{4m} =: s_j(k), \tag{3.38}$$

Proof. The result is obtained via estimating the infinite sum from above by an integral, see [Ste98]. ■

Remark 3.6 (Arbitrary precision)**.** It can be seen easily that for $\sigma_j = 1$, i.e., oversampling is not allowed, we may not always be able to achieve an arbitrary desired accuracy. As an example, for $k = M_j/2$ and $\sigma_j = 1$ we obtain

$$\sum_{r \in \mathbb{Z}\setminus\{0\}} \left(\frac{\frac{k}{\sigma_j M_j}}{\frac{k}{\sigma_j M_j} + r}\right)^{4m} = \sum_{r \in \mathbb{Z}\setminus\{0\}} \left(\frac{\frac{1}{2}}{\frac{1}{2} + r}\right)^{4m} \geq \left(\frac{\frac{1}{2}}{\frac{1}{2} - 1}\right)^{4m} = 1,$$

i.e., we can not force the error sum to approach the value 0.

On the other hand, an arbitrary precision can always be achieved by allowing $\sigma_j > 1$, since

$$s_j(k) \overset{|k|=M_j/2}{\leq} \frac{8m}{4m-1}\left(\frac{1}{2\sigma_j - 1}\right)^{4m}.$$

The upper limit tends to zero with growing m for fixed $\sigma_j > 1$. □

For the (adjoint) gradient as well as for the Hessian NFFT we need the derivatives of the window function in case of analytic differentiation. The derivatives are given by

$$\begin{aligned}\varphi_j'(x) &= \sigma_j M_j \cdot \left[B_{2m-1}(\sigma_j M_j x + 1/2) - B_{2m-1}(\sigma_j M_j x - 1/2)\right],\\ \varphi_j''(x) &= \sigma_j^2 M_j^2 \cdot \left[B_{2m-2}(\sigma_j M_j x + 1) - 2B_{2m-2}(\sigma_j M_j x) + B_{2m-2}(\sigma_j M_j x - 1)\right],\end{aligned}$$

see also [Pip15].

Modified B-spline

We introduce a shape parameter $b \in 1/2\mathbb{N}$ and define the modified B-spline window via

$$\varphi(\boldsymbol{x}) := \prod_{j=1}^{3} B_{2b}\left(\frac{\sigma_j M_j x_j}{m}\right),$$

cf. [Nes16a]. The support is the same as for the standard B-spline window, but the order $2b \in \mathbb{N}$ is allowed to vary. For the Fourier coefficients in the single dimensions $j \in \{1, 2, 3\}$ we obtain

$$c_k(\tilde{\varphi}_j) = \frac{m}{\sigma_j M_j b}\operatorname{sinc}^{2b}\left(\frac{m\pi k}{\sigma_j M_j b}\right).$$

Obviously, the case $b = m$ corresponds to the classical B-spline window, as introduced before.

For $\sigma_j = 1$ we have for $k \in \{-M_j/2, \ldots, M_j/2\}$

$$\frac{m\pi k}{\sigma_j M_j b} \in \left[-\frac{m\pi}{2b}, \frac{m\pi}{2b}\right] \subseteq (-\pi, \pi) \iff b > \frac{m}{2},$$

i.e., for $b \leq m/2$ some of the Fourier coefficients might be equal or very close to zero, so that the usage of the deconvolution coefficients (3.14) will lead to numerical instabilities. Thus, we will restrict our considerations to the case $b > m/2$ in the following.

Note that different shape parameters b_j may be applied in the single dimensions, but for the sake of simplicity we assume the same shape parameter b to be applied in all dimensions. This also applies for the Bessel and the Gaussian window function, which are introduced later on.

Theorem 3.7. *For $R \in \mathbb{N}$ we have for the modified B-spline window*

$$\sum_{|r|>R} c^2_{k+r\sigma_j M_j}(\tilde{\varphi}_j) \leq \frac{m^2}{\sigma_j^2 M_j^2 b^2} \frac{\left(\frac{k}{\sigma_j M_j} + R\right)^{1-4b} - \left(\frac{k}{\sigma_j M_j} - R\right)^{1-4b}}{\left(\frac{m\pi}{b}\right)^{4b}(4b-1)}.$$

Proof. We obtain

$$\sum_{|r|>R} c^2_{k+r\sigma_j M_j}(\tilde{\varphi}_j) = \frac{m^2}{\sigma_j^2 M_j^2 b^2} \sum_{|r|>R} \operatorname{sinc}^{4b}\left(\tfrac{m\pi}{b}\left[\tfrac{k}{\sigma_j M_j} + r\right]\right)$$
$$\leq \frac{m^2}{\sigma_j^2 M_j^2 b^2} \sum_{|r|>R} \frac{1}{\left(\frac{m\pi}{b}\right)^{4b}\left(\frac{k}{\sigma_j M_j} + r\right)^{4b}}.$$

Since $k/\sigma_j M_j \in [-1/2, 1/2)$ and $|r| \geq 1$ the expression

$$\frac{1}{\left(\frac{k}{\sigma_j M_j} + r\right)^{4b}} + \frac{1}{\left(\frac{k}{\sigma_j M_j} - r\right)^{4b}}$$

is monotonically decreasing in $r \in \mathbb{N}$. Thus we can estimate the series from above by an integral as follows.

$$\sum_{|r|>R} \frac{1}{\left(\frac{k}{\sigma_j M_j} + r\right)^{4b}} < \int_R^\infty \frac{\mathrm{d}r}{\left(\frac{k}{\sigma_j M_j} + r\right)^{4b}} + \int_R^\infty \frac{\mathrm{d}r}{\left(\frac{k}{\sigma_j M_j} - r\right)^{4b}}$$
$$= \frac{1}{4b-1}\left(\frac{1}{\left(\frac{k}{\sigma_j M_j} + R\right)^{4b-1}} - \frac{1}{\left(\frac{k}{\sigma_j M_j} - R\right)^{4b-1}}\right).$$

■

Corollary 3.8. *In summary, we can write in case of the modified B-spline window*

$$\sum_{r\in\mathbb{Z}\setminus\{0\}} \frac{c^2_{k+r\sigma_j M_j}(\tilde{\varphi}_j)}{c_k^2(\tilde{\varphi}_j)} \leq s_j(k)$$

with

$$s_j(k) := \sum_{0<|r|\leq R_k} \frac{c^2_{k+r\sigma_j M_j}(\tilde{\varphi}_j)}{c_k^2(\tilde{\varphi}_j)} + \frac{1}{\operatorname{sinc}^{4b}\left(\frac{m\pi k}{\sigma_j M_j b}\right)} \frac{\left(\frac{k}{\sigma_j M_j} + R_k\right)^{1-4b} - \left(\frac{k}{\sigma_j M_j} - R_k\right)^{1-4b}}{\left(\frac{m\pi}{b}\right)^{4b}(4b-1)} \tag{3.39}$$

for some $R_k \in \mathbb{N}$.

In our numerical tests, see Example 3.9 and Section 3.3.1, we compute for each k the upper bound $s_j(k)$, where we start with $R_k := 1$ and increase R_k step by step until the remainder of the series is comparatively small.

Example 3.9. In Figure 3.1 we plot the estimates $s_j(k)$, as given in (3.39), with respect to $|k|/\sigma_j M_j \in [0, 1/2]$, where we set $m := 4$. We consider different shape parameters $b > m/2 = 2$, $b \in 1/2\mathbb{N}$. For $b = m$ we plot the estimate (3.38), instead.

We have $\lim_{k\to 0} s_j(k) = 0$ only for $b = m$. Consequently, the standard B-spline window will be the optimal choice in case that the terms for $k \approx 0$ are dominating the approximation error, cf. (3.18). This is especially the case if the Fourier coefficients $\hat{f}_k$ tend to zero very fast. If all coefficients $\hat{f}_k$, $k \in \mathcal{I}_{\boldsymbol{M}}$, are of a comparable size, we expect that a shape parameter $b < m$ gives the best results. This has been verified by numerical tests, see Section 3.3.1. □

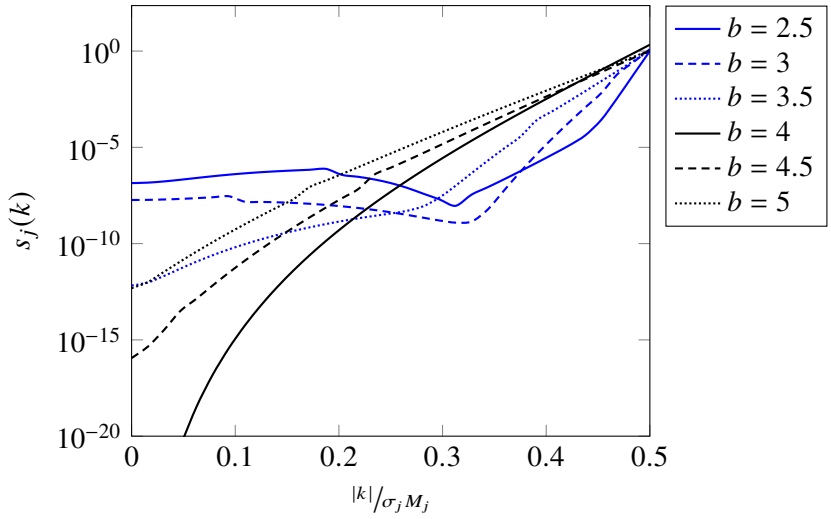


Figure 3.1: Upper bounds $s_j(k)$ for the modified B-spline window with respect to $|k|/\sigma_j M_j \in [0, 1/2]$. We applied the support parameter $m = 4$ and different shape parameters $b \in 1/2\mathbb{N}$.

Bessel

For some shape parameter $b > 0$ we define the Bessel window function via

$$\varphi(\boldsymbol{x}) := \prod_{j=1}^{3} \frac{1}{2} I_0\left(b\sqrt{m^2 - \sigma_j^2 M_j^2 x_j^2}\right) \cdot \chi_{[-m/\sigma_j M_j, m/\sigma_j M_j]}(x_j)$$

cf. [Kai66, Fou03, JMNM91, Pip15], where we denote by I_n the modified Bessel function

$$I_n(x) := \frac{1}{\pi} \int_0^\pi \mathrm{e}^{x \cos t} \cos(nt)\,\mathrm{d}t$$

of order $n \in \mathbb{N}_0$. Note that the window function is by definition already compactly supported with support (3.35).

The Fourier coefficients of its periodization are given by

$$c_k(\tilde{\varphi}_j) = \frac{1}{\sigma_j M_j} \begin{cases} \dfrac{\sinh\left(m\sqrt{b^2 - 4\pi^2 k^2/(\sigma_j^2 M_j^2)}\right)}{\sqrt{b^2 - 4\pi^2 k^2/(\sigma_j^2 M_j^2)}} & : |k| \le \frac{\sigma_j M_j b}{2\pi}, \\ m\,\operatorname{sinc}\left(m\sqrt{4\pi^2 k^2/(\sigma_j^2 M_j^2) - b^2}\right) & : \text{else.} \end{cases} \tag{3.40}$$

Theorem 3.10. *For some $R \in \mathbb{N}$, $R > \frac{|k|}{\sigma_j M_j} + \frac{b}{2\pi}$, we have for the Bessel window function*

$$\sum_{|r|>R} c^2_{k+r\sigma_j M_j}(\tilde{\varphi}_j) \le \frac{1}{\sigma_j^2 M_j^2} \frac{\ln\left|\frac{2\pi(|k|/\sigma_j M_j - R) - b}{2\pi(|k|/\sigma_j M_j - R) + b}\right| + \ln\left|\frac{2\pi(|k|/\sigma_j M_j + R) + b}{2\pi(|k|/\sigma_j M_j + R) - b}\right|}{4\pi b}.$$

Proof. If we choose $R > \frac{|k|}{\sigma_j M_j} + \frac{b}{2\pi}$, we obtain

$$\begin{aligned}\sum_{|r|>R} c^2_{k+r\sigma_j M_j}(\tilde{\varphi}_j) &= \frac{m^2}{\sigma_j^2 M_j^2} \sum_{r=R+1}^{\infty} \operatorname{sinc}^2\left(m\sqrt{4\pi^2(|k|/\sigma_j M_j \pm r)^2 - b^2}\right)\\ &\leq \frac{1}{\sigma_j^2 M_j^2} \sum_{r=R+1}^{\infty} \frac{1}{4\pi^2(|k|/\sigma_j M_j \pm r)^2 - b^2}\\ &< \frac{1}{\sigma_j^2 M_j^2} \int_R^{\infty} \frac{\mathrm{d}r}{4\pi^2(|k|/\sigma_j M_j \pm r)^2 - b^2},\end{aligned}$$

where

$$\begin{aligned}\int_R^{\infty} \frac{\mathrm{d}r}{4\pi^2(|k|/\sigma_j M_j \pm r)^2 - b^2} &= \frac{1}{2\pi}\left(\int_{2\pi(|k|/\sigma_j M_j + R)}^{\infty} + \int_{-\infty}^{2\pi(|k|/\sigma_j M_j - R)}\right)\frac{\mathrm{d}r}{r^2 - b^2}\\ &= \frac{\ln\left|\frac{2\pi(|k|/\sigma_j M_j - R)-b}{2\pi(|k|/\sigma_j M_j - R)+b}\right| + \ln\left|\frac{2\pi(|k|/\sigma_j M_j + R)+b}{2\pi(|k|/\sigma_j M_j + R)-b}\right|}{4\pi b}.\end{aligned}$$

■

Corollary 3.11. *In summary, we can write for the Bessel window function*

$$\sum_{r\in\mathbb{Z}\setminus\{0\}} \frac{c^2_{k+r\sigma_j M_j}(\tilde{\varphi}_j)}{c_k^2(\tilde{\varphi}_j)} \leq s_j(k)$$

with

$$s_j(k) := \sum_{0<|r|\leq R_k} \frac{c^2_{k+r\sigma_j M_j}(\tilde{\varphi}_j)}{c_k^2(\tilde{\varphi}_j)} + \frac{1}{c_k^2(\tilde{\varphi}_j)} \frac{\ln\left|\frac{2\pi(|k|/\sigma_j M_j - R)-b}{2\pi(|k|/\sigma_j M_j - R)+b}\right| + \ln\left|\frac{2\pi(|k|/\sigma_j M_j + R)+b}{2\pi(|k|/\sigma_j M_j + R)-b}\right|}{4\pi b\sigma_j^2 M_j^2} \tag{3.41}$$

for some $R_k \in \mathbb{N}$, $R_k > \frac{|k|}{\sigma_j M_j} + \frac{b}{2\pi}$.

In case of analytic differentiation for the different NFFT modules, as presented in Section 3.2.2, we need the first as well as the second derivatives. For the Bessel function I_0 we get

$$\begin{aligned}I_0'(x) &= \frac{1}{\pi}\int_0^{\pi} \mathrm{e}^{x\cos t}\cos t\,\mathrm{d}t = I_1(x),\\ I_0''(x) &= \frac{1}{\pi}\int_0^{\pi} \mathrm{e}^{x\cos t}\cos^2 t\,\mathrm{d}t = \frac{1}{\pi}\int_0^{\pi} \mathrm{e}^{x\cos t}\,\mathrm{d}t - \frac{1}{\pi}\int_0^{\pi} \mathrm{e}^{x\cos t}\sin t\cdot\sin t\,\mathrm{d}t = I_0(x) - \frac{I_1(x)}{x},\end{aligned}$$

by partial integration. For the Bessel window function as defined above we obtain for the single dimensions

$$\varphi_j'(x) = \begin{cases} -\dfrac{I_1(br_j(x))}{2r_j(x)}\cdot b\sigma_j^2 M_j^2 x & : |x| \neq {}^{m}/_{\sigma_j M_j},\\ \mp\dfrac{b^2 m\sigma_j M_j}{4} & : x = \pm{}^{m}/_{\sigma_j M_j},\end{cases}$$

where we set $r_j(x) := \sqrt{m^2 - \sigma_j^2 M_j^2 x^2}$, see also [Pip15]. The limit $\lim_{x\to 0}\frac{I_1(x)}{x} = \frac{1}{2}$ is easily obtained via l'Hospital. Note that the function φ_j is only differentiable on the interval $(-{}^{m}/_{\sigma_j M_j}, -{}^{m}/_{\sigma_j M_j})$. The

obtained derivative can be extended to a continuous function on the closed interval, since the one-sided limits exist.

For the second derivative we get in an analogue manner

$$\varphi_j''(x) = \begin{cases} \dfrac{b\sigma_j^2 M_j^2}{2r_j(x)^2}\left(b\sigma_j^2 M_j^2 x^2 I_0(br_j(x)) - \dfrac{I_1(br_j(x))}{r_j(x)}\left(\sigma_j^2 M_j^2 x^2 + m^2\right)\right) & : |x| \neq {}^{m}/_{\sigma_j M_j}, \\ \dfrac{b^4 m^2 \sigma_j^2 M_j^2}{16} - \dfrac{b^2 \sigma_j^2 M_j^2}{4} & : |x| = {}^{m}/_{\sigma_j M_j}. \end{cases}$$

Note that we can rewrite the second derivative for $|x| \neq {}^{m}/_{\sigma_j M_j}$ as

$$\frac{b\sigma_j^2 M_j^2}{2}\left(-bI_0(br_j(x)) + \frac{bI_1(br_j(x))}{br_j(x)} + b^3 m^2 \frac{I_0(br_j(x))}{b^2 r_j(x)^2} - 2b^3 m^2 \frac{I_1(br_j(x))}{b^3 r_j(x)^3}\right)$$

and in order to compute the limit for $x \to \pm {}^{m}/_{\sigma_j M_j}$ it remains to calculate

$$\lim_{x\to 0} \frac{I_0(x) - \frac{2}{x} I_1(x)}{x^2} = \lim_{x\to 0} \frac{\frac{1}{\pi}\int_0^\pi \mathrm{e}^{x\cos t}(1 - 2\sin^2 t)\mathrm{d}t}{x^2} = \frac{1}{8},$$

which is again obtained via l'Hospital.

Remark 3.12. As already pointed out in Remark 3.3, the presented error estimates for the analytic differentiation approach are only valid for differentiable window functions. This is due to the fact that the relation

$$c_k(\tilde{\varphi}_j') = 2\pi \mathrm{i} k \, c_k(\tilde{\varphi}_j),$$

which is exploited in order to obtain (3.27), does not hold in case of non differentiable functions.

Analytic differentiation is nevertheless applicable. In case of the Bessel window function we obtain by partial integration

$$c_k(\tilde{\varphi}_j') = -\mathrm{i}\sin\left(\frac{2\pi k m}{\sigma_j M_j}\right) + 2\pi \mathrm{i} k \, c_k(\tilde{\varphi}_j).$$

Since the partial derivatives are still piecewise continuous, the Fourier coefficients tend to zero as well, which is a consequence of the well-known Riemann-Lebesgue lemma. Indeed, in case of the Bessel window function we have, cf. (3.40),

$$c_k(\tilde{\varphi}_j) \sim \frac{1}{2\pi k}\sin\left(\frac{2\pi k m}{\sigma_j M_j}\right) \quad \text{as } k \to \infty,$$

and thus $c_k(\tilde{\varphi}_j') \to 0$ as $k \to \infty$. □

Window functions without compact support

In the case that the window function is not compactly supported we simply replace the function by a truncated version so that the same error formulas are valid, see (3.18) and (3.19).

As an example we consider the Gaussian window function

$$\varphi(\boldsymbol{x}) = \prod_{j=1}^{3} \frac{1}{\sqrt{\pi b}} \, \mathrm{e}^{-\sigma_j^2 M_j^2 x_j^2 / b}.$$

The Fourier coefficients in the single dimensions are given by

$$c_k(\tilde{\varphi}_j) = \frac{1}{\sigma_j M_j} \mathrm{e}^{-b\pi^2 k^2/(\sigma_j^2 M_j^2)},$$

where $b > 0$ denotes again the so called shape parameter.

In case of the truncated version, see (3.10), the Fourier coefficients read as

$$c_k(\tilde{\varphi}_{\mathrm{t},j}) = \frac{1}{\sqrt{\pi b}} \int_{-m/\sigma_j M_j}^{m/\sigma_j M_j} \mathrm{e}^{-\sigma_j^2 M_j^2 x^2/b} \mathrm{e}^{-2\pi \mathrm{i} k x} \mathrm{d}x = \frac{\mathrm{e}^{-b\pi^2 k^2/(\sigma_j^2 M_j^2)}}{\sigma_j M_j} \mathrm{Re}\left[\mathrm{erf}\left(\frac{m}{\sqrt{b}} + \mathrm{i}\frac{\pi k \sqrt{b}}{\sigma_j M_j}\right)\right]$$

and tend to zero in k only with order 1.

The estimation of the error sums

$$\sum_{r \in \mathbb{Z}\setminus\{0\}} c^2_{k+r\sigma_j M_j}(\tilde{\varphi}_{\mathrm{t},j})$$

proves to be particularly difficult due to the presence of the complex valued error function. For other not compactly supported window functions we expect similarly complicated expressions.

In our numerical examples we simply set

$$\sum_{r \in \mathbb{Z}\setminus\{0\}} \frac{c^2_{k+r\sigma_j M_j}(\tilde{\varphi}_{\mathrm{t},j})}{c^2_k(\tilde{\varphi}_{\mathrm{t},j})} \approx \sum_{0<|r|\le R_k} \frac{c^2_{k+r\sigma_j M_j}(\tilde{\varphi}_{\mathrm{t},j})}{c^2_k(\tilde{\varphi}_{\mathrm{t},j})} =: s_j(k), \tag{3.42}$$

i.e., we truncate the sums at $|r| = R_k$, where we choose $R_k \in \mathbb{N}$ large enough.

Remark 3.13. Commonly, see [PST01] for instance, the NFFT deconvolution step is performed by using the Fourier coefficients $c_k(\tilde{\varphi}_j)$ of the non truncated function, i.e., we set $\hat{d}_{\boldsymbol{k}} := c_{\boldsymbol{k}}^{-1}(\tilde{\varphi})$ in (3.16) and obtain

$$\|f - f_{\approx}\|^2 = \sum_{\boldsymbol{k} \in I_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \left[\left(1 - \frac{c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})}{c_{\boldsymbol{k}}(\tilde{\varphi})}\right)^2 + \sum_{\boldsymbol{r} \in \mathbb{Z}^3\setminus\{\boldsymbol{0}\}} \frac{c^2_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})}{c^2_{\boldsymbol{k}}(\tilde{\varphi})}\right].$$

In other words, we obtain a fundamentally different representation of the approximation error, compared to (3.18), if the Fourier coefficients of the non truncated Gaussian are used. □

Remark 3.14. Alternatively, we may separate the resulting approximation error as follows. We have

$$\begin{aligned} f_{\approx}(\boldsymbol{x}) &= \sum_{\boldsymbol{l} \in I_{\boldsymbol{M}_{\mathrm{o}}}} g_{\boldsymbol{l}} \tilde{\varphi}(\boldsymbol{x} - \boldsymbol{l} \odot \boldsymbol{M}_{\mathrm{o}}^{-1}) + \sum_{\boldsymbol{l} \in I_{\boldsymbol{M}_{\mathrm{o}}}} g_{\boldsymbol{l}} (\tilde{\varphi}_{\mathrm{t}} - \tilde{\varphi})(\boldsymbol{x} - \boldsymbol{l} \odot \boldsymbol{M}_{\mathrm{o}}^{-1}) \\ &= \sum_{\boldsymbol{k} \in I_{\boldsymbol{M}}} \hat{d}_{\boldsymbol{k}} \hat{f}_{\boldsymbol{k}} \sum_{\boldsymbol{r} \in \mathbb{Z}^3} c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}) \mathrm{e}^{2\pi \mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\mathsf{T}}\boldsymbol{x}} + \sum_{\boldsymbol{l} \in I_{\boldsymbol{M}_{\mathrm{o}}}} g_{\boldsymbol{l}} (\tilde{\varphi}_{\mathrm{t}} - \tilde{\varphi})(\boldsymbol{x} - \boldsymbol{l} \odot \boldsymbol{M}_{\mathrm{o}}^{-1}) \end{aligned}$$

and thus

$$f(\boldsymbol{x}) - f_{\approx}(\boldsymbol{x}) = E_{\mathrm{a}}(\boldsymbol{x}) + E_{\mathrm{t}}(\boldsymbol{x}),$$

cf. [Ste98], where we define the aliasing error

$$E_{\mathrm{a}}(\boldsymbol{x}) := \sum_{\boldsymbol{k} \in I_{\boldsymbol{M}}} \hat{f}_{\boldsymbol{k}} \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^{\mathsf{T}}\boldsymbol{x}} - \sum_{\boldsymbol{k} \in I_{\boldsymbol{M}}} \hat{d}_{\boldsymbol{k}} \hat{f}_{\boldsymbol{k}} \sum_{\boldsymbol{r} \in \mathbb{Z}^3} c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}) \mathrm{e}^{2\pi \mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\mathsf{T}}\boldsymbol{x}}$$

as well as the truncation error

$$E_{\mathrm{t}}(\boldsymbol{x}) := \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}} g_{\boldsymbol{l}}(\tilde{\varphi} - \tilde{\varphi}_{\mathrm{t}})(\boldsymbol{x} - \boldsymbol{l} \odot \boldsymbol{M}_{\mathrm{o}}^{-1}).$$

The Fourier space representation of the aliasing sum is favorable for the computation of the $\mathcal{L}_2$-norm, which is given by

$$\|E_{\mathrm{a}}\|_{\mathcal{L}_2}^2 = \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \left(1 - \hat{d}_{\boldsymbol{k}} c_{\boldsymbol{k}}(\tilde{\varphi})\right)^2 + \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \sum_{\boldsymbol{r} \in \mathbb{Z}^3 \setminus \{\boldsymbol{0}\}} \hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}+\boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}),$$

see the proof of Lemma 3.1.

In order to predict the aliasing error, we need to estimate the sums

$$\sum_{\boldsymbol{r} \in \mathbb{Z}^3 \setminus \{\boldsymbol{0}\}} c_{\boldsymbol{k}+\boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}) = \frac{1}{|\mathcal{I}_{\boldsymbol{M}_{\mathrm{o}}}|} \sum_{\boldsymbol{r} \in \mathbb{Z}^3 \setminus \{\boldsymbol{0}\}} \mathrm{e}^{-b\pi^2 \boldsymbol{k} \odot \boldsymbol{M}_{\mathrm{o}}^{-1}},$$

which is possible based on estimating the infinite sum by an integral, see [Ste98]. However, in order to predict the total error, we need to accurately estimate the truncation error E_{t} in the $\mathcal{L}_2$-norm, which is by no means straight forward. We remark that the truncation error is typically estimated by using the infinity norm, see [Ste98] for instance. □

3.3.1 One-dimensional setting

We presented estimates of the form

$$\sum_{r \in \mathbb{Z} \setminus \{0\}} \frac{c_{k+r\sigma M}^2(\tilde{\varphi})}{c_k^2(\tilde{\varphi})} \leq s(k) \tag{3.43}$$

for the different window functions, see (3.38), (3.39), (3.41) and (3.42). Note that we omit the indices j in this section since we are in a one-dimensional setting, i.e., we consider the approximation of univariate trigonometric polynomials

$$f(x) = \sum_{k=-M/2}^{M/2-1} \hat{f}_k \, \mathrm{e}^{2\pi \mathrm{i} k x}.$$

Of course, the work flows of the Algorithms 3.1 (NFFT) and 3.2 (adjoint NFFT) as well as the error estimates, cf. (3.18) and (3.19), are basically the same. We obtain the following estimates by making use of the upper limits $s(k)$.

$$\Delta f^2 := \|f - f_{\approx}\|_{\mathcal{L}_2}^2 \leq \begin{cases} \displaystyle\sum_{k=-M/2}^{M/2-1} |\hat{f}_k|^2 s(k) & \text{for } \hat{d}_k := \dfrac{1}{c_k(\tilde{\varphi}_{\mathrm{t}})}, \quad (3.44) \\ \displaystyle\sum_{k=-M/2}^{M/2-1} |\hat{f}_k|^2 \frac{s(k)}{1+s(k)} & \text{for } \hat{d}_k := \dfrac{c_k(\tilde{\varphi}_{\mathrm{t}})}{\sum\limits_{r \in \mathbb{Z}} c_{k+r\sigma M}^2(\tilde{\varphi}_{\mathrm{t}})}. \quad (3.45) \end{cases}$$

The second estimate is obtained based on the fact that the function $y(x) := x \cdot (1+x)^{-1}$ is monotonically increasing for $x := c_k^{-2}(\tilde{\varphi}) \sum_{r \in \mathbb{Z} \setminus \{0\}} c_{k+r\sigma M}^2(\tilde{\varphi}) \geq 0$.

In the following examples we consider two different sets of Fourier coefficients, namely

$$\hat{f}_k := \frac{1}{1+k^2}, \quad k \in \mathcal{I}_{64}, \tag{3.46}$$

and

$$\hat{f}_k := \mathrm{e}^{-(k/5)^2}, \quad k \in \mathcal{I}_{64}. \tag{3.47}$$

Note that the coefficients (3.47) tend to zero exponentially fast for growing $|k|$, i.e., for large values of $|k|$ the factors $\hat{f}_k$ only have an insignificant influence on the overall error. In contrast, the coefficients (3.46) tend to zero rather slowly.

We used a simple implementation of the univariate NFFT in MATLAB in order to verify the theoretical estimates. In order to measure the actual errors, we considered $N = 500$ randomly chosen nodes $x_j \in [-1/2, 1/2)$ and computed

$$\Delta f^2 \approx \frac{1}{N} \sum_{j=1}^{N} \left[f(x_j) - f_{\approx}(x_j) \right]^2, \tag{3.48}$$

where $f(x_j)$ are the exact values of the trigonometric polynomial and $f_{\approx}(x_j)$ denote the approximate values computed by the univariate NFFT. This can be understood as a simple Monte Carlo ansatz [KW86] for the approximation of the integral $\|f - f_{\approx}\|^2$.

In case of the modified B-spline, the Bessel or the Gaussian window function, the choice of the shape parameter b influences the resulting accuracy a lot. The shape parameter may be chosen appropriately based on the above discussed error estimates, which is also considered in the following numerical examples. Instead of minimizing a worst case error, see [Ste98, DS99, PS03, GL04, Jac09] for instance, we optimize the shape of the window function based on the given Fourier coefficients. The considered numerical examples have already been presented in [Nes16a].

Remark 3.15 (Relative errors)**.** Commonly, one is also interested in a relative error measure. In case of estimating the error in the $\mathcal{L}_2$-norm we may take into consideration the relative error

$$\Delta f_{\mathrm{rel}}^2 := \frac{\|f - f_{\approx}\|_{\mathcal{L}_2}^2}{\|f\|_{\mathcal{L}_2}^2},$$

where we can compute $\|f\|_{\mathcal{L}_2}^2$ via

$$\|f\|_{\mathcal{L}_2}^2 = \sum_{k=-M/2}^{M/2-1} \hat{f}_k^2,$$

provided that the Fourier coefficients $\hat{f}_k$ are known. For the two considered sets of Fourier coefficients we obtain

$$\|f\|_{\mathcal{L}_2}^2 \approx 1.61 \iff \Delta f_{\mathrm{rel}}^2 \approx 6.20 \cdot 10^{-1} \|f - f_{\approx}\|_{\mathcal{L}_2}^2 \quad \text{for } \hat{f}_k \text{ defined via (3.46)},$$

and

$$\|f\|_{\mathcal{L}_2}^2 \approx 2.80 \iff \Delta f_{\mathrm{rel}}^2 \approx 3.57 \cdot 10^{-1} \|f - f_{\approx}\|_{\mathcal{L}_2}^2 \quad \text{for } \hat{f}_k \text{ defined via (3.47)}.$$

Thus, in order to estimate the resulting relative approximation errors we have to multiply the absolute errors, as presented in the following examples, by a factor of $6.2 \cdot 10^{-1}$ or $3.57 \cdot 10^{-1}$, respectively. □

B-spline

In Figure 3.2 we plot the estimated errors (3.44) and (3.45) for the B-spline window, where we considered different support parameters m and oversampling factors σ. The estimates $s(k)$ were computed via (3.38). We also plot the measured errors (3.48), obtained by applying the standard deconvolution approach. Note that the estimates regarding the two deconvolution approaches do not significantly differ. A small difference is obtained for the slowly decreasing coefficients (3.47) in case that no oversampling is applied, i.e., $\sigma = 1$.

It can be seen that the measured errors match very well with the computed estimates. A higher accuracy is achieved by using a larger support parameter m, i.e., a B-spline of a higher order, or by applying an oversampling factor $\sigma > 1$. Note that in case of the fast decreasing Fourier coefficients (3.47) the errors are already very small even without oversampling.

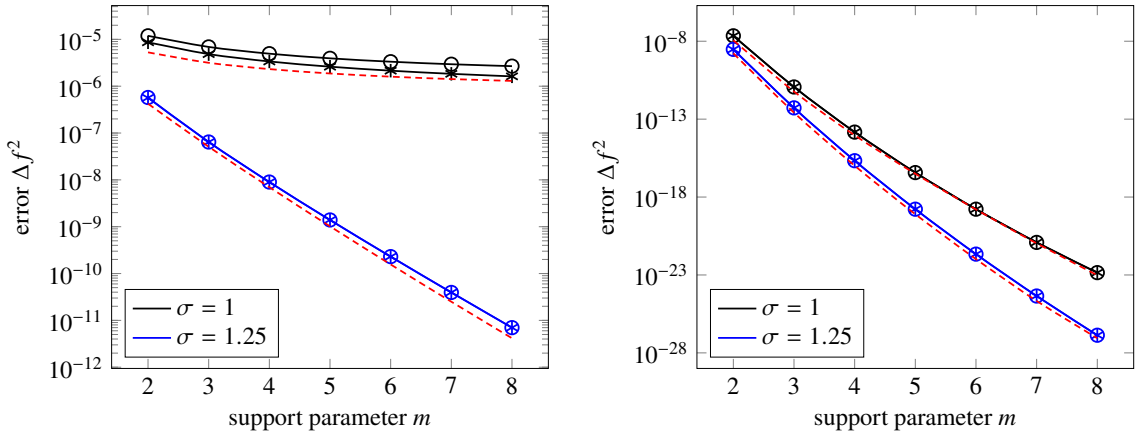


Figure 3.2: Predicted quadratic errors (3.44), marked by o, and (3.45), represented by $*$, for the B-spline window function. We consider the Fourier coefficients (3.46) and (3.47), from left to right. The corresponding measured errors (standard deconvolution) are represented by the red dashed lines.

Modified B-spline

In Figure 3.3 we plot the estimated errors (3.44) and (3.45) for the modified B-spline window, where we applied the support parameter $m = 4$. We considered different shape parameters $b \in {}^1/_2\mathbb{N}$, $b > {}^m/_2$, as well as different oversampling factors σ. The estimates $s(k)$ were computed via (3.39) for $b \neq m$ and via (3.38) in the case $b = m$.

The standard B-spline window, where $b = m$, seems to be the optimal choice in case of fast decreasing coefficients. For the slowly decreasing coefficients (3.47) the choice of a shape parameter $b < m$ can lead to significantly smaller approximation errors, as already mentioned in Example 3.9.

Based on the results we suggest the following simple procedure in order to tune the shape parameter b, cf. [Nes16a].

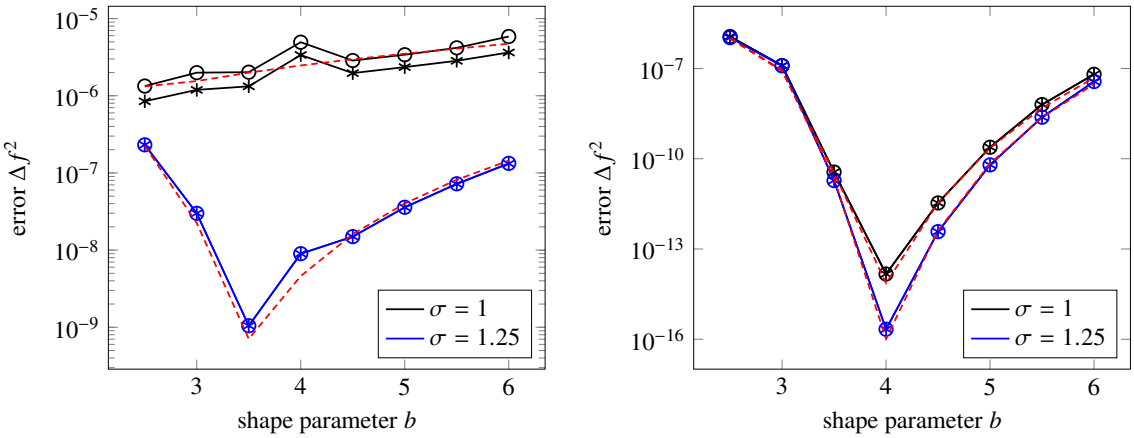


Figure 3.3: Predicted quadratic errors (3.44), marked by o, and (3.45), represented by $*$, for the modified B-spline window function with support parameter $m = 4$. We consider the Fourier coefficients (3.46) and (3.47), from left to right. The corresponding measured errors (standard deconvolution) are represented by the red dashed lines.

Algorithm 3.3 (Shape parameter tuning for B-splines)**.**
Input: Fourier coefficients $\hat{f}_k$, $k \in \mathcal{I}_M$, support parameter $m \in \mathbb{N}$, oversampling factor $\sigma \geq 1$.

i) Set $b_{\text{opt}} := m$ and compute the resulting approximation error via (3.44) or (3.45).

ii) Decrease b_{opt} step by step, i.e., $b_{\text{opt}} \mapsto b_{\text{opt}} - 1/2$, as long as the predicted approximation error decreases.

Output: Optimal shape parameter b_{opt} and approximate error Δf^2. □

Bessel window

The numerical results, presented in the following, give an impression of how the shape parameter b influences the resulting approximation errors in case of the Bessel window function. In Figure 3.4 we plot the estimated errors (3.44) and (3.45) with respect to the shape parameter $b \in \mathbb{R}$, where we applied the support parameter $m = 4$. The estimates $s(k)$ were obtained based on (3.41).

It can be seen that the size of the approximation error is highly influenced by the shape parameter b. Furthermore, it strongly depends on the Fourier coefficients $\hat{f}_k$ which shape parameter will be optimal. Again, the difference between the two deconvolution approaches is only visible for slowly decreasing Fourier coefficients for $\sigma = 1$. The measured errors coincide with the predicted errors.

We remark that commonly the shape parameter

$$b_0 := \frac{(2\sigma - 1)\pi}{\sigma} \approx \begin{cases} 3.1416 & : \sigma = 1, \\ 3.7699 & : \sigma = 1.25, \end{cases} \tag{3.49}$$

is applied, see [PS03, Appendix].

The optimal shape parameters, cf. Figure 3.4, are very close to the suggested standard values (3.49) for the slowly decreasing Fourier coefficients (3.46). In contrast, in case of the exponentially decreasing coefficients (3.47) the optimal shape parameters adopt fundamentally different values.

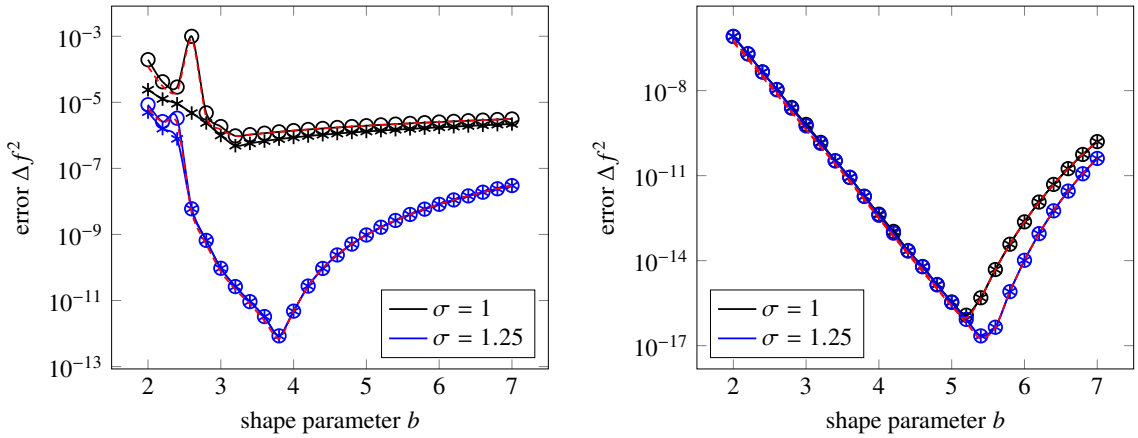


Figure 3.4: Predicted quadratic errors (3.44), marked by o, and (3.45), represented by $*$, for the Bessel window function with support parameter $m = 4$. We consider the Fourier coefficients (3.46) and (3.47), from left to right. The corresponding measured errors (standard deconvolution) are represented by the red dashed lines.

We suppose that in many cases the error depends on the shape parameter b as depicted in Figure 3.4 for the coefficients (3.47). That is, that the error increases with growing distance $|b - b_{\text{opt}}|$ from the uniquely determined optimal shape parameter b_{opt}. Based on this assumption, we suggest the following simple bisection procedure in order to tune the shape parameter b, cf. [Nes16a]. The same method is going to be applied in case of the Gaussian window function. We remark that the variability of the shape parameter for the Bessel window has also been studied in [FS03], where the underlying Fourier coefficients could not be taken into account.

Algorithm 3.4 (Shape parameter tuning for Bessel and Gauss)**.**
Input: Fourier coefficients $\hat{f}_k$, $k \in \mathcal{I}_M$, support parameter $m \in \mathbb{N}$, oversampling factor $\sigma \geq 1$.

i) Set $b_{\text{opt}} := b_0$, cf. (3.49) or (3.50).

ii) Start with step size $d := {}^1\!/_2\, b_{\text{opt}}$.

iii) Set $b_{\text{left}} := b_{\text{opt}} - d$ and $b_{\text{right}} := b_{\text{opt}} + d$. Compute the three resulting approximation errors via (3.44) or (3.45). Determine the new value for b_{opt}.

iv) If the old and the new b_{opt} coincide, choose a smaller step size via $d \mapsto {}^d\!/_2$.

v) Repeat iii) and iv) until the optimal error does not change significantly anymore.

Output: Optimal shape parameter b_{opt} and approximate error Δf^2. □

We applied Algorithm 3.4 in order to determine the optimal shape parameter for the Bessel window function. For this purpose, we considered different values for the support parameter m and the oversampling factor σ. The obtained optimal shape parameters are listed in Table 3.1. In the case $m = 4$ we determine the optimal shape parameters, which are also nearly identifiable based on the results plotted in Figure 3.4. The optimal shape parameters are in general closer to the suggested standard value b_0, see (3.49), for the slowly decreasing coefficients (3.46). Substantially different optimal shape parameters have been obtained in case of the fast decreasing coefficients (3.47).

	Coefficients (3.46)		Coefficients (3.47)	
	$\sigma = 1$	$\sigma = 1.25$	$\sigma = 1$	$\sigma = 1.25$
$m = 2$	4.0743	5.0364	5.5101	5.5776
$m = 3$	3.1416	4.0350	5.3751	5.6015
$m = 4$	3.1539	3.8067	5.2094	5.4597
$m = 5$	3.1907	3.7294	5.0437	5.3622
$m = 6$	3.2398	3.7340	4.8781	5.2462
$m = 7$	3.2398	3.6705	4.7063	5.1358
$m = 8$	3.3379	3.6862	4.5406	5.0216

Table 3.1: Tuned optimal shape parameters for the Bessel window.

Gaussian

We proceeded analogously to the Bessel window and examined the influence of the shape parameter b on the resulting approximation errors. In Figure 3.5 we plot the estimated errors (3.44) and (3.45) for the truncated Gaussian, where we applied the support parameter $m = 4$. The estimates $s(k)$ were obtained based on (3.42). In order to evaluate the complex valued error function we made use of the Faddeeva package [JCW] in MATLAB.

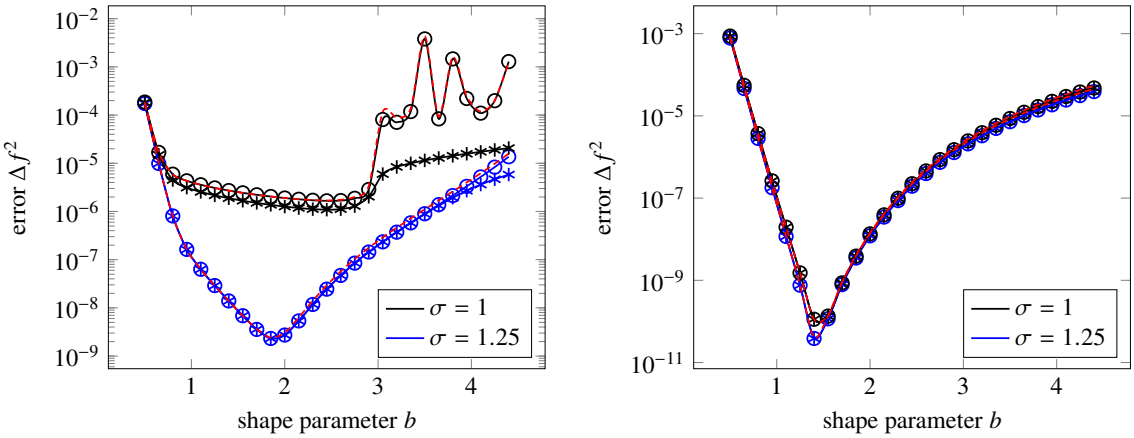


Figure 3.5: Predicted quadratic errors (3.44), marked by o, and (3.45), represented by $*$, for the Gaussian window function with support parameter $m = 4$. We consider the Fourier coefficients (3.46) and (3.47), from left to right. The corresponding measured errors (standard deconvolution) are represented by the red dashed lines.

Again, the results obtained with the optimized deconvolution approach are negligibly better. If the Fourier coefficients tend to zero rapidly or in case of oversampling, the optimal errors concerning the two deconvolution approaches coincide. For large values of the shape parameter b some Fourier coefficients $c_k(\tilde{\varphi}_t)$ are close to a zero point and thus the error plot may yield some unexpected jumps in case of the standard deconvolution approach, see left hand side of Figure 3.5.

We remark that commonly the shape parameter is chosen via

$$b_0 := \frac{2\sigma}{2\sigma - 1} \frac{m}{\pi}, \tag{3.50}$$

see [Ste98, DS99, GL04]. For the parameters $m = 4$ and $\sigma \in \{1, 1.25\}$, as considered in this example, the standard shape parameters adopt the following values.

$$b_0 \approx \begin{cases} 2.5465 & : m = 4, \sigma = 1\,, \\ 2.1221 & : m = 4, \sigma = 1.25\,. \end{cases}$$

As depicted in Figure 3.5, the optimal shape parameters are close to the suggested standard values b_0 only in case of slowly decreasing Fourier coefficients. For the rapidly decreasing coefficients the optimal shape parameters adopt significantly different values, as already observed for the Bessel window function.

Commonly, the Fourier coefficients of the non truncated Gaussian are used in NFFT methods, i.e., $\hat{d}_k := c_k^{-1}(\varphi)$, see Remark 3.13. We considered also this deconvolution approach and estimated the resulting approximation errors via

$$\begin{aligned} \Delta f^2 &= \sum_{k=-M/2}^{M/2-1} |\hat{f}_k|^2 \left(\left(1 - \frac{c_k(\tilde{\varphi}_\mathrm{t})}{c_k(\tilde{\varphi})} \right)^2 + \sum_{r \in \mathbb{Z}\setminus\{0\}} \frac{c_{k+r\sigma M}^2(\tilde{\varphi}_\mathrm{t})}{c_k^2(\tilde{\varphi})} \right) \\ &\approx \sum_{k=-M/2}^{M/2-1} |\hat{f}_k|^2 \left(\left(1 - \frac{c_k(\tilde{\varphi}_\mathrm{t})}{c_k(\tilde{\varphi})} \right)^2 + \sum_{0<|r|\le R_k} \frac{c_{k+r\sigma M}^2(\tilde{\varphi}_\mathrm{t})}{c_k^2(\tilde{\varphi})} \right). \end{aligned} \tag{3.51}$$

The results are given in Figure 3.6.

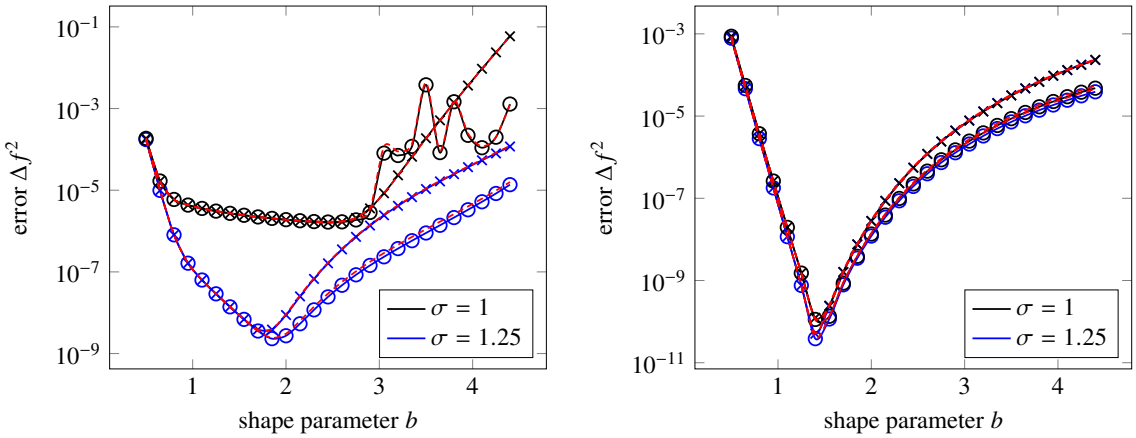


Figure 3.6: Predicted quadratic errors (3.44), marked by o, and (3.51), represented by x, for the Gaussian window function with support parameter $m = 4$. We consider the Fourier coefficients (3.46) and (3.47), from left to right. The corresponding measured errors are represented by the red dashed lines.

The optimal errors based on the deconvolution $\hat{d}_k := c_k^{-1}(\tilde{\varphi}_\mathrm{t})$ are marginally better than those obtained by setting $\hat{d}_k := c_k^{-1}(\tilde{\varphi})$. Again, for rapidly decreasing Fourier coefficients the difference is not visible at all.

If the error behaves as depicted for the rapidly decreasing Fourier coefficients, the optimal shape parameter b_opt can be determined via Algorithm 3.4. We applied Algorithm 3.4 in order to tune b_opt for the Gaussian window function, where we considered different values for the support parameter m and the oversampling factor σ. The obtained optimal shape parameters are given in Table 3.2.

Again, we obtain agreement with the results plotted in Figure 3.5 for $m = 4$. In general, the tuned optimal shape parameters are closer to the suggested standard values for slowly decreasing coefficients. Fundamentally different results are obtained if the Fourier coefficients tend to zero rapidly.

	Coefficients (3.46)		Coefficients (3.47)		Standard values	
	$\sigma = 1$	$\sigma = 1.25$	$\sigma = 1$	$\sigma = 1.25$	$\sigma = 1$	$\sigma = 1.25$
$m = 2$	0.7759	0.7626	0.7709	0.7647	1.2732	1.0610
$m = 3$	1.6114	1.2434	1.1116	1.1004	1.9099	1.5915
$m = 4$	2.4669	1.8900	1.4672	1.4382	2.5465	2.1221
$m = 5$	3.3323	2.5697	1.8278	1.7822	3.1831	2.6526
$m = 6$	3.7600	3.2453	2.1934	2.1324	3.8197	3.1831
$m = 7$	4.5956	3.8297	2.5763	2.4878	4.4563	3.7136
$m = 8$	4.9338	4.4762	2.9692	2.8474	5.0930	4.2441

Table 3.2: Tuned optimal shape parameters as well as suggested standard values b_0, see (3.50), for the Gaussian window.

Comparison of the different window functions and summary

As a short summary we consider all introduced window functions, namely the (modified) B-spline, the Bessel as well as the Gaussian window function, and compare the resulting approximation errors. In Figure 3.7 we plot the predicted approximation errors for all window functions, where we applied different combinations of the parameters m and σ. The shape parameters were tuned automatically via Algorithm 3.3 or Algorithm 3.4. Obviously, the Bessel window yields the smallest error in most cases. For the fast decreasing coefficients (3.47), the B-spline window performs better if a small support parameter m is applied.

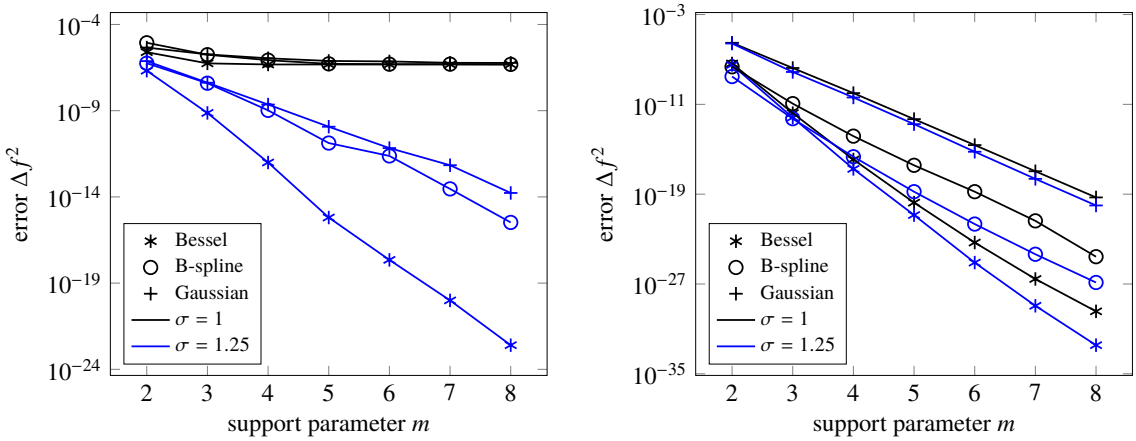


Figure 3.7: Comparison of the predicted approximation errors (3.44) for the different window functions. We consider the Fourier coefficients (3.46) and (3.47), from left to right, and plot the error with respect to the support parameter m as well as for different oversampling factors σ. The optimal shape parameters have been determined via Algorithm 3.3 or Algorithm 3.4, respectively.

3.3.2 The multivariate case

In order to discuss later the approximation errors regarding the 3d-periodic fast Ewald summation, see Section 4.3, we now pay attention to the three dimensional case. Analogously to the univariate setting, we assume the Fourier coefficients $\hat{f}_{\boldsymbol{k}}$ to be known. Of course, this is not the case in all possible applications, but in case of the 3d-periodic Coulomb problem the Fourier coefficients are well known, see Section 2.2.1 and especially (2.31).

We aim to predict the resulting approximation errors accurately and, in addition, as efficiently as possible. We will see that a tensor product structure of the Fourier coefficients $\hat{f}_{\boldsymbol{k}}$ is particularly advantageous.

Since the Fourier coefficients of the NFFT window function are of a tensor product structure, cf. (3.34), we can separate the estimation of the error sums regarding the single dimensions as follows. In case of the standard deconvolution (3.14) we obtain

$$\begin{aligned}
\|f - f_{\approx}\|_{\mathcal{L}_2}^2 &= \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \sum_{\boldsymbol{r}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}} \frac{c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}^2(\tilde{\varphi}_{\mathrm{t}})}{c_{\boldsymbol{k}}^2(\tilde{\varphi}_{\mathrm{t}})} \\
&\leq \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \left(\prod_{j=1}^{3} (1 + s_j(k_j)) - 1 \right) \qquad (3.52) \\
&= \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \left(\sum_{j=1}^{3} s_j(k_j) + \sum_{j_1<j_2} s_{j_1}(k_{j_1}) s_{j_2}(k_{j_2}) + s_1(k_1) s_2(k_2) s_3(k_3) \right), \qquad (3.53)
\end{aligned}$$

where we use the upper bounds

$$\sum_{r\in\mathbb{Z}\setminus\{0\}} \frac{c_{k+r\sigma_j M_j^2}(\tilde{\varphi}_{\mathrm{t},j})}{c_k^2(\tilde{\varphi}_{\mathrm{t},j})} \leq s_j(k),$$

depending on the given window function.

The computation of the resulting error is now possible very efficiently with only $O(M_1 + M_2 + M_3)$ arithmetic operations, if also the squared Fourier coefficients $|\hat{f}_{\boldsymbol{k}}|^2$ are of a tensor product structure, see Example 3.16. If the three dimensions are not separable, the numerical complexity for predicting the $\mathcal{L}_2$-error via evaluating (3.53) is $O(M_1 M_2 M_3)$, which roughly corresponds to the complexity of the NFFT itself.

In the following we consider only the B-spline window function, for which the upper bounds $s_j(k_j)$ are given in (3.38). The presented estimates have been verified based on a simple implementation of the trivariate NFFT in MATLAB. In order to measure the actual approximation error, we considered $N = 10^5$ randomly chosen nodes $\boldsymbol{x}_j$, uniquely distributed in $\mathbb{T}^3$, and evaluated (3.48).

Example 3.16. We consider the coefficients

$$\hat{f}_{\boldsymbol{k}} := \mathrm{e}^{-\alpha^2\|\boldsymbol{k}\|^2}, \quad \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}, \qquad (3.54)$$

where $\alpha \in \mathbb{R}$. The fast evaluation of the resulting approximation error (3.53) is straight forward since the three dimensions can be separated via

$$\hat{f}_{\boldsymbol{k}}^2 = \mathrm{e}^{-2\alpha^2\|\boldsymbol{k}\|^2} = \mathrm{e}^{-2\alpha^2 k_1^2} \cdot \mathrm{e}^{-2\alpha^2 k_2^2} \cdot \mathrm{e}^{-2\alpha^2 k_3^2}.$$

If $\boldsymbol{M} = (M, M, M)^\top$ and the same oversampling factor σ is applied in all three dimensions, we only have to compute the two sums

$$S := \sum_{k=-M/2}^{M/2-1} \mathrm{e}^{-2\alpha^2 k^2} \quad \text{and} \quad R := \sum_{k=-M/2}^{M/2-1} \frac{8m}{4m-1} \left(\frac{|k|}{|k| - \sigma M} \right)^{4m} \mathrm{e}^{-2\alpha^2 k^2}, \tag{3.55}$$

in case of the B-spline window, and the error is estimated via

$$\|f - f_{\approx}\|_{\mathcal{L}_2}^2 \leq 3S^2 R + 3SR^2 + R^3. \tag{3.56}$$

In the non cubic case with $\boldsymbol{M} = (M_1, M_2, M_3)$, the computation of six sums S_j, R_j, where $j \in \{1, 2, 3\}$, is required. The estimate has been verified based on a numerical example in MATLAB, where we set $\alpha := 0.2$ and $\boldsymbol{M} := (64, 64, 64)$ in (3.54). A comparison of the measured errors with the computed estimates (3.56) for different support parameters m and oversampling factors σ is given in Figure 3.8. The actual errors can be estimated from above fairly well, where we notice a deviation factor less than 2.5, approximately. □

Example 3.17. We consider the coefficients

$$\hat{f}_{\boldsymbol{k}} := \frac{1}{1 + \|\boldsymbol{k}\|^2} = \frac{1}{1 + k_1^2 + k_2^2 + k_3^2}, \quad \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}, \tag{3.57}$$

which are not of a tensor product structure. Thus, the direct computation of the approximation error (3.53) would require $O(M_1 M_2 M_3)$ arithmetic operations.

In the following we consider the case $\boldsymbol{M} = (M, M, M)^\top$ with $M = 64$ and proceed as follows in order to estimate the errors more efficiently. Note that we have $1 \leq 1 + \|\boldsymbol{k}\|^2 \leq 1 + 3 \cdot 32^2 = 3073$. Suppose that we have given an approximation of the function $f(x) = x^{-2}$ by a short sum of exponential terms, i.e.,

$$\frac{1}{x^2} \approx \sum_{i=1}^{n} c_i \, \mathrm{e}^{-\omega_i x}, \tag{3.58}$$

which holds for every $x \in [1, 3073]$. Based on this we are able to approximate the squared Fourier coefficients via

$$|\hat{f}_{\boldsymbol{k}}|^2 = \frac{1}{\left(1 + \|\boldsymbol{k}\|^2\right)^2} \approx \sum_{i=1}^{n} c_i \, \mathrm{e}^{-\omega_i} \, \mathrm{e}^{-\omega_i (k_1^2 + k_2^2 + k_3^2)}$$

for all $\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}$, $\boldsymbol{M} = (64, 64, 64)$. Then, we only have to compute the $2n$ sums

$$S_i := \sum_{k=-M/2}^{M/2-1} \mathrm{e}^{-\omega_i k^2} \quad \text{and} \quad R_i := \sum_{k=-M/2}^{M/2-1} \frac{8m}{4m-1} \left(\frac{|k|}{|k| - \sigma M} \right)^{4m} \mathrm{e}^{-\omega_i k^2},$$

where $i \in \{1, \ldots, n\}$, and the error can be approximated by

$$\|f - f_{\approx}\|_{\mathcal{L}_2}^2 \approx \sum_{i=1}^{n} c_i \, \mathrm{e}^{-\omega_i} \left(3S_i^2 R_i + 3S_i R_i^2 + R_i^3 \right), \tag{3.59}$$

which only requires $O(nM)$ instead of $O(M^3)$ arithmetic operations. The computation of an approximation of the form (3.58) is possible based on the ESPRIT method [RK89], a Prony like method [PT13] or the Remez algorithm, as presented in [Hac05], for instance.

Note that [Hac05] serves a variety of such approximations for the function $f(x) = x^{-1}$, amongst others. We consider an approximation of the form

$$\frac{1}{x} \approx \sum_{i=1}^{n} \tilde{c}_i \, \mathrm{e}^{-\omega_i x}, \quad x \in [1, 4 \cdot 10^3],$$

where at least $n = 7$ exponential terms are required in order to keep the maximum relative approximation error below 1. Hereby, we achieve a maximum absolute approximation error of $\epsilon_{\text{abs}} \approx 1.1 \cdot 10^{-4}$ and a maximum relative error $\epsilon_{\text{rel}} \approx 4.4 \cdot 10^{-1}$. By differentiation with respect to x we obtain

$$\frac{1}{x^2} \approx \sum_{i=1}^{n} -\tilde{c}_i \omega_i \, \mathrm{e}^{-\omega_i x} =: \sum_{i=1}^{n} c_i \, \mathrm{e}^{-\omega_i x}, \quad x \in [1, 4 \cdot 10^4],$$

which yields slightly larger approximation errors ($\epsilon_{\text{abs}} \approx 5.5 \cdot 10^{-3}$, $\epsilon_{\text{rel}} \approx 6.9 \cdot 10^{-1}$). We applied this approximation with only $n = 7$ exponential terms to approximate (3.53) via (3.59) and compared to the values obtained from the direct evaluation, see Table 3.3.

	$\sigma = 1$		$\sigma = 1.25$	
	approx	direct	approx	direct
$m = 2$	6.701e-02	6.128e-02	2.858e-03	2.711e-03
$m = 3$	4.060e-02	3.633e-02	3.490e-04	3.207e-04
$m = 4$	2.975e-02	2.629e-02	5.076e-05	4.582e-05
$m = 5$	2.398e-02	2.103e-02	8.061e-06	7.193e-06
$m = 6$	2.049e-02	1.787e-02	1.351e-06	1.196e-06

Table 3.3: Approximated errors, see (3.59), and exact errors, see (3.53), for the Fourier coefficients $\hat{f}_{\boldsymbol{k}} = (1 + \|\boldsymbol{k}\|^2)^{-1}$, $\boldsymbol{k} \in \mathcal{I}_{(64,64,64)}$.

A comparison of the measured errors with the computed estimates for different support parameters m and oversampling factors σ is given in Figure 3.8. The actual errors are estimated from above very well. In contrast to the fast decreasing coefficients, as considered in Example 3.16, we recognize somewhat larger deviation factors up to 3.35, approximately. □

Remark 3.18 (Relative errors)**.** As already pointed out regarding the one-dimensional setting, see Remark 3.15, we may estimate also the resulting relative errors by computing

$$\|f\|_{\mathcal{L}_2}^2 = \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{f}_{\boldsymbol{k}}^2.$$

For the considered Fourier coefficients we obtain

$$\|f\|_{\mathcal{L}_2}^2 \approx 2.46 \cdot 10^2 \iff \Delta f_{\text{rel}}^2 \approx 4.06 \cdot 10^{-3} \|f - f_{\approx}\|_{\mathcal{L}_2}^2 \quad \text{for } \hat{f}_{\boldsymbol{k}} \text{ defined via (3.54)},$$

and

$$\|f\|_{\mathcal{L}_2}^2 \approx 9.67 \iff \Delta f_{\text{rel}}^2 \approx 1.03 \cdot 10^{-1} \|f - f_{\approx}\|_{\mathcal{L}_2}^2 \quad \text{for } \hat{f}_{\boldsymbol{k}} \text{ defined via (3.57)}.$$

□

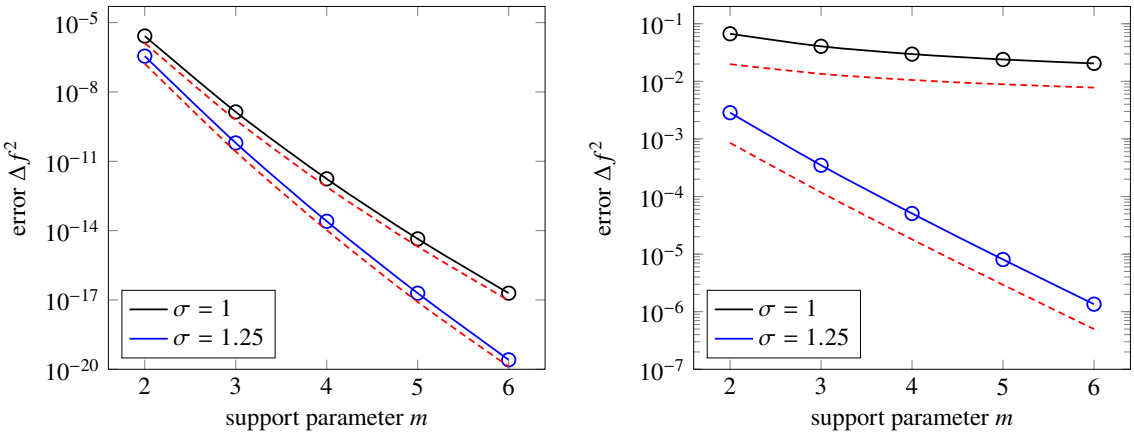


Figure 3.8: Predicted quadratic errors for the trivariate NFFT, see (3.56) and (3.59), where we consider the Fourier coefficients (3.54) with $\alpha := 0.2$ (left) and (3.57) (right), respectively. We plot the errors with respect to the support parameter m, where we applied the B-spline window with different oversampling factors σ (see labels). The according measured errors (3.48) are represented by the red dashed lines.

In the following examples we consider the gradient NFFT. Again, we measured the resulting errors in MATLAB based on a simple implementation of the gradient NFFT. We considered $N = 10^5$ uniquely distributed nodes and approximated the actual errors via

$$\Delta f_{\text{grad}}^2 := \|\nabla f - \nabla f_{\approx}\|_2^2 \approx \frac{1}{N} \sum_{j=1}^{N} \|\nabla f(\boldsymbol{x}_j) - \nabla f_{\approx}(\boldsymbol{x}_j)\|^2.$$

Example 3.19. We estimated the errors of the gradient NFFT for the Fourier coefficients (3.54). For this purpose, we applied the $i\boldsymbol{k}$- as well as the analytic differentiation approach with the B-spline window function.

In case of differentiation in Fourier space we obtain

$$\|\nabla f - \nabla f_{\text{ik}}\|^2 \leq 4\pi^2 \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 (k_1^2 + k_2^2 + k_3^2) \left(\prod_{j=1}^{3} (1 + s_j(k_j)) - 1 \right),$$

cf. (3.52). Analogously to (3.56), we obtain in the cubic case, i.e., $\boldsymbol{M} = (M, M, M)^{\top}$ and $\boldsymbol{\sigma} = (\sigma, \sigma, \sigma)^{\top}$, the estimate

$$\|\nabla f - \nabla f_{\text{ik}}\|^2 \leq 12\pi^2 \left(2\tilde{S}SR + \tilde{S}R^2 + 2\tilde{R}RS + \tilde{R}S^2 + \tilde{R}R^2 \right), \tag{3.60}$$

where we additionally define

$$\tilde{S} := \sum_{k=-M/2}^{M/2-1} k^2 \, \mathrm{e}^{-2\alpha^2 k^2} \quad \text{and} \quad \tilde{R} := \sum_{k=-M/2}^{M/2-1} \frac{8m}{4m-1} \frac{k^{4m+2}}{(|k| - \sigma M)^{4m}} \, \mathrm{e}^{-2\alpha^2 k^2},$$

and S, R are given in (3.55).

A fundamentally different analysis is necessary in case of analytic differentiation, where we have

$$\|\nabla f - \nabla f_{\mathrm{ad}}\|^2 = \sum_{\boldsymbol{k}\in I_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \left(\sum_{\boldsymbol{r}\in\mathbb{Z}^3} \|\boldsymbol{k} + \boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}\|^2 \frac{c^2_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})}{c^2_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})} - \|\boldsymbol{k}\|^2 \right).$$

In order to estimate the error we have to derive upper bounds of sums of the form

$$\sum_{r\in\mathbb{Z}} (k + r\sigma M)^2 \frac{c^2_{k+r\sigma M}(\tilde{\varphi})}{c^2_k(\tilde{\varphi})} = k^2 + \sum_{r\in\mathbb{Z}\setminus\{0\}} (k + r\sigma M)^2 \frac{c^2_{k+r\sigma M}(\tilde{\varphi})}{c^2_k(\tilde{\varphi})}.$$

For the B-spline window we obtain

$$\begin{aligned}\sum_{r\in\mathbb{Z}\setminus\{0\}} (k + r\sigma M)^2 \frac{c^2_{k+r\sigma M}(\tilde{\varphi})}{c^2_k(\tilde{\varphi})} &= k^2 \sum_{r\in\mathbb{Z}\setminus\{0\}} \left(\frac{|k|}{|k| + r\sigma M} \right)^{4(m-1/2)} \\ &\leq \frac{8m-4}{4m-3} \frac{k^{4m}}{(|k| - \sigma M)^{4m-2}} =: s_{\mathrm{ad}}(k),\end{aligned}$$

analogously to (3.38).

By making use of the upper bounds $s(k)$ and $s_{\mathrm{ad}}(k)$ we can write

$$\|\nabla f - \nabla f_{\mathrm{ad}}\|^2 \leq \sum_{\boldsymbol{k}\in I_{\boldsymbol{M}}} |\hat{f}_{\boldsymbol{k}}|^2 \left(\begin{array}{rl} & (k_1^2 + s_{\mathrm{ad}}(k_1))(1 + s(k_2))(1 + s(k_3)) \\ + & (1 + s(k_1))(k_2^2 + s_{\mathrm{ad}}(k_2))(1 + s(k_3)) \\ + & (1 + s(k_1))(1 + s(k_2))(k_3^2 + s_{\mathrm{ad}}(k_3)) - \|\boldsymbol{k}\|^2 \end{array} \right),$$

where we also exploit the tensor product structure of the window function $\tilde{\varphi}$.

Finally, we obtain in the cubic case

$$\|\nabla f - \nabla f_{\mathrm{ad}}\|^2 \leq 12\pi^2 \left(TR^2 + \tilde{S}R^2 + 2TSR + 2\tilde{S}SR + TS^2 \right), \tag{3.61}$$

where we additionally define

$$T := \sum_{k=-M/2}^{M/2-1} \frac{(8m-4)k^{4m}}{(4m-3)(|k| - \sigma M)^{4m-2}} \, \mathrm{e}^{-2\alpha^2 k^2}.$$

A comparison between the measured errors and the presented estimates is given in Figure 3.9. It can be seen that we achieve a very good agreement. □

Example 3.20. We estimated the errors of the gradient NFFT for the slowly decreasing Fourier coefficients 3.57. By making use of the techniques applied in Examples 3.17 and 3.19, we obtain

$$\|\nabla f - \nabla f_{\mathrm{ik}}\|^2 \leq 12\pi^2 \sum_{i=1}^{N} c_i \, \mathrm{e}^{-\omega_i} \left(2\tilde{S}_i S_i R_i + \tilde{S}_i R_i^2 + 2\tilde{R}_i R_i S_i + \tilde{R}_i S_i^2 + \tilde{R}_i R_i^2 \right),$$

$$\|\nabla f - \nabla f_{\mathrm{ad}}\|^2 \leq 12\pi^2 \sum_{i=1}^{N} c_i \, \mathrm{e}^{-\omega_i} \left(T_i R_i^2 + \tilde{S}_i R_i^2 + 2T_i S_i R_i + 2\tilde{S}_i S_i R_i + T_i S_i^2 \right),$$

cf. (3.60) and (3.61). Here, we additionally define the sums

$$\tilde{S}_i := \sum_{k=-M/2}^{M/2-1} k^2 \, \mathrm{e}^{-\omega_i k^2} \quad \text{and} \quad \tilde{R}_i := \sum_{k=-M/2}^{M/2-1} \frac{8m}{4m-1} \frac{k^{4m+2}}{(|k| - \sigma M)^{4m}} \, \mathrm{e}^{-\omega_i k^2},$$

as well as

$$T_i := \sum_{k=-M/2}^{M/2-1} \frac{(8m-4)k^{4m}}{(4m-3)(|k| - \sigma M)^{4m-2}} \, \mathrm{e}^{-\omega_i k^2}.$$

Note that S_i and R_i have already been defined in Example 3.17.

A comparison between the measured errors and the presented estimates is given in Figure 3.9. Again, in contrast to the fast decreasing coefficients we obtain somewhat larger deviations. □

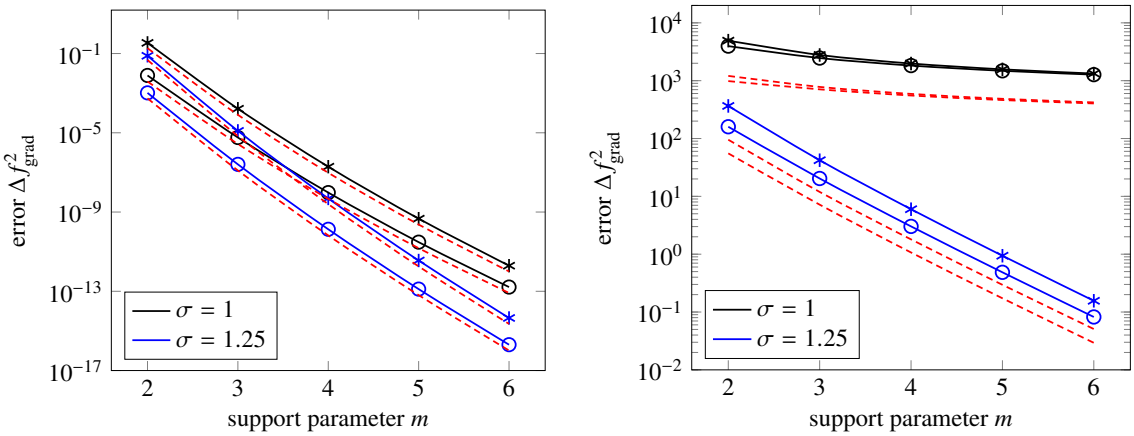


Figure 3.9: Predicted quadratic errors for the gradient NFFT in case of i$\boldsymbol{k}$-differentiation (o) as well as for the analytic differentitation approach ($*$). We consider the Fourier coefficients (3.54) with $\alpha := 0.2$ (left) and (3.57) (right), respectively, and plot the errors with respect to the support parameter m. We applied the B-spline window with different oversampling factors σ (see legend). The according measured errors are represented by the red dashed lines.

Remark 3.21 (Relative errors)**.** In order to estimate the relative error

$$\Delta f_{\text{grad,rel}}^2 := \frac{\|\nabla f - \nabla f_{\approx}\|_{\mathcal{L}_2}^2}{\|\nabla f\|_{\mathcal{L}_2}^2}$$

we have to compute

$$\|\nabla f\|_{\mathcal{L}_2}^2 = 4\pi^2 \sum_{\boldsymbol{k} \in I_{\boldsymbol{M}}} \|\boldsymbol{k}\|^2 \hat{f}_{\boldsymbol{k}}^2.$$

For the above considered sets of Fourier coefficients we obtain

$$\|\nabla f\|_{\mathcal{L}_2}^2 \approx 1.82 \cdot 10^5 \iff \Delta f_{\text{grad,rel}}^2 \approx 5.49 \cdot 10^{-6} \|\nabla f - \nabla f_{\approx}\|_{\mathcal{L}_2}^2, \quad \hat{f}_{\boldsymbol{k}} \text{ defined via (3.54)},$$

and

$$\|\nabla f\|_{\mathcal{L}_2}^2 \approx 1.82 \cdot 10^4 \iff \Delta f_{\text{grad,rel}}^2 \approx 5.48 \cdot 10^{-5} \|\nabla f - \nabla f_{\approx}\|_{\mathcal{L}_2}^2, \quad \hat{f}_{\boldsymbol{k}} \text{ defined via (3.57)}.$$

□

3.4 NFFT based fast summation

A smooth periodic function is well approximated by a trigonometric polynomial. Given such an approximation, the fast evaluation of the function at nonequispaced nodes is possible based on the NFFT, as introduced in Section 3.2. However, in several applications, as for example the computation of electrostatic interactions under open or partially periodic boundary conditions, see Section 2.1, one has to deal with nonperiodic functions. In order to evaluate nonperiodic functions at nonequispaced nodes efficiently, one may embed the given function into a periodic framework or rather periodize it. Afterward, the constructed periodic function may be treated via the NFFT.

This roughly explains the basic idea of the NFFT based fast summation approach, which is discussed in the following. We start by considering symmetric univariate functions in Section 3.4.1, where we discuss different approaches to periodize a given function as well as resulting advantages and disadvantages. Radial functions in higher dimensions are studied in Section 3.4.2. The introduced techniques are of great importance for the algorithms presented in Section 4.1.2.

3.4.1 One-dimensional case

We consider the following problem. Let $K : \mathbb{R} \to \mathbb{R}$ be a continuous, smooth and symmetric function, i.e., $K(x) = K(|x|)$ for all $x \in \mathbb{R}$. For a given set of nodes $x_j \in [-L/2, L/2]$ and coefficients $\alpha_j \in \mathbb{R}$, $j = 1, \dots, N$, we are interested in a fast evaluation of the sums

$$f_j := \sum_{i=1}^{N} \alpha_i K(x_i - x_j), \quad j = 1, \dots, N.$$

If the kernel function K does not allow the separation of the source and the target nodes, i.e., K is not of the form $K(x - y) = K_1(x)K_2(y)$, a direct evaluation of the sums f_j may be applied. This requires undesirable $O(N^2)$ arithmetic operations.

Note that we have $x_i - x_j \in [-L, L]$. In order to achieve an arithmetic complexity of $O(N \log N)$ we might compute an approximation of K in terms of a trigonometric polynomial

$$K(x) \approx \sum_{l=-M/2}^{M/2-1} \hat{b}_l \, \mathrm{e}^{2\pi \mathrm{i} l x / h} \tag{3.62}$$

with $M \in 2\mathbb{N}$ and a period $h \geq 2L$, since $x \in [-L, L]$ in our setting.

Based on (3.62) we can approximate the sums f_j via

$$f_j \approx \sum_{i=1}^{N} \alpha_i \sum_{l=-M/2}^{M/2-1} \hat{b}_l \, \mathrm{e}^{2\pi \mathrm{i} l (x_i - x_j)/h} = \sum_{l=-M/2}^{M/2-1} \hat{b}_l \left(\sum_{i=1}^{N} \alpha_i \, \mathrm{e}^{2\pi \mathrm{i} l x_i / h} \right) \mathrm{e}^{-2\pi \mathrm{i} l x_j / h},$$

where the inner sums can be computed efficiently via the one-dimensional adjoint NFFT. Afterward, the outer sums are approximated via the one-dimensional NFFT.

In the following we discuss three different ways of how to compute an approximation of the form (3.62). We remark that the presented approaches are actually not limited to symmetric functions.

Variant I - Periodization

This method is based on the Poisson summation formula, see Lemma 2.2. If the Fourier transform

$$\hat{K}(\xi) := \int_{-\infty}^{\infty} K(x) \, \mathrm{e}^{-2\pi \mathrm{i} \xi x} \, \mathrm{d}x, \quad \xi \in \mathbb{R},$$

is known analytically and $K(x)$ is sufficiently small for all $|x| > h$, we have

$$K(x) \approx K_h(x) := \sum_{n\in\mathbb{Z}} K(x+hn) = \frac{1}{h}\sum_{l\in\mathbb{Z}} \hat{K}\left(\tfrac{l}{h}\right) \mathrm{e}^{2\pi \mathrm{i} l x/h} \approx \frac{1}{h}\sum_{l=-M/2}^{M/2-1} \hat{K}\left(\tfrac{l}{h}\right) \mathrm{e}^{2\pi \mathrm{i} l x/h} \tag{3.63}$$

for some $M \in 2\mathbb{N}$ large enough. In other words, we set $\hat{b}_l := h^{-1}\hat{K}(l/h)$ in (3.62).

The presented approach is obviously limited to functions that decay fast enough in the interval $[-h/2, h/2]$. Whenever $K(\pm L)$ is relatively large, we have to choose $h \gg 2L$ and a large cutoff parameter M will be required in order to achieve a certain accuracy. See Figure 3.10 for a graphical illustration.

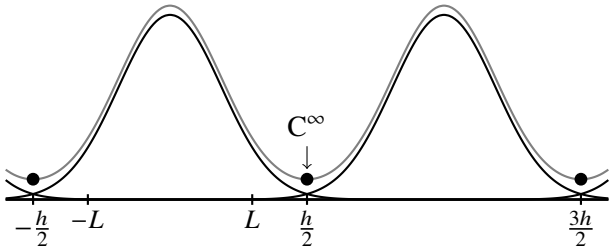


Figure 3.10: The periodic function $K_h(x) \in C^\infty(h\mathbb{T})$. We only have $K(x) \approx K_h(x)$ if $K(h/2)$ is already very small.

Variant II - Truncation

We simply truncate the kernel K at $x = \pm h/2$ and consider the periodic version

$$\tilde{K}(x) := K(x+hn) \quad \text{with } n \in \mathbb{Z} : x + hn \in [-h/2, h/2).$$

The resulting function is periodic with period h and continuous, but not smooth, i.e., $K \in C^0(h\mathbb{T})$, see Figure 3.11. This results in a rather slow second order convergence in Fourier space, i.e., we expect that a relatively large cutoff parameter M will be needed in (3.62) to achieve a good approximation. The Fourier coefficients of the periodic function $\tilde{K}$ are given by

$$c_l(\tilde{K}) = \frac{1}{h}\int_{-h/2}^{h/2} K(x)\,\mathrm{e}^{-2\pi \mathrm{i} l x/h}\,\mathrm{d}x =: \hat{b}_l.$$

If the Fourier coefficients are not known analytically, we may sample the kernel function K at M equally spaced nodes and approximate the coefficients via the FFT.

Variant III - Regularization

We choose $h > 2L$ and construct a periodic function, which is smooth up to some higher order $p \in \mathbb{N}$, cf. [PS03]. The resulting function is named regularization or regularized kernel function and is of the form

$$K_\mathrm{R}(x) := \begin{cases} K(x) & : x \in [-L, L], \\ K_\mathrm{B}(x) & : x \in (L, h-L), \end{cases}$$

i.e., $K_\mathrm{R} : [-L, h-L) \to \mathbb{R}$.

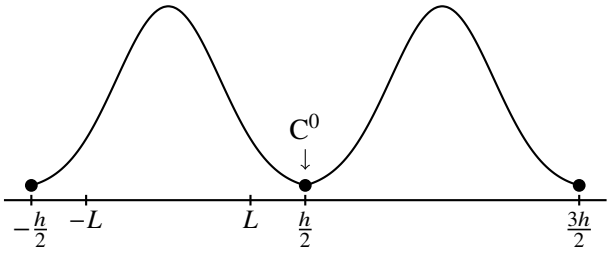


Figure 3.11: If we simply truncate the kernel K at $x = \pm h/2$, the periodic function $\tilde{K}$ is not smooth and the Fourier coefficients tend to zero only with order 2.

In order to construct the function K_B, we claim that the $2p$ interpolation conditions

$$\frac{\mathrm{d}}{\mathrm{d}x^k} K_\mathrm{B}(x)\Big|_{x=L} = \frac{\mathrm{d}}{\mathrm{d}x^k} K(x)\Big|_{x=L}, \tag{3.64}$$

$$\frac{\mathrm{d}}{\mathrm{d}x^k} K_\mathrm{B}(x)\Big|_{x=h-L} = \frac{\mathrm{d}}{\mathrm{d}x^k} K(x)\Big|_{x=-L}, \tag{3.65}$$

where $k = 0, \ldots, p-1$, are fulfilled. Thus, the periodic version of the regularized kernel

$$\tilde{K}_\mathrm{R}(x) := K_\mathrm{R}(x + hn) \quad \text{with } n \in \mathbb{Z} : x + hn \in [-L, h - L)$$

is smooth and periodic, more precisely $\tilde{K}_\mathrm{R} \in C^{p-1}(h\mathbb{T})$. For all $x \in [-L, L]$ we get

$$K(x) = \tilde{K}_\mathrm{R}(x) = \sum_{l \in \mathbb{Z}} c_l(\tilde{K}_\mathrm{R})\, \mathrm{e}^{2\pi \mathrm{i} l x/h} \approx \sum_{l=-M/2}^{M/2-1} c_l(\tilde{K}_\mathrm{R})\, \mathrm{e}^{2\pi \mathrm{i} l x/h} =: T_M(K_\mathrm{R})(x)$$

for $M \in 2\mathbb{N}$ large enough. The Fourier coefficients $c_l(\tilde{K}_\mathrm{R}) =: \hat{b}_l$, $l = -M/2, \ldots, M/2 - 1$, can be computed via the FFT after sampling the function K_R at M equispaced nodes in the interval $[-h/2, h/2)$, i.e.,

$$\hat{b}_l := \frac{1}{M} \sum_{j=-M/2}^{M/2-1} \tilde{K}_\mathrm{R}\left(\frac{jh}{M}\right) \mathrm{e}^{-2\pi \mathrm{i} jl/h}, \quad l = -M/2, \ldots, M/2 - 1.$$

A graphical illustration of Variant III is given in Figure 3.12.

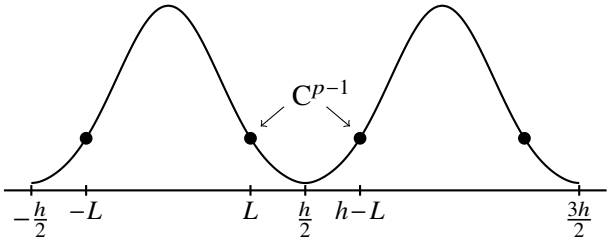


Figure 3.12: The regularized kernel function K_R is obtained by constructing a smooth transition between L and $h - L$. The obtained periodic function $\tilde{K}_\mathrm{R}$ is in the space $C^{p-1}(h\mathbb{T})$ for some $p \in \mathbb{N}$.

In general, the function K_B is a polynomial of degree $2p - 1$ and is computed via the so called two-point Taylor interpolation, as presented in the following lemma.

Lemma 3.22. *Let an interval $[m-r, m+r], r > 0$, and the interpolation values $a_j = f^{(j)}(m-r)$, $b_j = f^{(j)}(m+r)$, $j = 0, \dots, p-1$, be given. For $y := \frac{x-m}{r}$ the polynomial*

$$P(x) = \sum_{j=0}^{p-1} B(p, j, y) r^j a_j + \sum_{j=0}^{p-1} B(p, j, -y)(-r)^j b_j, \tag{3.66}$$

of degree $2p-1$, where the basis polynomials $B(p, j, y)$ are defined via

$$B(p, j, y) := \sum_{k=0}^{p-1-j} \binom{p-1+k}{k} \frac{1}{j! 2^p 2^k} (1-y)^p (1+y)^{k+j},$$

satisfies the interpolation conditions $P^{(j)}(m-r) = a_j$, $P^{(j)}(m+r) = b_j$, $j = 0, \dots, p-1$.

Proof. See Corollary 2.2.6 in [AW93] or Proposition 3.2 in [FS04]. ∎

In our setting we have to set $m := {}^{h}/_{2}$, $r := {}^{h}/_{2} - L$, $a_j := K^{(j)}(L)$ and $b_j := K^{(j)}(-L)$ in Lemma 3.22. Since we assume the kernel function K to be symmetric we additionally have $b_j = (-1)^j a_j$.

Remark 3.23 (Error estimate)**.** For a symmetric kernel function K we have $b_j = (-1)^j a_j$ and, thus, the representation (3.66) of the interpolating polynomial simplifies to

$$P(x) = \sum_{j=0}^{p-1} r^j a_j \big(B(p, j, y) + B(p, j, -y)\big).$$

An upper bound for the p-th derivative of P is given by

$$\max_{x \in [m, m+r]} |P^{(p)}(x)| \le p! \left(\frac{3}{2}\right)^p \sum_{l=0}^{p-2} \binom{p-1+l}{l} \frac{r^{-1-l}}{2^l (p-1-l)!} |a_{p-1-l}|, \tag{3.67}$$

see [FS04, Theorem 3.4]. Furthermore, we have

$$\max_{|x| \le {}^{h}/_{2}} \left|\tilde{K}_{\mathrm{R}}(x) - T_M(K_{\mathrm{R}})(x)\right| \le C \cdot \frac{h^{p-1}}{(p-1)\pi^p M^{p-1}} \int_0^{{}^{h}/_{2}} \left|K_{\mathrm{R}}^{(p)}(x)\right| \mathrm{d}x \tag{3.68}$$

with a constant $C \approx 4$, cf. [FS04, Lemma 4.1]. With (3.67) we get

$$\begin{aligned} \int_0^{{}^{h}/_{2}} \left|K_{\mathrm{R}}^{(p)}(x)\right| \mathrm{d}x &= \int_0^L \left|K^{(p)}(x)\right| \mathrm{d}x + \int_L^{{}^{h}/_{2}} \left|K_{\mathrm{B}}^{(p)}(x)\right| \mathrm{d}x \\ &\le \int_0^L \left|K^{(p)}(x)\right| \mathrm{d}x + p! \left(\frac{3}{2}\right)^p \left(\frac{|a_{p-1}|}{(p-1)!} + \sum_{l=1}^{p-2} \binom{p-1+l}{l} \frac{r^{-l}}{2^l (p-1-l)!} |a_{p-1-l}|\right) \end{aligned} \tag{3.69}$$

with $r := {}^{h}/_{2} - L$ and $a_j := K^{(j)}(L)$, $j = 1, \dots, p-1$.

Thus, when considering a fixed function K in a fixed interval $[-L, L]$ the following holds. If we increase the size r of the regularization domain as well as the period h, the worst case error decreases, provided that the number of Fourier coefficients M grows at least proportional with h.

In the limit for $r \to \infty$ we obtain

$$\frac{D^{p-1} C}{(p-1)\pi^p} \left(\int_0^L \left|K^{(p)}(x)\right| \mathrm{d}x + p! \left(\frac{3}{2}\right)^p \frac{|a_{p-1}|}{(p-1)!}\right)$$

with $D := {}^{h}/_{M}$. □

Example 3.24. As an example, we considered the symmetric function $K(x) := \mathrm{e}^{-x^2}$ in the interval $[-L, L]$ with $L := 1$. In order to compute the interpolating polynomial, we considered different periods $h > 2L$ or rather different values for $r = {}^{h}/_{2} - L$ and applied the degree of smoothness $p := 8$. We fixed 32 equispaced nodes in the interval $[-L, L]$, i.e., $D = 1/16$, and increased the period h as well as the number of Fourier coefficients M such that ${}^{h}/_{M} = {}^{1}/_{16}$.

In Figure 3.13 we plot the measured approximation errors

$$\|\tilde{K}_{\mathrm{R}} - T_M(K_{\mathrm{R}})\|_\infty := \max_{|x| \le {}^{h}/_{2}} \left|\tilde{K}_{\mathrm{R}}(x) - T_M(K_{\mathrm{R}})(x)\right|$$

as well as the corresponding estimates obtained via (3.68) and (3.69) with $C := 4$. Therefor, we considered periods $h \in [2.5, 4]$ or rather mesh sizes $M \in \{40, \dots, 64\}$. We can see that the actual errors are only roughly estimated from above. However, the estimates give at least a good qualitative description of what is observed. □

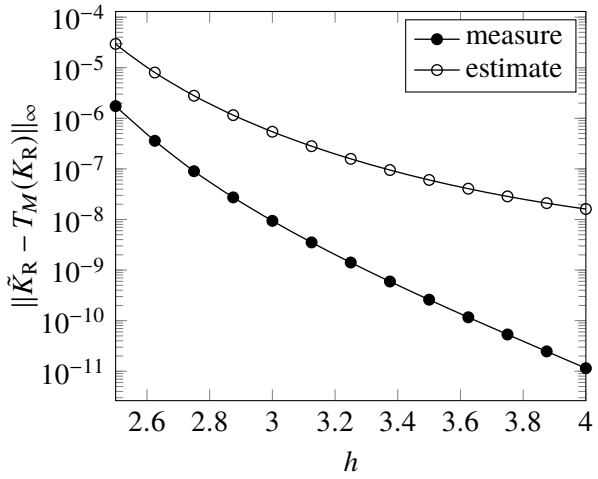


Figure 3.13: Measured and estimated approximation errors for the regularized kernel K_{R}, where we set $K(x) := \mathrm{e}^{-x^2}$ and $p := 8$. We considered different periods $h > 2L = 2$, where we chose the number of Fourier coefficients $M \in 2\mathbb{N}$ such that ${}^{h}/_{M} = {}^{1}/_{16}$.

3.4.2 Fast summation with radial kernels in higher dimensions

In the case of higher dimensions we restrict our considerations to radial kernel functions, i.e., we want to compute

$$f_j := \sum_{i=1}^{N} \alpha_i K(\|\boldsymbol{x}_i - \boldsymbol{x}_j\|),$$

where the kernel Function K, the coefficients $\alpha_j \in \mathbb{R}$ and nodes $\boldsymbol{x}_j \in \mathbb{R}^d$, $j = 1, \dots, N$, are given.

If we have an approximation of the form

$$K(\|\boldsymbol{x}\|) \approx \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}} \hat{b}_{\boldsymbol{l}}\, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l}^{\mathsf{T}} \boldsymbol{H}^{-1} \boldsymbol{x}} \tag{3.70}$$

with $\boldsymbol{M} \in 2\mathbb{N}^d$ large enough and a regular matrix $\boldsymbol{H} \in \mathbb{R}^{d\times d}$ such that $\boldsymbol{H}^{-1}\boldsymbol{x} \in [-1/2, 1/2]^d$, we obtain

$$f_j \approx \sum_{\boldsymbol{l}\in\mathcal{I}_{\boldsymbol{M}}} \hat{b}_{\boldsymbol{l}} \left(\sum_{i=1}^{N} \alpha_i \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l}^\top \boldsymbol{H}^{-1}\boldsymbol{x}_i} \right) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{l}^\top \boldsymbol{H}^{-1}\boldsymbol{x}_j},$$

analogously to the one-dimensional case. An efficient computation is again possible via the NFFT and its adjoint.

The periodization as well as the truncation techniques work analogously to the one-dimensional case. Hence, we restrict our considerations to the regularization approach in the following.

Regularization

We assume that we have $\|\boldsymbol{x}_j\| \leq L/2$ or rather $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| \leq L$ and construct the regularized kernel function $K_{\mathrm{R}} : [-h/2, h/2]^d \to \mathbb{R}$ as follows.

$$K_{\mathrm{R}}(\boldsymbol{x}) := \begin{cases} K(\|\boldsymbol{x}\|) & : \|\boldsymbol{x}\| \leq L, \\ K_{\mathrm{B}}(\|\boldsymbol{x}\|) & : L < \|\boldsymbol{x}\| \leq h/2, \\ K_{\mathrm{B}}(h/2) & : \boldsymbol{x} \in [-h/2, h/2]^d, \|\boldsymbol{x}\| > h/2, \end{cases}$$

where we choose $h \geq 2L$. The function $K_{\mathrm{B}} : [L, h/2]$ is constructed such that

$$\frac{\mathrm{d}}{\mathrm{d}x^k} K_{\mathrm{B}}(x)\bigg|_{x=L} = \frac{\mathrm{d}}{\mathrm{d}x^k} K(x)\bigg|_{x=L} \qquad \forall\, k = 0, \ldots, p-1, \tag{3.71}$$

$$\frac{\mathrm{d}}{\mathrm{d}x^k} K_{\mathrm{B}}(x)\bigg|_{x=h/2} = 0 \qquad \forall\, k = 1, \ldots, p-1. \tag{3.72}$$

Thus, the function

$$\tilde{K}_{\mathrm{R}}(\boldsymbol{x}) := K_{\mathrm{R}}(\boldsymbol{x} + h\boldsymbol{n}) \quad \text{ with } \boldsymbol{n} \in \mathbb{Z}^d : \boldsymbol{x} + h\boldsymbol{n} \in [-h/2, h/2)^d$$

is periodic with period h in each direction and smooth of degree $p-1$, i.e., $\tilde{K}_{\mathrm{R}} \in C^{p-1}(h\mathbb{T}^d)$, cf. [PSN04]. For a graphical illustration see Figure 3.14.

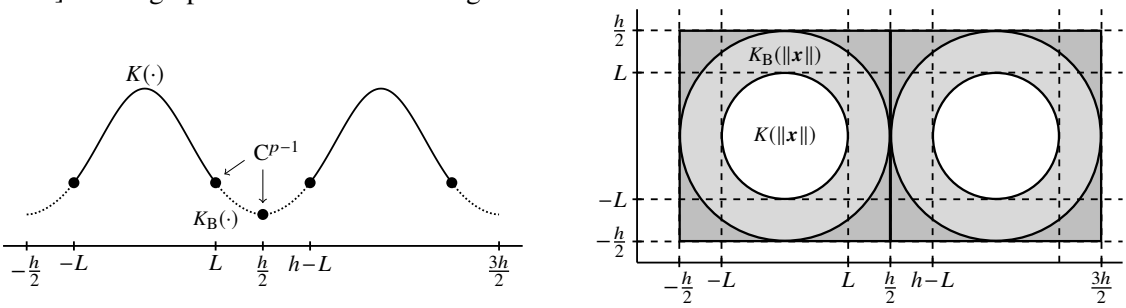


Figure 3.14: The function $K_{\mathrm{B}} : [L, h/2] \to \mathbb{R}$ is constructed such that the values of the first $p-1$ derivatives coincide with those of the kernel function K at $x = L$ and that the first $p-1$ derivatives vanish at $x = h/2$ (left). The regularized kernel $K_{\mathrm{R}} : [-h/2, h/2]^d \to \mathbb{R}$ is radial over the domain $\{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\| \leq h/2\}$ and is equal to $K_{\mathrm{B}}(h/2)$ otherwise. Thus, the periodic function $\tilde{K}_{\mathrm{R}}$ is continuous up to order $p-1$ (right).

The periodic function $\tilde{K}_{\mathrm{R}}$ can be approximated by a trigonometric polynomial and we obtain

$$K(\|\boldsymbol{x}\|) = \tilde{K}_{\mathrm{R}}(\boldsymbol{x}) \approx \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}} \hat{b}_{\boldsymbol{l}}\, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l}^{\top} \boldsymbol{x}/h} \quad \forall\, \boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\| \leq L,$$

where $\boldsymbol{M} \in 2\mathbb{N}^d$. This is equivalent to (3.70) with $\boldsymbol{H} = h\boldsymbol{I}_{d\times d}$, where we denote by $\boldsymbol{I}_{d\times d}$ the identity matrix. The Fourier coefficients $\hat{b}_{\boldsymbol{l}}$ may be computed via the FFT after sampling the function K_{R} at $|\mathcal{I}_{\boldsymbol{M}}|$ equispaced nodes in $[-h/2, h/2)^2$, i.e.,

$$\hat{b}_{\boldsymbol{l}} := \frac{1}{|\mathcal{I}_{\boldsymbol{M}}|} \sum_{\boldsymbol{j} \in \mathcal{I}_{\boldsymbol{M}}} K_{\mathrm{R}}\left(h(\boldsymbol{x} \odot \boldsymbol{M}^{-1})\right) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{j}^{\top}(\boldsymbol{l} \odot \boldsymbol{M}^{-1})}, \quad \boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}.$$

Note that we now only have $2p-1$ interpolation conditions, see (3.71) and (3.72), instead of $2p$ as in the one-dimensional case. This is due to the fact that the function value for $x = h/2$ is unknown. Thus, a modified version of the two point Taylor interpolation approach is used, see Lemma 3.25. In our setting we have to set $m := L/2 + h/4$, $r := h/4 - L/2$, $a_j := K^{(j)}(L)$ and $b_j := 0$.

Lemma 3.25. *Let an interval $[m-r, m+r], r > 0$, and the interpolation values $a_j = K^{(j)}(m-r)$, $j = 0, \ldots, p-1$, and $b_j = K^{(j)}(m+r)$, $j = 1, \ldots, p-1$, be given. For $y = \frac{x-m}{r}$ the polynomial*

$$\begin{aligned} Q(x) := &\sum_{j=0}^{p-2} I(p-1, j, y) r^{j+1} a_{j+1} + \sum_{j=0}^{p-2} I(p-1, j, -y)(-r)^{j+1} b_{j+1} \\ &- \sum_{j=0}^{p-2} I(p-1, j, -1) r^{j+1} a_{j+1} - \sum_{j=0}^{p-2} I(p-1, j, 1)(-r)^{j+1} b_{j+1} + a_0 \end{aligned}$$

of degree $2p-2$ satisfies the interpolation conditions $Q^{(j)}(m-r) = a_j$, $j = 0, \ldots, p-1$, and $Q^{(j)}(m+r) = b_j$, $j = 1, \ldots, p-1$. The polynomials $I(p, j, y)$ are given by

$$I(p, j, y) := \sum_{k=0}^{p-1-j} \binom{p-1+k}{k} \frac{1}{j! 2^p 2^k} \sum_{l=0}^{p} \frac{p!}{(p-l)!} \frac{(k+j)!}{(k+1+j+l)!} (1-y)^{p-l} (1+y)^{k+j+1+l}.$$

Proof. See Theorem C.2 in [NPP15a]. ∎

4 The P^{2}NFFT method for charges and dipoles

The particle-particle NFFT method (P^{2}NFFT) combines the Ewald formulas, as introduced in Section 2.2, and the different NFFT modules, see Section 3.2, in order to compute the electrostatic particle interactions efficiently. The method has originally been developed for the treatment of pure charge systems under 3d-periodic boundary conditions, see [PP11, PP13, Pip15]. An extension of the method to mixed periodic and open boundary conditions was introduced in [NP13, NPP15a, NPP15b].

In order to generalize the method for systems containing both charges as well as dipoles we need the adjoint gradient NFFT and the Hessian NFFT, as introduced in Section 3.2.2. The latter is solely needed for the computation of the dipole forces. A summary of the P^{2}NFFT method for charges and dipoles is given in Section 4.1, where we discuss the different types of periodic boundary conditions separately. Section 4.2 contains a detailed description of the current implementation, where we summarize the P^{2}NFFT for various types of periodic boundary conditions in Algorithm 4.3. The approximation errors, caused by applying the different NFFT modules, are discussed in Section 4.3. Note that the generalization of the method for dipole-dipole and charge-dipole interactions has also been considered in [Nes16b, HNP17].

4.1 Computational approach

The efficient evaluation of the potentials $\phi(j)$, fields $\boldsymbol{E}(j)$ and field gradients $\boldsymbol{G}(j)$ is more or less straight forward in the 3d-periodic case, since the Fourier coefficients are known analytically, see Section 2.2.1. The 3d-periodic case is considered in Section 4.1.1. In order to treat also mixed as well as open boundary conditions we combine the corresponding Ewald formulas with the NFFT based fast summation approach, which is described in Section 4.1.2.

4.1.1 3d-periodic boundary conditions

In the following we restrict our considerations to the computation of the potentials, fields as well as field gradients. The energies, forces as well as the dipole torques are obtained easily afterward by simple multiplications, see (2.13), (2.14), (2.15) and (2.16).

Computation of the potentials

Applying the Ewald splitting (2.17) to the single summands of the potential, see (2.3), and computing the spherical limit (2.29) gives, cf. Lemma 2.5,

$$\phi(j) = \phi^{\text{short}}(j) + \phi^{\text{F}}(j) + \phi^{\mathbf{0}}(j) - \phi^{\text{self}}(j),$$

where we define the short range potential

$$\phi^{\text{short}}(j) := {\sum_{\boldsymbol{n}\in\mathcal{S}}}' \left(\sum_{i=1}^{N_{\text{c}}} q_i \frac{\operatorname{erfc}(\alpha\|\boldsymbol{x}_{ij}+\boldsymbol{L}\boldsymbol{n}\|)}{\|\boldsymbol{x}_{ij}+\boldsymbol{L}\boldsymbol{n}\|} + \sum_{i=N_{\text{c}}+1}^{N} \boldsymbol{\mu}_i^{\top} \nabla_{\boldsymbol{x}_i} \frac{\operatorname{erfc}(\alpha\|\boldsymbol{x}_{ij}+\boldsymbol{L}\boldsymbol{n}\|)}{\|\boldsymbol{x}_{ij}+\boldsymbol{L}\boldsymbol{n}\|} \right) \tag{4.1}$$

with $\mathcal{S} := \mathbb{Z}^3$, and the Fourier space potential

$$\phi^{\mathrm{F}}(j) := \frac{1}{\pi V} \sum_{\boldsymbol{k} \in \mathbb{Z}^3} \hat{\psi}(\boldsymbol{k}) \left(\sum_{i=1}^{N_{\mathrm{c}}} q_i \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{L}^{-1} \boldsymbol{x}_i} + \sum_{i=N_{\mathrm{c}}+1}^{N} \boldsymbol{\mu}_i^\top \nabla_{\boldsymbol{x}_i} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{L}^{-1} \boldsymbol{x}_i} \right) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{L}^{-1} \boldsymbol{x}_j}$$

with the Fourier coefficients $\hat{\psi}(\boldsymbol{k})$ as defined in (2.31). The $\boldsymbol{k} = \boldsymbol{0}$ parts and the self potentials are defined in (2.32) and (2.24), respectively, and can be computed easily with only $O(N)$ arithmetic operations.

As already mentioned before, the complementary error function as well as its gradient tend to zero exponentially fast and, thus, the short range parts can be obtained via a direct summation, i.e., the infinite sum is simply truncated so that only distances $\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\|$ smaller than some appropriate near field cutoff radius r_{cut} have to be considered. If a sufficiently homogeneous particle distribution is assumed, this can be realized with a linked cell algorithm [FS02] in $O(N)$ arithmetic operations, since the single particles only interact with a fixed number of neighbors. In case of very heterogeneous distributions a combination with the FMM is possible, as presented in [DK00].

Also the Fourier coefficients $\hat{\psi}(\boldsymbol{k})$ tend to zero exponentially fast so that the infinite Fourier space sum can be truncated as well. In the following we define the structure factors

$$S_{\mathrm{c}}(\boldsymbol{k}) := \sum_{i=1}^{N_{\mathrm{c}}} q_i \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{L}^{-1} \boldsymbol{x}_i} \tag{4.2}$$

and

$$S_{\mathrm{d}}(\boldsymbol{k}) := \sum_{i=N_{\mathrm{c}}+1}^{N} \boldsymbol{\mu}_i^\top \nabla_{\boldsymbol{x}_i} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{L}^{-1} \boldsymbol{x}_i} = \sum_{i=N_{\mathrm{c}}+1}^{N} (\boldsymbol{L}^{-1} \boldsymbol{\mu}_i)^\top \nabla_{\boldsymbol{y}_i} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{y}_i} \Big|_{\boldsymbol{y}_i = \boldsymbol{L}^{-1} \boldsymbol{x}_i}. \tag{4.3}$$

In order to compute the Fourier space potentials efficiently we truncate the infinite sum, i.e., we replace the set $\mathbb{Z}^3$ by an index set $\mathcal{I}_{\boldsymbol{M}} \subset \mathbb{Z}^3$, as introduced in (3.2), for some appropriate cutoff parameter $\boldsymbol{M} \in 2\mathbb{N}^3$. The structure factors $S_{\mathrm{c}}(\boldsymbol{k})$, $\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}$, are approximated via the adjoint NFFT, and for the approximation of $S_{\mathrm{d}}(\boldsymbol{k})$, $\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}$, we use the adjoint gradient NFFT. We denote the obtained approximations by $S_{\mathrm{c},\approx}(\boldsymbol{k})$ and $S_{\mathrm{d},\approx}(\boldsymbol{k})$, respectively. Note that the NFFT algorithms have to be applied to the nodes $-\boldsymbol{L}^{-1}\boldsymbol{x}_j \in \mathbb{T}^3$. In order to approximate the sums $S_{\mathrm{d}}(\boldsymbol{k})$ via the adjoint NFFT we have to set $\boldsymbol{f}_j := \boldsymbol{L}^{-1}\boldsymbol{\mu}_j$ in (3.31), as indicated in (4.3), which results from applying the chain rule.

Afterward, the outer sums are obtained by applying the NFFT, i.e.,

$$\begin{aligned} \phi^{\mathrm{F}}(j) \approx \phi_{\mathrm{t}}^{\mathrm{F}}(j) &:= \frac{1}{\pi V} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k}) \left[S_{\mathrm{c}}(\boldsymbol{k}) + S_{\mathrm{d}}(\boldsymbol{k}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{L}^{-1} \boldsymbol{x}_j} && (4.4) \\ &\approx \frac{1}{\pi V} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k}) \left[S_{\mathrm{c},\approx}(\boldsymbol{k}) + S_{\mathrm{d},\approx}(\boldsymbol{k}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{L}^{-1} \boldsymbol{x}_j} \\ &\approx \frac{1}{\pi V} \left(\sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k}) \left[S_{\mathrm{c},\approx}(\boldsymbol{k}) + S_{\mathrm{d},\approx}(\boldsymbol{k}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{L}^{-1} \boldsymbol{x}_j} \right)_{\approx} =: \phi_{\mathrm{p2nfft}}^{\mathrm{F}}(j). && (4.5) \end{aligned}$$

In summary, several approximations are combined. Namely, at first the infinite Ewald sums are truncated and, secondly, the remaining finite sums are approximated based on the adjoint NFFT, the adjoint gradient NFFT and the NFFT, which we shortly summarize via

$$\phi^{\mathrm{F}}(j) \overset{\text{Ewald truncation}}{\approx} \phi_{\mathrm{t}}^{\mathrm{F}}(j) \overset{\text{NFFT modules}}{\approx} \phi_{\mathrm{p2nfft}}^{\mathrm{F}}(j).$$

Note that the approximations $S_{\mathrm{d},\approx}(\boldsymbol{k})$ can be obtained via the $\mathrm{i}\boldsymbol{k}$-differentiation approach or the analytic differentiation approach, i.e.,

$$S_{\mathrm{d},\approx}(\boldsymbol{k}) \in \left\{S_{\mathrm{d,ik}}(\boldsymbol{k}), S_{\mathrm{d,ad}}(\boldsymbol{k})\right\},$$

following the notations introduced by (3.32) and (3.33).

Computation of the fields

The Ewald representations of the fields, as defined in (2.12), read as

$$\boldsymbol{E}(j) = \boldsymbol{E}^{\mathrm{short}}(j) + \boldsymbol{E}^{\mathrm{F}}(j) + \boldsymbol{E}^{\boldsymbol{0}}(j) - \boldsymbol{E}^{\mathrm{self}}(j),$$

where the short range parts, the Fourier space parts and the $\boldsymbol{k} = \boldsymbol{0}$ terms are directly obtained by applying the operator $-\nabla_{\boldsymbol{x}_j}$ to the corresponding potential terms, i.e., we have

$$\boldsymbol{E}^{\mathrm{short}}(j) = -\nabla_{\boldsymbol{x}_j}\phi^{\mathrm{short}}(j), \tag{4.6}$$

$$\boldsymbol{E}^{\mathrm{F}}(j) = -\frac{\nabla_{\boldsymbol{x}_j}}{\pi V}\sum_{\boldsymbol{k}\in\mathbb{Z}^3}\hat{\psi}(\boldsymbol{k})\left[S_{\mathrm{c}}(\boldsymbol{k}) + S_{\mathrm{d}}(\boldsymbol{k})\right]\mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_j},$$

$$\boldsymbol{E}^{\boldsymbol{0}}(j) = -\frac{4\pi}{3V}\boldsymbol{D}, \tag{4.7}$$

where the $\boldsymbol{k} = \boldsymbol{0}$ part is easily obtained from (2.36).

The self field $\boldsymbol{E}^{\mathrm{self}}(j)$ is obtained by computing the limits

$$\lim_{\boldsymbol{x}\to\boldsymbol{0}}\nabla\frac{\mathrm{erf}(\alpha\|\boldsymbol{x}\|)}{\|\boldsymbol{x}\|} = \boldsymbol{0},$$

which follows due to symmetry, and

$$\begin{aligned}&\lim_{\boldsymbol{x}\to\boldsymbol{0}}\nabla(\boldsymbol{\mu}^\top\nabla)\frac{\mathrm{erf}(\alpha\|\boldsymbol{x}\|)}{\|\boldsymbol{x}\|}\\ &= \lim_{\boldsymbol{x}\to\boldsymbol{0}}\left(\frac{\mathrm{erf}(\alpha\|\boldsymbol{x}\|)}{\|\boldsymbol{x}\|^3} - \frac{2\alpha}{\sqrt{\pi}\|\boldsymbol{x}\|^2}\mathrm{e}^{-\alpha^2\|\boldsymbol{x}\|^2}\right)\left(\boldsymbol{\mu} - \frac{3\boldsymbol{\mu}^\top\boldsymbol{x}}{\|\boldsymbol{x}\|^2}\boldsymbol{x}\right) + \frac{4\alpha^3}{\sqrt{\pi}\|\boldsymbol{x}\|^2}\mathrm{e}^{-\alpha^2\|\boldsymbol{x}\|^2}(\boldsymbol{\mu}^\top\boldsymbol{x})\boldsymbol{x} = \frac{4\alpha^3}{3\sqrt{\pi}}\boldsymbol{\mu},\end{aligned}$$

which follows from

$$\lim_{r\to 0}\frac{\mathrm{erf}(\alpha r)}{r^3} - \frac{2\alpha}{\sqrt{\pi}r^2}\mathrm{e}^{-\alpha^2 r^2} = \frac{4\alpha^3}{3\sqrt{\pi}},$$

where we simply apply l'Hospital's rule. In summary, we end up with the following representation of the self field.

$$\boldsymbol{E}^{\mathrm{self}}(j) = \begin{cases}\boldsymbol{0} & : j = 1,\dots,N_{\mathrm{c}},\\ -\frac{4\alpha^3}{3\sqrt{\pi}}\boldsymbol{\mu}_j & : j = N_{\mathrm{c}}+1,\dots,N.\end{cases} \tag{4.8}$$

The computation of the self fields and $\boldsymbol{k} = \boldsymbol{0}$ parts is again straight forward. Analogously to the short range parts of the potentials, the short range fields can be obtained via a direct summation scheme.

The Fourier space parts are approximated in the same manner as the Fourier space parts of the potentials, where we use the gradient NFFT instead of the standard NFFT in order to compute the outer sums.

This can be expressed via

$$\boldsymbol{E}^{\mathrm{F}}(j) \approx \boldsymbol{E}^{\mathrm{F}}_{\mathrm{t}}(j) := -\frac{\nabla_{\boldsymbol{x}_j}}{\pi V} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k}) \left[S_{\mathrm{c}}(\boldsymbol{k}) + S_{\mathrm{d}}(\boldsymbol{k})\right] \mathrm{e}^{-2\pi \mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_j} \tag{4.9}$$

$$\approx -\frac{1}{\pi V}\left(\nabla_{\boldsymbol{x}_j} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k}) \left[S_{\mathrm{c},\approx}(\boldsymbol{k}) + S_{\mathrm{d},\approx}(\boldsymbol{k})\right] \mathrm{e}^{-2\pi \mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_j}\right)_{\approx} =: \boldsymbol{E}^{\mathrm{F}}_{\mathrm{p2nfft}}(j). \tag{4.10}$$

Note that the outer sums can be obtained via the $\mathrm{i}\boldsymbol{k}$-differentiation or the analytic differentiation approach, see (3.25) and (3.27). In order to compute the partial derivatives correctly, we have to pay attention to the multiplication with the matrix $\boldsymbol{L}$. With $\boldsymbol{y}_j := \boldsymbol{L}^{-1}\boldsymbol{x}_j$ we obtain $\nabla_{\boldsymbol{x}_j} = \boldsymbol{L}^{-\top}\nabla_{\boldsymbol{y}_j}$ and thus

$$\boldsymbol{E}^{\mathrm{F}}_{\mathrm{t}}(j) = -\boldsymbol{L}^{-\top}\nabla_{\boldsymbol{y}_j} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \frac{\hat{\psi}(\boldsymbol{k})}{\pi V} \left[S_{\mathrm{c}}(\boldsymbol{k}) + S_{\mathrm{d}}(\boldsymbol{k})\right] \mathrm{e}^{-2\pi \mathrm{i}\boldsymbol{k}^{\top}\boldsymbol{y}_j}\Big|_{\boldsymbol{y}_j=\boldsymbol{L}^{-1}\boldsymbol{x}_j},$$

i.e., in order to approximate the Fourier space parts of the fields we apply the gradient NFFT for the nodes $-\boldsymbol{L}^{-1}\boldsymbol{x}_j$ and multiply by $\boldsymbol{L}^{-\top}$ afterward.

Remark 4.1 (Systematic errors)**.** In Section 2.3.3 we presented error estimates regarding the approximation

$$\boldsymbol{F}^{\mathrm{F}}(j) \approx \boldsymbol{F}^{\mathrm{F}}_{\mathrm{t}}(j) := q_j \boldsymbol{E}^{\mathrm{F}}_{\mathrm{t}}(j), \quad j = 1, \ldots, N_{\mathrm{c}}, \tag{4.11}$$

amongst others. Note that we now deal with non symmetric index sets $\mathcal{I}_{\boldsymbol{M}}$, cf. (3.2), in order to make the NFFT algorithms applicable. This gives rise to the following problem.

In case of charge-charge interactions the diagonal terms with $i = j$ in the exact Fourier space parts of the forces are given by, cf. (2.85),

$$\frac{2\mathrm{i}q_j^2}{V} \sum_{\boldsymbol{k}\in\mathbb{Z}^3} \hat{\psi}(\boldsymbol{k})(\boldsymbol{L}^{-\top}\boldsymbol{k}) = \boldsymbol{0}, \quad j = 1, \ldots, N_{\mathrm{c}},$$

which is due to the underlying symmetry. This does not hold for the approximation obtained via (4.9) and (4.11), since the index set $\mathcal{I}_{\boldsymbol{M}}$ is not symmetric. We have

$$\frac{2\mathrm{i}q_j^2}{V} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})(\boldsymbol{L}^{-\top}\boldsymbol{k}) = \frac{2\mathrm{i}q_j^2}{V} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}\setminus\mathcal{I}^*_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})(\boldsymbol{L}^{-\top}\boldsymbol{k}),$$

where we define the symmetric index set

$$\mathcal{I}^*_{\boldsymbol{M}} := \{-M_1/2 + 1, \ldots, M_1/2 - 1\} \times \ldots \times \{-M_3/2 + 1, \ldots, M_3/2 - 1\} \subset \mathcal{I}_{\boldsymbol{M}}.$$

Thus, the presented approximation causes the computation of undesired self interactions, which results in systematic errors of the form

$$\Delta\boldsymbol{F}^{\mathrm{F,sys}}(j) := \frac{2\mathrm{i}q_j^2}{V} \sum_{\boldsymbol{k}\in\mathbb{Z}^3\setminus(\mathcal{I}_{\boldsymbol{M}}\setminus\mathcal{I}^*_{\boldsymbol{M}})} \hat{\psi}(\boldsymbol{k})(\boldsymbol{L}^{-\top}\boldsymbol{k}), \quad j = 1, \ldots, N_{\mathrm{c}},$$

where we can replace the index set $\mathbb{Z}^3 \setminus (\mathcal{I}_{\boldsymbol{M}} \setminus \mathcal{I}^*_{\boldsymbol{M}})$ by $\mathcal{I}_{-\boldsymbol{M}} \setminus \mathcal{I}^*_{\boldsymbol{M}}$, due to the underlying symmetry. This means that only terms with $k_1 = M_1/2 \vee k_2 = M_2/2 \vee k_3 = M_3/2$ are present and, thus, we may assume these systematic errors to be negligible compared to the off-diagonal contributions to the overall error.

These systematic errors are not incorporated in the estimates as presented in Section 2.3.3. We denote the resulting systematic error in the forces of the charges by

$$\Delta F_{\text{c.c.}}^{\text{F,sys}} := \sqrt{\frac{1}{N_{\text{c}}} \sum_{j=1}^{N_{\text{c}}} \left\| \Delta \boldsymbol{F}^{\text{F,sys}}(j) \right\|^2} = \frac{2\sqrt{Q}}{V\sqrt{N_{\text{c}}}} \left\| \sum_{\boldsymbol{k} \in \mathcal{I}_{-\boldsymbol{M}} \setminus \mathcal{I}_{\boldsymbol{M}}^*} \hat{\psi}(\boldsymbol{k}) (\boldsymbol{L}^{-\top} \boldsymbol{k}) \right\|. \tag{4.12}$$

In an analogue manner, we obtain the following systematic errors in case of dipole-dipole interactions, cf. (2.91).

$$\Delta \boldsymbol{F}^{\text{F,sys}}(j) := \frac{8\pi^2 \mathrm{i} \|\boldsymbol{\mu}_j\|^2}{V} \sum_{\boldsymbol{k} \in \mathcal{I}_{-\boldsymbol{M}} \setminus \mathcal{I}_{\boldsymbol{M}}^*} \hat{\psi}(\boldsymbol{k}) \|\boldsymbol{L}^{-\top} \boldsymbol{k}\|^2 \cos^2 \gamma_{j,\boldsymbol{k}}(\boldsymbol{L}^{-\top} \boldsymbol{k}), \quad j = N_{\text{c}} + 1, \ldots, N.$$

In order to estimate the overall systematic rms force error in the forces of the dipoles, we may replace the cosine terms by the mean value $\pi^{-1} \int_0^\pi \cos^2 x \, \mathrm{d}x = 1/2$ to obtain

$$\Delta F_{\text{d.d.}}^{\text{F,sys}} := \sqrt{\frac{1}{N_{\text{d}}} \sum_{j=N_{\text{c}}+1}^{N} \left\| \Delta \boldsymbol{F}^{\text{F,sys}}(j) \right\|^2} \approx \frac{4\pi^2 \sqrt{\mathcal{M}}}{V\sqrt{N_{\text{d}}}} \left\| \sum_{\boldsymbol{k} \in \mathcal{I}_{-\boldsymbol{M}} \setminus \mathcal{I}_{\boldsymbol{M}}^*} \hat{\psi}(\boldsymbol{k}) \|\boldsymbol{L}^{-\top} \boldsymbol{k}\|^2 (\boldsymbol{L}^{-\top} \boldsymbol{k}) \right\|. \tag{4.13}$$

This all refers to the approximation errors caused by truncating the Fourier space part in the Ewald summation formulas. A discussion of the errors resulting from the application of the different NFFT variants follows in Section 4.3.

We remark that the influence of the above discussed systematic errors on the overall errors has not been studied so far. In other words, it is widely unknown how much the usage of non symmetric index sets $\mathcal{I}_{\boldsymbol{M}}$ might influence the accuracy in general. □

Computation of the field gradients

The Ewald splitting of the negative field gradient reads as

$$\boldsymbol{G}(j) = \boldsymbol{G}^{\text{short}}(j) + \boldsymbol{G}^{\text{F}}(j) - \boldsymbol{G}^{\text{self}}(j),$$

where the short range part as well as the Fourier space part are directly obtained by applying the operator $-\nabla_{\boldsymbol{x}_j} \nabla_{\boldsymbol{x}_j}^\top$ to the corresponding potential terms, i.e., we have

$$\begin{aligned} \boldsymbol{G}^{\text{short}}(j) &= -\nabla_{\boldsymbol{x}_j} \nabla_{\boldsymbol{x}_j}^\top \phi^{\text{short}}(j), \\ \boldsymbol{G}^{\text{F}}(j) &= -\frac{\nabla_{\boldsymbol{x}_j} \nabla_{\boldsymbol{x}_j}^\top}{\pi V} \sum_{\boldsymbol{k} \in \mathbb{Z}^3} \hat{\psi}(\boldsymbol{k}) \left[S_{\text{c}}(\boldsymbol{k}) + S_{\text{d}}(\boldsymbol{k}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^\top \boldsymbol{L}^{-1} \boldsymbol{x}_j}. \end{aligned} \tag{4.14}$$

The self interaction term $\boldsymbol{G}^{\text{self}}(j)$ is computed analogously to the self field $\boldsymbol{E}^{\text{self}}(j)$, see (4.8). We obtain

$$\boldsymbol{G}^{\text{self}}(j) = \begin{cases} \frac{4\alpha^3}{3\sqrt{\pi}} q_j \boldsymbol{I}_{3,3} & : j = 1, \ldots, N_{\text{c}}, \\ \boldsymbol{0}_{3,3} & : j = N_{\text{c}} + 1, \ldots, N. \end{cases} \tag{4.15}$$

Again, the short range parts can be obtained via a direct summation and the computation of the self terms is straight forward.

In order to approximate the Fourier space parts efficiently, we truncate the infinite sum over $\boldsymbol{k} \in \mathbb{Z}^3$, compute the inner sums via the adjoint as well as the adjoint gradient NFFT and evaluate the outer sums with the Hessian NFFT, i.e.,

$$\begin{aligned} \boldsymbol{G}^{\mathrm{F}}(j) \approx \boldsymbol{G}^{\mathrm{F}}_{\mathrm{t}}(j) &:= -\frac{\nabla_{\boldsymbol{x}_j} \nabla_{\boldsymbol{x}_j}^{\top}}{\pi V} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k}) \left[S_{\mathrm{c}}(\boldsymbol{k}) + S_{\mathrm{d}}(\boldsymbol{k}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{L}^{-1} \boldsymbol{x}_j} \\ &\approx -\frac{1}{\pi V} \left(\nabla_{\boldsymbol{x}_j} \nabla_{\boldsymbol{x}_j}^{\top} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k}) \left[S_{\mathrm{c},\approx}(\boldsymbol{k}) + S_{\mathrm{d},\approx}(\boldsymbol{k}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{L}^{-1} \boldsymbol{x}_j} \right)_{\approx} =: \boldsymbol{G}^{\mathrm{F}}_{\mathrm{p2nfft}}(j). \end{aligned} \tag{4.16}$$

Again, we have to pay attention to the multiplication with the matrix $\boldsymbol{L}$. We obtain

$$\boldsymbol{G}^{\mathrm{F}}_{\mathrm{t}}(j) = \boldsymbol{L}^{-\top} \left(-\nabla_{\boldsymbol{y}_j} \nabla_{\boldsymbol{y}_j}^{\top} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \frac{\hat{\psi}(\boldsymbol{k})}{\pi V} \left[S_{\mathrm{c}}(\boldsymbol{k}) + S_{\mathrm{d}}(\boldsymbol{k}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{y}_j} \Big|_{\boldsymbol{y}_j = \boldsymbol{L}^{-1} \boldsymbol{x}_j} \right) \boldsymbol{L}^{-1},$$

i.e., we can simply apply the Hessian NFFT for the nodes $-\boldsymbol{L}^{-1}\boldsymbol{x}_j$ and multiply with $\boldsymbol{L}^{-\top}$ from left and by $\boldsymbol{L}^{-1}$ from right afterward.

Finally, the Fourier space parts of the forces acting on the dipoles may be obtained via

$$\boldsymbol{F}^{\mathrm{F}}(j) \approx \boldsymbol{G}^{\mathrm{F}}_{\mathrm{p2nfft}}(j) \boldsymbol{\mu}_j, \quad j = N_{\mathrm{c}} + 1, \ldots, N.$$

The P²NFFT method is summarized in Algorithm 4.3, where 3d-periodic boundary conditions are obtained for $\rho = 3$.

Remark 4.2 (Complexity)**.** The computation of the self interactions and the $\boldsymbol{k} = \boldsymbol{0}$ parts is straight forward and requires only $O(N)$ arithmetic operations. The same applies to the near field computations, provided that the particles are equally distributed. In order to approximate the Fourier space parts we need $O(N + |\mathcal{I}_{\boldsymbol{M}}| \log |\mathcal{I}_{\boldsymbol{M}}|)$ arithmetic operations, cf. Section 3.2. A more detailed discussion of the overall complexity is given later on in Remark 5.14. □

Remark 4.3 (Relation to other particle-mesh methods)**.** We remark that the classical P³M algorithm [EHL80, HE88] is in principle equivalent to the P²NFFT method for 3d-periodic boundary conditions, provided that the B-spline window function without oversampling is applied. In the context of P³M, the deconvolution step of the adjoint NFFT, see Step 3 in Algorithm 3.2, the multiplication with the Fourier coefficients $\hat{\psi}(\boldsymbol{k})$ and the deconvolution step of the NFFT, see Step 1 in Algorithm 3.1, are summarized via introducing the so called influence function. The optimal influence function depends on the applied deconvolution approach and on which rms error shall be minimized, see [DH98a, DH98b, BCH12].

Also other particle-mesh approaches may be interpreted as special cases of the P²NFFT method. For a detailed discussion on the relation of P²NFFT to P³M and other particle-mesh algorithms we refer to [Pip15, Sections 4.6 and 4.7]. □

4.1.2 Mixed periodic and open boundary conditions

In the following we combine the Ewald summation formulas for mixed periodic and open boundary conditions, see Sections 2.2.2, 2.2.3 and 2.2.4, with the NFFT based fast summation approach, see

Section 3.4, in order to generalize the P^{2}NFFT method to mixed periodic and open boundary conditions, cf. [NP13, NPP15a, NPP15b, Nes16b].

We also refer to other Fourier based algorithms handling mixed periodic boundary conditions, as for example the MMM2D, the MMM1D and the ELC method, see [AH02, AdJH02, AH05], or the spectrally accurate Ewald method for 1d- and 2d-periodic constraints, see [LT12, ST17]. Furthermore, also the fast multipole method can handle all the considered types of periodic boundary conditions, cf. [KS04, KD11, Kab12]. For algorithms showing a higher arithmetic complexity we refer to [WA97, YB99, BS04, Bró04].

2d-periodic

The computation of the short range parts $\phi^{\text{short}}(j)$, $\boldsymbol{E}^{\text{short}}(j)$ and $\boldsymbol{G}^{\text{short}}(j)$ may be realized analogously to the 3d-periodic case, see (4.1), (4.6) and (4.14), respectively. For this purpose, the lattice $\mathcal{S} = \mathbb{Z}^3$ is replaced by $\mathcal{S} = \mathbb{Z}^2 \times \{0\}$. In addition, the self interaction terms $\phi^{\text{self}}(j)$, $\boldsymbol{E}^{\text{self}}(j)$ and $\boldsymbol{G}^{\text{self}}(j)$ are the same as in the 3d-periodic case, see (2.24), (4.8) and (4.15), respectively, and can be computed easily.

The Fourier space parts of the potentials are given by, cf. (2.38),

$$\phi^{\text{F}}(j) = \frac{1}{\|\boldsymbol{\ell}_1 \times \boldsymbol{\ell}_2\|} \sum_{\boldsymbol{k} \in \mathbb{Z}^2} \sum_{i=1}^{N} \xi_i \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^{\top} \tilde{\boldsymbol{L}}^{-1} \tilde{\boldsymbol{x}}_{ij}} \Theta^{\text{p2}} \left(\|\tilde{\boldsymbol{L}}^{-\top} \boldsymbol{k}\|, x_{ij,3} \right), \tag{4.17}$$

where the function Θ^{p2} is defined in (2.39). Here, we assume periodic boundary conditions with respect to $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$, which we claim to be of the form

$$\begin{pmatrix} \boldsymbol{\ell}_1 & \boldsymbol{\ell}_2 \end{pmatrix} = \begin{pmatrix} l_{11} & l_{21} \\ l_{12} & l_{22} \\ 0 & 0 \end{pmatrix} =: \begin{pmatrix} \tilde{\boldsymbol{L}} \\ 0 \quad 0 \end{pmatrix}.$$

Note that $\Theta^{\text{p2}}(k, \cdot)$ is a symmetric C^∞ function for each $k \geq 0$. In addition, it is easy to verify that $\Theta^{\text{p2}}(k, r)$ tends to zero exponentially fast for fixed $k > 0$ and $|r| \to \infty$. In contrast, in the special case $k = 0$ we have $\Theta^{\text{p2}}(0, r) \approx -2\pi|r|$ for $|r|$ large enough. We plot some examples in Figure 4.1

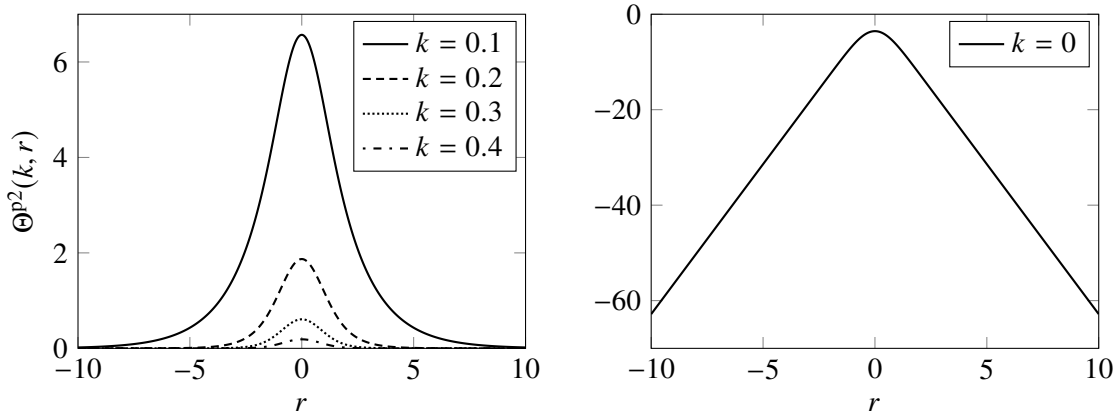


Figure 4.1: We plot the function $\Theta^{\text{p2}}(k, \cdot)$ for $\alpha := 1$ and different values of k.

It is reasonable to assume that the coordinates of the particles with respect to the nonperiodic dimension are absolutely bounded, i.e., we have $|x_{j,3}| \leq {}^{L}/_{2}$, $j = 1, \ldots, N$, for some $L > 0$. Thus, we have $-L \leq x_{ij,3} \leq L$ for all $i, j = 1, \ldots, N$.

We recall that the function values $\Theta^{\mathrm{p2}}(k,r)$ tend to zero exponentially fast for growing k, see Remark 2.9. Thus, we may truncate the Fourier series (4.17), i.e., we only include $\boldsymbol{k} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}$ with $\tilde{\boldsymbol{M}} = (M_1, M_2) \in 2\mathbb{N}^2$ large enough. Furthermore, we choose a period

$$h \geq 2L$$

and approximate each function $\Theta^{\mathrm{p2}}(k,\cdot)$, $k \in \{\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|, \boldsymbol{k} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}\}$, by a trigonometric polynomial. This ansatz is closely related to the NFFT based fast summation, see Section 3.4.1.

Basically, there are two practicable approaches to compute these approximations, namely the periodization, see Variant I in Section 3.4.1, and the regularization approach, see Variant III in Section 3.4.1. The combination of both approaches offers the following advantages.

The periodization approach is limited to functions that decay sufficiently fast in the interval $[-h/2, h/2)$. This does not apply for the regularization technique. In other words, in the case that k is small or equal to zero, we can apply the regularization approach. This allows the choice of a relatively short period $h \geq 2L$. To this end, we need to precompute the approximating trigonometric polynomial, which involves the computation of the kernel function $\Theta^{\mathrm{p2}}(k,r)$ and its derivatives at $r = \pm h/2$, the evaluation of the regularized kernel at equispaced nodes as well as the computation of the Fourier coefficients based on the FFT, as described in Section 3.4.1. For kernels that are sufficiently small in $r = \pm h/2$ we can simply switch over to applying the periodization approach, which does not require costly precomputations, since the Fourier coefficients are known in an analytically closed form, see Lemma 4.5.

Regularization: The function $\Theta^{\mathrm{p2}}(0,\cdot)$ is not well localized, i.e., a periodization of the form (3.63) does not exist. In addition, for small values of $k > 0$ the h-periodization exists, but the function value $\Theta^{\mathrm{p2}}(k, \pm h/2)$ might not be small enough, i.e., we would have to choose $h \gg 2L$ or we will obtain quite inaccurate results when applying the periodization approach.

Consequently, we apply the regularization approach in order to approximate the functions $\Theta^{\mathrm{p2}}(k,\cdot)$ for small $k \geq 0$ by trigonometric polynomials, i.e., we define the functions

$$K_{\mathrm{R}}(k,x) := \begin{cases} \Theta^{\mathrm{p2}}(k,x) & : x \in [-L, L], \\ K_{\mathrm{B}}(k,x) & : x \in (L, h-L), \end{cases} \tag{4.18}$$

where for each k the polynomial $K_{\mathrm{B}}(k,\cdot)$ is computed via the two-point Taylor interpolation, see Lemma 3.22. In other words, $K_{\mathrm{B}}(k,\cdot)$ is constructed such that the $2p$ interpolation conditions (3.64) and (3.65) are fulfilled. We refer to $p \in \mathbb{N}$ as the degree of smoothness, since $K_{\mathrm{R}}(k,\cdot) \in \mathrm{C}^{p-1}([-L, h-L])$.

The periodic functions

$$\tilde{K}_{\mathrm{R}}(k,x) := K_{\mathrm{R}}(x + hn) \quad \text{with } n \in \mathbb{Z} : x + hn \in [-L, h-L),$$

are smooth on the torus $h\mathbb{T}$, or more precisely $\tilde{K}_{\mathrm{R}}(k,\cdot) \in \mathrm{C}^{p-1}(h\mathbb{T})$, and may be approximated by trigonometric polynomials, i.e.,

$$\tilde{K}_{\mathrm{R}}(k,x) \approx \sum_{l=-M_3/2}^{M_3/2-1} \hat{b}_{k,l}\, \mathrm{e}^{2\pi \mathrm{i} l x/h},$$

where $M_3 \in 2\mathbb{N}$ should be chosen large enough. The Fourier coefficients $\hat{b}_{k,l}$ are simply approximated via the FFT by using equispaced samples of $\tilde{K}_{\mathrm{R}}(k,\cdot)$ in $[-h/2, h/2)$, i.e.,

$$\hat{b}_{k,l} := \frac{1}{M_3} \sum_{j=-M_3/2}^{M_3/2-1} \tilde{K}_{\mathrm{R}}\left(k, \tfrac{jh}{M_3}\right) \mathrm{e}^{-2\pi \mathrm{i} jl/M_3}. \tag{4.19}$$

Since we have $x_{ij,3} \in [-L, L]$, we obtain

$$\Theta^{\mathrm{p2}}(\|\tilde{\boldsymbol{L}}^{-\top}\boldsymbol{k}\|, x_{ij,3}) \approx \sum_{l=-M_3/2}^{M_3/2-1} \hat{b}_{\|\tilde{\boldsymbol{L}}^{-\top}\boldsymbol{k}\|,l}\, \mathrm{e}^{2\pi \mathrm{i} l x_{ij,3}/h}$$

for all considered values $\|\tilde{\boldsymbol{L}}^{-\top}\boldsymbol{k}\| = k$.

The described regularization approach requires the computation of the derivatives up to degree $p-1$ of the functions $\Theta^{\mathrm{p2}}(k,\cdot)$. In the following we derive the needed expressions.

For $k \neq 0$ we set

$$\Theta_1(k,r) := \mathrm{e}^{2\pi k r}\operatorname{erfc}\left(\frac{\pi k}{\alpha} + \alpha r\right)$$

and

$$\Theta_{\pm}(k,r) = \Theta_1(k,r) \pm \Theta_1(k,-r),$$

which implies

$$\Theta^{\mathrm{p2}}(k,r) = \frac{\Theta_+(k,r)}{2k}.$$

The following lemma provides a recursion for the computation of the n-th derivative of the function $\Theta_+(k,\cdot)$ or rather $\Theta^{\mathrm{p2}}(k,\cdot)$, see also [NPP15a].

Lemma 4.4. *For $n \geq 2$ we have*

$$\frac{\partial^n}{\partial r^n}\Theta_+(k,r) = 4\pi^2 k^2 \frac{\partial^{n-2}}{\partial r^{n-2}}\Theta_+(k,r) - \frac{\partial^{n-2}}{\partial r^{n-2}} 8\alpha\sqrt{\pi}k\, \mathrm{e}^{-\alpha^2 r^2 - \pi^2 k^2/\alpha^2}. \tag{4.20}$$

Proof. We easily compute

$$\begin{aligned}\frac{\partial}{\partial r}\Theta_+(k,r) = {} & 2\pi k\, \mathrm{e}^{2\pi k r}\operatorname{erfc}\left(\frac{\pi k}{\alpha} + \alpha r\right) - \frac{2\alpha}{\sqrt{\pi}}\, \mathrm{e}^{2\pi k r}\, \mathrm{e}^{-(\pi k/\alpha + \alpha r)^2} \\ & - 2\pi k\, \mathrm{e}^{-2\pi k r}\operatorname{erfc}\left(\frac{\pi k}{\alpha} - \alpha r\right) + \frac{2\alpha}{\sqrt{\pi}}\, \mathrm{e}^{-2\pi k r}\, \mathrm{e}^{-(\pi k/\alpha - \alpha r)^2} = 2\pi k \Theta_-(k,r).\end{aligned}$$

In an analogue manner we get

$$\frac{\partial}{\partial r}\Theta_-(k,r) = 2\pi k \Theta_+(k,r) - \frac{4\alpha}{\sqrt{\pi}}\, \mathrm{e}^{-\pi^2 k^2/\alpha^2 - \alpha^2 r^2}$$

and thus

$$\frac{\partial^2}{\partial r^2}\Theta_+(k,r) = 4\pi^2 k^2 \Theta_+(k,r) - 8\alpha\sqrt{\pi}k\, \mathrm{e}^{-\alpha^2 r^2 - \pi^2 k^2/\alpha^2}.$$

We obtain (4.20) by differentiating $n-2$ times with respect to r on both sides. ∎

For $k = 0$ we obtain

$$\frac{\partial}{\partial r}\Theta^{\mathrm{p2}}(0,r) = -2\pi \operatorname{erf}(\alpha r) \quad \text{and} \quad \frac{\partial^2}{\partial r^2}\Theta^{\mathrm{p2}}(0,r) = -4\pi\alpha\, \mathrm{e}^{-\alpha^2 r^2}.$$

Thus, we can simply refer to the well known derivatives of the Gaussian function via

$$\frac{\partial^n}{\partial r^n}\Theta^{\mathrm{p2}}(0,r) = -4\pi\alpha \frac{\partial^{n-2}}{\partial r^{n-2}}\, \mathrm{e}^{-\alpha^2 r^2} \quad \forall\, n \geq 2. \tag{4.21}$$

Periodization: For $k \neq 0$ we consider the error

$$\sum_{n\in\mathbb{Z}} \Theta^{\mathrm{p2}}(k, r + hn) - \Theta^{\mathrm{p2}}(k, r) = \sum_{n\in\mathbb{Z}\setminus\{0\}} \Theta^{\mathrm{p2}}(k, r + hn). \tag{4.22}$$

Since $\Theta^{\mathrm{p2}}(k, x)$ is monotonically decreasing for $|x| \to \infty$ and $h > 2L$, we have

$$\Theta^{\mathrm{p2}}(k, r + hn) \leq \Theta^{\mathrm{p2}}(k, L - h) = \Theta^{\mathrm{p2}}(k, h - L) \quad \forall\, r \in [-L, L], n \in \mathbb{Z} \setminus \{0\}.$$

In other words, if $\Theta^{\mathrm{p2}}(k, h - L)$ is sufficiently small, the error (4.22) is small and we may consider the periodized version. The Fourier coefficients of the periodized function can be expressed in terms of the Fourier transform of $\Theta^{\mathrm{p2}}(k, r)$ with respect to r, see the Poisson summation formula (2.25). The corresponding analytic Fourier coefficients are stated in the following lemma.

Lemma 4.5. *For $k \neq 0$ we have*

$$\hat{\Theta}^{\mathrm{p2}}(k, \xi) := \int_{-\infty}^{\infty} \Theta^{\mathrm{p2}}(k, r)\, \mathrm{e}^{-2\pi \mathrm{i} r\xi} \mathrm{d}r = \frac{1}{\pi(k^2 + \xi^2)}\, \mathrm{e}^{-\pi^2 k^2/\alpha^2 - \pi^2\xi^2/\alpha^2}. \tag{4.23}$$

Proof. We use the short notations $\Theta'_+ := \frac{\partial}{\partial r}\Theta_+(k, r)$. As we have already seen, the function Θ_+ is a solution of the following second order differential equation, cf. (4.20) for $n := 2$.

$$\Theta''_+ = 4\pi k^2 \Theta_+ - 8\alpha\sqrt{\pi} k\, \mathrm{e}^{-\pi^2 k^2/\alpha^2}\, \mathrm{e}^{-\alpha^2 r^2}.$$

If we compute the above defined Fourier integral on both sides, we obtain

$$-4\pi^2\xi^2 \hat{\Theta}_+ = 4\pi^2 k^2 \hat{\Theta}_+ - 8\pi k\, \mathrm{e}^{-\pi^2 k^2/\alpha^2}\, \mathrm{e}^{-\pi^2\xi^2/\alpha^2},$$

where we set $\hat{\Theta}_+ := \hat{\Theta}_+(k, \xi)$, as introduced above. The obtained equation can be rewritten as

$$\hat{\Theta}_+ = \frac{2k\, \mathrm{e}^{-\pi^2 k^2/\alpha^2 - \pi^2\xi^2/\alpha^2}}{\pi(k^2 + \xi^2)}.$$

This completes the proof, since $\Theta^{\mathrm{p2}}(k, \cdot) = (2k)^{-1}\Theta_+(k, \cdot)$. ∎

Following the periodization approach, we obtain the following approximation of the function $\Theta^{\mathrm{p2}}(k, \cdot)$ via the Poisson summation formula.

$$\Theta^{\mathrm{p2}}(k, r) \approx \frac{1}{h} \sum_{l=-M_3/2}^{M_3/2-1} \hat{\Theta}^{\mathrm{p2}}(k, h^{-1} l)\, \mathrm{e}^{2\pi \mathrm{i} l r/h},$$

where $\hat{\Theta}^{\mathrm{p2}}(k, {}^{l}/_{h})$ is known in an analytically closed form, see Lemma 4.5.

Remark 4.6 (Periodization)**.** The Fourier transform (4.23) can also be computed as follows

$$\hat{\Theta}^{\mathrm{p2}}(k, \xi) = 2\sqrt{\pi} \int_0^{\alpha} \frac{\mathrm{e}^{-\pi^2 k^2/z^2}}{z^2} \int_{-\infty}^{\infty} \mathrm{e}^{-z^2 r^2}\, \mathrm{e}^{-2\pi \mathrm{i} r\xi} \mathrm{d}r\, \mathrm{d}z,$$

where we use the integral representation (2.41) of the function Θ^{p2}. The order of the integrals can be changed since for $k \neq 0$ all the integrals exist.

Via (2.27) we obtain

$$\hat{\Theta}^{\mathrm{p2}}(k, \xi) = 2\pi \int_0^{\alpha} \frac{\mathrm{e}^{-\pi^2 k^2/z^2 - \pi^2\xi^2/z^2}}{z^3} \mathrm{d}z,$$

where the remaining integral can be computed easily and we end up with the same result. The inverse Fourier transform gives

$$\Theta^{\mathrm{p2}}(k,r) = \int_{-\infty}^{\infty} \hat{\Theta}^{\mathrm{p2}}(k,\xi)\, \mathrm{e}^{2\pi \mathrm{i} r\xi} \mathrm{d}\xi.$$

Since $\hat{\Theta}^{\mathrm{p2}}(k,\xi)$ tends to zero very fast for $|\xi| \to \infty$, we may truncate the infinite integral and then approximate the finite integral via the trapezoidal quadrature rule, i.e.,

$$\Theta^{\mathrm{p2}}(k,r) \approx \int_{-K/2}^{K/2} \hat{\Theta}^{\mathrm{p2}}(k,\xi)\, \mathrm{e}^{2\pi \mathrm{i} r\xi} \mathrm{d}\xi \approx \frac{K}{m} \sum_{l=-m/2}^{m/2-1} \Theta^{\mathrm{p2}}\left(k, \tfrac{lK}{m}\right) \mathrm{e}^{2\pi \mathrm{i} rlK/m}$$

for $K > 0$ and $m \in 2\mathbb{N}$ large enough. It is easy to see that this approach is equivalent to the considered periodization approach. We remark that the method presented in [LT12] is based on this idea. □

Whereas the method presented in [LT12] depends on the periodization approach, cf. Variant I in Section 3.4.1, the truncation approach, cf. Variant II in Section 3.4.1, is applied in [MTPM02]. A more detailed comparison to these methods can be found in [NPP15a].

Summary: We summarize the precomputations that are necessary for the P^{2}NFFT method under 2d-periodic boundary conditions in the following algorithm.

Algorithm 4.1 (Precomputations 2d-periodic P^{2}NFFT)**.**
Input: Splitting parameter α, far field cutoff $\tilde{\boldsymbol{M}} = (M_1, M_2)^{\top} \in 2\mathbb{N}^2$, mesh size $M_3 \in 2\mathbb{N}$, maximum distance $L > 0$, period $h \geq 2L$, matrix $\tilde{\boldsymbol{L}} \in \mathbb{R}^{2\times 2}$.

For all $\boldsymbol{k} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}$:

i) Set $k := \|\tilde{\boldsymbol{L}}^{-\top} \boldsymbol{k}\|$.

ii) If $k = 0$ or $\Theta^{\mathrm{p2}}(k, h-L) \geq \varepsilon$:

   a) Construct the regularized kernel $K_{\mathrm{R}}(k,\cdot)$, see (4.18).

   b) Compute the coefficients $\hat{b}_{k,l}$, $l = -M_3/2, \ldots, M_3/2 - 1$, via the inverse FFT, see (4.19).

iii) Else, i.e., $\Theta^{\mathrm{p2}}(k, h-L) < \varepsilon$: Set

$$\hat{b}_{k,l} := \frac{\mathrm{e}^{-\pi^2 k^2/\alpha^2 - \pi^2 l^2/(\alpha^2 h^2)}}{h\pi(k^2 + l^2/h^2)}$$

   for all $l = -M_3/2, \ldots, M_3/2 - 1$.

iv) Set $\hat{\psi}^{\mathrm{p2}}(\boldsymbol{k}, l) := \hat{b}_{k,l}$, $l = -M_3/2, \ldots, M_3/2 - 1$.

Output: Fourier coefficients $\hat{\psi}^{\mathrm{p2}}(\boldsymbol{k}, l)$ for $(\boldsymbol{k}, l) \in \mathcal{I}_{\boldsymbol{M}}$, $\boldsymbol{M} := (M_1, M_2, M_3)^{\top}$. □

In the current implementation of the algorithm, as summarized in Section 4.2, we apply $\varepsilon := 10^{-16}$ within the second step of the above described precomputation procedure.

With the precomputed Fourier coefficients $\hat{\psi}^{\mathrm{p2}}(\boldsymbol{k}, l)$ and the matrix

$$\boldsymbol{L}_{\mathrm{p2}} := \begin{pmatrix} & \tilde{\boldsymbol{L}} & \\ & & 0 \\ 0 & 0 & h \end{pmatrix}$$

we can write

$$\phi^{\mathrm{F}}(j) \approx \phi_{\mathrm{t}}^{\mathrm{F}}(j) := \sum_{(\boldsymbol{k},l)\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p2}}(\boldsymbol{k},l)\,[S_{\mathrm{c}}(\boldsymbol{k},l) + S_{\mathrm{d}}(\boldsymbol{k},l)]\ \mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}^{\top},l)\boldsymbol{L}_{\mathrm{p2}}^{-1}\boldsymbol{x}_j}, \tag{4.24}$$

where we define the structure factors

$$S_{\mathrm{c}}(\boldsymbol{k},l) := \sum_{i=1}^{N_{\mathrm{c}}} q_i\,\mathrm{e}^{2\pi\mathrm{i}(\boldsymbol{k}^{\top},l)\boldsymbol{L}_{\mathrm{p2}}^{-1}\boldsymbol{x}_i}, \tag{4.25}$$

$$S_{\mathrm{d}}(\boldsymbol{k},l) := \sum_{i=N_{\mathrm{c}}+1}^{N} \boldsymbol{\mu}_i^{\top}\nabla_{\boldsymbol{x}_i}\,\mathrm{e}^{2\pi\mathrm{i}(\boldsymbol{k}^{\top},l)\boldsymbol{L}_{\mathrm{p2}}^{-1}\boldsymbol{x}_i}. \tag{4.26}$$

Note that the approximation (4.24) has the same structure as the one obtained for 3d-periodic boundary conditions, see (4.4). We just formally replace $(\pi V)^{-1}\hat{\psi}(\boldsymbol{k})$ by $\hat{\psi}^{\mathrm{p2}}(\boldsymbol{k},l)$ and $\boldsymbol{L}$ by $\boldsymbol{L}_{\mathrm{p2}}$.

In an analogue manner we obtain the following approximations for the Fourier space parts of the fields and the negative field gradients.

$$\boldsymbol{E}^{\mathrm{F}}(j) \approx \boldsymbol{E}_{\mathrm{t}}^{\mathrm{F}}(j) := -\nabla_{\boldsymbol{x}_j} \sum_{(\boldsymbol{k},l)\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p2}}(\boldsymbol{k},l)\,[S_{\mathrm{c}}(\boldsymbol{k},l) + S_{\mathrm{d}}(\boldsymbol{k},l)]\ \mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}^{\top},l)\boldsymbol{L}_{\mathrm{p2}}^{-1}\boldsymbol{x}_j},$$

$$\boldsymbol{G}^{\mathrm{F}}(j) \approx \boldsymbol{G}_{\mathrm{t}}^{\mathrm{F}}(j) := -\nabla_{\boldsymbol{x}_j}\nabla_{\boldsymbol{x}_j}^{\top} \sum_{(\boldsymbol{k},l)\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p2}}(\boldsymbol{k},l)\,[S_{\mathrm{c}}(\boldsymbol{k},l) + S_{\mathrm{d}}(\boldsymbol{k},l)]\ \mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}^{\top},l)\boldsymbol{L}_{\mathrm{p2}}^{-1}\boldsymbol{x}_j}.$$

Since the structure is the same as for 3d-periodic boundary conditions, we can proceed analogously, i.e., the sums $S_{\mathrm{c}}(\boldsymbol{k},l)$ are computed via the adjoint NFFT in three dimensions and $S_{\mathrm{d}}(\boldsymbol{k},l)$ are obtained by applying the adjoint gradient NFFT. In order to approximate $\phi^{\mathrm{F}}(j)$ the outer sums are evaluated via the NFFT, i.e.,

$$\phi_{\mathrm{t}}^{\mathrm{F}}(j) \approx \left(\sum_{(\boldsymbol{k},l)\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p2}}(\boldsymbol{k},l)\left[S_{\mathrm{c},\approx}(\boldsymbol{k},l) + S_{\mathrm{d},\approx}(\boldsymbol{k},l)\right]\ \mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}^{\top},l)\boldsymbol{L}_{\mathrm{p2}}^{-1}\boldsymbol{x}_j} \right)_{\approx} =: \phi_{\mathrm{p2nfft}}^{\mathrm{F}}(j). \tag{4.27}$$

For $\boldsymbol{E}^{\mathrm{F}}(j)$ and $\boldsymbol{G}^{\mathrm{F}}(j)$ we use the gradient NFFT and the Hessian NFFT, respectively.

$$\boldsymbol{E}_{\mathrm{t}}^{\mathrm{F}}(j) \approx -\left(\nabla_{\boldsymbol{x}_j} \sum_{(\boldsymbol{k},l)\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p2}}(\boldsymbol{k},l)\left[S_{\mathrm{c},\approx}(\boldsymbol{k},l) + S_{\mathrm{d},\approx}(\boldsymbol{k},l)\right]\ \mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}^{\top},l)\boldsymbol{L}_{\mathrm{p2}}^{-1}\boldsymbol{x}_j} \right)_{\approx} =: \boldsymbol{E}_{\mathrm{p2nfft}}^{\mathrm{F}}(j), \tag{4.28}$$

$$\boldsymbol{G}_{\mathrm{t}}^{\mathrm{F}}(j) \approx -\left(\nabla_{\boldsymbol{x}_j}\nabla_{\boldsymbol{x}_j}^{\top} \sum_{(\boldsymbol{k},l)\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p2}}(\boldsymbol{k},l)\left[S_{\mathrm{c},\approx}(\boldsymbol{k},l) + S_{\mathrm{d},\approx}(\boldsymbol{k},l)\right]\ \mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}^{\top},l)\boldsymbol{L}_{\mathrm{p2}}^{-1}\boldsymbol{x}_j} \right)_{\approx} =: \boldsymbol{G}_{\mathrm{p2nfft}}^{\mathrm{F}}(j). \tag{4.29}$$

We refer to Algorithm 4.3 for a summary of the P²NFFT method, where 2d-periodic boundary conditions are obtained for $\rho = 2$.

1d-periodic

The computation of the near field terms $\phi^{\mathrm{short}}(j)$, $\boldsymbol{E}^{\mathrm{short}}(j)$ and $\boldsymbol{G}^{\mathrm{short}}(j)$ is done analogously to the 3d-periodic case, see (4.1), (4.6) and (4.14), respectively. For this purpose, the lattice $\mathcal{S} = \mathbb{Z}^3$ is replaced by $\mathcal{S} = \mathbb{Z}\times\{0\}^2$. In addition, the self interaction terms $\phi^{\mathrm{self}}(j)$, $\boldsymbol{E}^{\mathrm{self}}(j)$ and $\boldsymbol{G}^{\mathrm{self}}(j)$ are the same as in the 3d-periodic case, see (2.24), (4.8) and (4.15), respectively, and can be computed easily.

The Fourier space parts of the potentials are given by

$$\phi^{\mathrm{F}}(j) = \frac{1}{|l_{11}|} \sum_{k \in \mathbb{Z}} \sum_{i=1}^{N} \xi_i \, \mathrm{e}^{2\pi \mathrm{i} k x_{ij,1}/l_{11}} \Theta^{\mathrm{p1}} \left(k/l_{11}, \|\tilde{\boldsymbol{x}}_{ij}\| \right), \tag{4.30}$$

where the function Θ^{p1} is defined in (2.53). Here, we assume periodic boundary conditions with respect to $\boldsymbol{\ell}_1$ and that this vector has the form

$$\boldsymbol{\ell}_1 = \begin{pmatrix} l_{11} \\ 0 \\ 0 \end{pmatrix}.$$

Note that $\Theta^{\mathrm{p1}}(\kappa, \cdot)$ is a symmetric C^∞ function for each $\kappa \geq 0$. For $\kappa > 0$ the functions $\Theta^{\mathrm{p1}}(\kappa, \cdot)$ have similar properties as the functions $\Theta^{\mathrm{p2}}(k, \cdot)$, which can be seen based on the very similar integral representations, see (2.41) and (2.58). In the special case $\kappa = 0$ we have $\Theta^{\mathrm{p1}}(0, r) \to -\infty$ for $|r| \to \infty$. The function is continuous in $r = 0$ with

$$\Theta^{\mathrm{p1}}(0, 0) := \lim_{r \to 0} \Theta^{\mathrm{p1}}(0, r) = \lim_{r \to 0} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n! n} (\alpha r)^{2n} = 0,$$

which is obtained via (2.49). We also easily see that $\Theta^{\mathrm{p1}}(0, \cdot) \in \mathrm{C}^\infty(\mathbb{R})$. An example is shown in Figure 4.2.

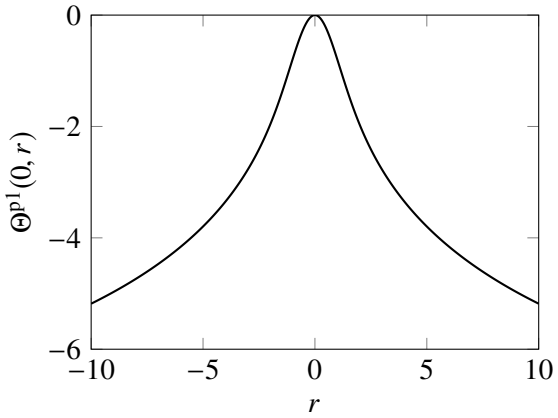


Figure 4.2: We plot the function $\Theta^{\mathrm{p1}}(0, \cdot)$ for the splitting parameter $\alpha := 1$.

It is reasonable to assume that the norms of the particle coordinates with respect to the nonperiodic dimensions are bounded, i.e., we have $\|\tilde{\boldsymbol{x}}_j\| \leq {}^{L}/_{2}$, $j = 1, \ldots, \mathbb{N}$, where $L > 0$. Thus, we have $\|\tilde{\boldsymbol{x}}_{ij}\| \leq L$ for all $i, j = 1, \ldots, N$.

We recall that the function values $\Theta^{\mathrm{p1}}(\kappa, r)$ tend to zero exponentially fast for growing κ, see Remark 2.13. Thus, we can truncate the Fourier series (4.30), i.e., we only include $k \in \mathcal{I}_{M_1}$ with $M_1 \in 2\mathbb{N}$ large enough, choose again $h > 2L$ and approximate each function $\Theta^{\mathrm{p1}}(\kappa, \cdot)$, $\kappa \in \{{}^{k}/_{l_{11}}, \kappa \in \mathcal{I}_{M_1}\}$, by a bivariate trigonometric polynomial. This ansatz is closely related to the NFFT based fast summation approach for radial kernels, see Section 3.4.2. Similarly to the 2d-periodic case we apply both, the regularization as well as the periodization approach, now in $d = 2$ dimensions, cf. Section 3.4.2.

Regularization: As already described for 2d-periodic boundary conditions, the regularization approach is applied for $\kappa = 0$ and relatively small values $\kappa > 0$. In the 1d-periodic setting we have to apply the regularization for radial kernels, see Section 3.4.2, with respect to the two nonperiodic dimensions. The regularized kernels $K_{\mathrm{R}}(\kappa, \cdot) : [-h/2, h/2]^2 \to \mathbb{R}$ are defined via

$$K_{\mathrm{R}}(\kappa, \boldsymbol{x}) := \begin{cases} \Theta^{\mathrm{p1}}(\kappa, \|\boldsymbol{x}\|) & : \|\boldsymbol{x}\| \leq L, \\ K_{\mathrm{B}}(\kappa, \|\boldsymbol{x}\|) & : L < \|\boldsymbol{x}\| \leq h/2, \\ K_{\mathrm{B}}(\kappa, h/2) & : \boldsymbol{x} \in [-h/2, h/2]^2, \|\boldsymbol{x}\| > h/2, \end{cases} \tag{4.31}$$

where the functions $K_{\mathrm{B}}(\kappa, \cdot)$ fulfill the $2p - 1$ interpolation conditions (3.71), (3.72).

Consequently, the periodic functions

$$\tilde{K}_{\mathrm{R}}(\kappa, \boldsymbol{x}) := K_{\mathrm{R}}(\kappa, \boldsymbol{x} + h\boldsymbol{n}) \quad \text{with } \boldsymbol{n} \in \mathbb{Z}^2 : \boldsymbol{x} + h\boldsymbol{n} \in [-h/2, h/2)^2,$$

are smooth on the two dimensional torus $h\mathbb{T}^2$, or more precisely $\tilde{K}_{\mathrm{R}}(\kappa, \cdot) \in \mathrm{C}^{p-1}(h\mathbb{T}^2)$, see Figure 3.14. Thus, we may approximate the functions by trigonometric polynomials, i.e.,

$$\tilde{K}_{\mathrm{R}}(\kappa, \boldsymbol{x}) \approx \sum_{\boldsymbol{l} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}} \hat{b}_{\kappa, \boldsymbol{l}} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l}^{\top} \boldsymbol{x}/h},$$

where $\tilde{\boldsymbol{M}} = (M_2, M_3) \in 2\mathbb{N}^2$ is chosen large enough. The Fourier coefficients $\hat{b}_{\kappa, \boldsymbol{l}}$ are simply approximated via the two dimensional FFT by using equispaced samples of $\tilde{K}_{\mathrm{R}}(\kappa, \cdot)$ in $[-h/2, h/2)^2$, i.e.,

$$\hat{b}_{\kappa, \boldsymbol{l}} := \frac{1}{M_2 M_3} \sum_{\boldsymbol{j} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}} \tilde{K}_{\mathrm{R}}(\kappa, (h\boldsymbol{j}) \odot \tilde{\boldsymbol{M}}^{-1}) \, \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{j}^{\top} (\boldsymbol{l} \odot \tilde{\boldsymbol{M}}^{-1})}. \tag{4.32}$$

Since we have $\|\tilde{\boldsymbol{x}}_{ij}\| \leq L$, we obtain

$$\Theta^{\mathrm{p1}}(k/l_{11}, \|\tilde{\boldsymbol{x}}_{ij}\|) \approx \sum_{\boldsymbol{l} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}} \hat{b}_{k/l_{11}, \boldsymbol{l}} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l}^{\top} \tilde{\boldsymbol{x}}_{ij}/h}$$

for all considered values of $\kappa = k/l_{11}$.

In order to construct the regularized kernels $K_{\mathrm{R}}(\kappa, \cdot)$ we need to compute the partial derivatives of $\Theta^{\mathrm{p1}}(\kappa, r) = K_0(\pi^2\kappa^2/\alpha^2, \alpha^2 r^2)$ with respect to r. Note that the relation [Har08]

$$\frac{\partial K_\nu(x, y)}{\partial y} = -K_{\nu+1}(x, y) \tag{4.33}$$

gives an easy way to compute the required derivatives. The identity can be verified easily by using the integral representation (2.2.3). The evaluation of the partial derivatives is possible based on the recursion as presented in the following lemma, see also [NPP15a, Lemma 5.6].

Lemma 4.7. *For $x \in \mathbb{R}$ and $k \in \mathbb{N}_0$ we have*

$$\frac{\partial^{2k}}{\partial y^{2k}} K_0(x, y^2) = \sum_{l=0}^{k} \alpha_{k,l} K_{k+l}(x, y^2) y^{2l}$$

$$\frac{\partial^{2k+1}}{\partial y^{2k+1}} K_0(x, y^2) = \sum_{l=0}^{k} \beta_{k,l} K_{k+1+l}(x, y^2) y^{2l+1},$$

where the coefficients $\alpha_{k,l}$ and $\beta_{k,l}$ are given recursively by

$$\alpha_{0,0} = 1,$$
$$\beta_{k,l} = \begin{cases} -2\alpha_{k,k} & : l = k, \\ -2\alpha_{k,l} + 2(l+1)\alpha_{k,l+1} & : else, \end{cases}$$
$$\alpha_{k,l} = \begin{cases} \beta_{k-1,0} & : k \geq 1, l = 0, \\ -2\beta_{k-1,l-1} & : k \geq 1, l = k, \\ -2\beta_{k-1,l-1} + (2l+1)\beta_{k-1,l} & : k \geq 1, else. \end{cases}$$

Proof. The result is obtained via induction in k, where we have to apply the chain rule as well as relation (4.33). ∎

In the case $\kappa = 0$ we have to compute the derivatives of $\ln(\alpha^2 r^2)$ as well as $\Gamma(0, \alpha^2 r^2)$. Since $\ln(\alpha^2 r^2) = 2\ln(\alpha) + 2\ln(r)$ the computation of $\frac{\mathrm{d}^n}{\mathrm{d}r^n} \ln(\alpha^2 r^2)$ is straight forward. The derivatives

$$\frac{\mathrm{d}^n}{\mathrm{d}r^n} \Gamma(0, \alpha^2 r^2) = \alpha^n \frac{\mathrm{d}^n}{\mathrm{d}x^n} \Gamma(0, x^2)\Big|_{x=\alpha r}$$

are obtained based on the following result, cf. [NPP15a, Lemma 5.7].

Lemma 4.8. *For $k \in \mathbb{N}_0$ we have*

$$\frac{\mathrm{d}^{2k}}{\mathrm{d}x^{2k}} \Gamma(0, x^2) = 2\,\mathrm{e}^{-x^2} \left(\sum_{l=1}^{k} \frac{(2k-1)!}{(k-l)!} x^{-2l} + p_k(x) \right),$$
$$\frac{\mathrm{d}^{2k+1}}{\mathrm{d}x^{2k+1}} \Gamma(0, x^2) = -2\,\mathrm{e}^{-x^2} \left(\sum_{l=0}^{k} \frac{(2k)!}{(k-l)!} x^{-(2l+1)} + q_k(x) \right),$$

where p_k and q_k are polynomials of degree $2k-2$ and $2k-1$, respectively, fulfilling the recursion

$$q_0(x) = 0$$
$$p_1(x) = 2$$
$$q_k(x) = 2x p_k(x) - \frac{\mathrm{d}}{\mathrm{d}x} p_k(x)$$
$$p_{k+1}(x) = 2x q_k(x) - \frac{\mathrm{d}}{\mathrm{d}x} q_k(x) + \frac{2(2k)!}{k!}.$$

Proof. Via the definition of the incomplete Gamma function, see (2.48), we get

$$\frac{\mathrm{d}}{\mathrm{d}x} \Gamma(0, x) = -\frac{\mathrm{e}^{-x}}{x} \quad \text{and thus} \quad \frac{\mathrm{d}}{\mathrm{d}x} \Gamma(0, x^2) = -\frac{2\,\mathrm{e}^{-x^2}}{x}.$$

Based on that, the result is obtained via induction in $k \in \mathbb{N}$. ∎

Periodization: For $\kappa \neq 0$ we consider the error

$$\sum_{\boldsymbol{n} \in \mathbb{Z}^2} \Theta^{\mathrm{p1}}(\kappa, \|\boldsymbol{r} + h\boldsymbol{n}\|) - \Theta^{\mathrm{p1}}(\kappa, \|\boldsymbol{r}\|) = \sum_{\boldsymbol{n} \in \mathbb{Z}^2 \setminus \{\boldsymbol{0}\}} \Theta^{\mathrm{p1}}(\kappa, \|\boldsymbol{r} + h\boldsymbol{n}\|).$$

Since $\Theta^{\mathrm{p1}}(\kappa, x)$ is monotonically decreasing for $|x| \to \infty$ and $h > 2L$, we have

$$\Theta^{\mathrm{p1}}(\kappa, \|\boldsymbol{r} + h\boldsymbol{n}\|) \leq \Theta^{\mathrm{p1}}(\kappa, h - L) \quad \forall\, \|\boldsymbol{r}\| \leq L, \boldsymbol{n} \in \mathbb{Z}^2 \setminus \{\boldsymbol{0}\}.$$

Thus, if $\Theta^{\mathrm{p1}}(\kappa, h - L)$ is sufficiently small, we may consider the periodized function instead and use the analytic Fourier coefficients based on the Poisson summation formula (2.25). For this purpose, we have to compute the Fourier transform

$$\int_{\mathbb{R}^2} \Theta^{\mathrm{p1}}(\kappa, \|\boldsymbol{r}\|)\, \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{r}^\top \boldsymbol{\xi}}\, \mathrm{d}\boldsymbol{r} =: \hat{\Theta}^{\mathrm{p1}}(\kappa, \|\boldsymbol{\xi}\|). \tag{4.34}$$

Note that we exploit the well known fact that the Fourier transform of a radial function is radial itself.

Lemma 4.9. *For $\kappa \neq 0$ we have*

$$\hat{\Theta}^{\mathrm{p1}}(\kappa, \|\boldsymbol{\xi}\|) = \frac{1}{\pi(\kappa^2 + \|\boldsymbol{\xi}\|^2)}\, \mathrm{e}^{-\pi^2\kappa^2/\alpha^2 - \pi^2\|\boldsymbol{\xi}\|^2/\alpha^2}.$$

Proof. We insert the integral representation of the function Θ^{p1}, see (2.58), and get the following representation of the integral in (4.34).

$$\int_{\mathbb{R}^2} \Theta^{\mathrm{p1}}(\kappa, \|\boldsymbol{r}\|)\, \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{r}^\top \boldsymbol{\xi}}\, \mathrm{d}\boldsymbol{r} = \int_1^\infty \frac{\mathrm{e}^{-\frac{\pi^2\kappa^2 t}{\alpha^2}}}{t} \int_{\mathbb{R}^2} \mathrm{e}^{-\frac{\alpha^2\|\boldsymbol{r}\|^2}{t}}\, \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{r}^\top \boldsymbol{\xi}}\, \mathrm{d}\boldsymbol{r}\, \mathrm{d}t.$$

Now, we just have to compute the Fourier integral of a Gaussian, cf. (2.27), to obtain

$$\hat{\Theta}^{\mathrm{p1}}(\kappa, \|\boldsymbol{\xi}\|) = \int_1^\infty \frac{\mathrm{e}^{-\frac{\pi^2\kappa^2 t}{\alpha^2}}}{t} \cdot \frac{\pi t}{\alpha^2}\, \mathrm{e}^{-\frac{\pi^2\|\boldsymbol{\xi}\|^2 t}{\alpha^2}}\, \mathrm{d}t = \frac{1}{\pi(\kappa^2 + \|\boldsymbol{\xi}\|^2)}\, \mathrm{e}^{-\pi^2\kappa^2/\alpha^2 - \pi^2\|\boldsymbol{\xi}\|^2/\alpha^2}.$$

■

Following the periodization approach, we obtain an approximation of the function $\Theta^{\mathrm{p1}}(\kappa, \|\cdot\|)$ via the Poisson summation formula (2.26), where $d = 2$. We end up with an approximation of the form

$$\Theta^{\mathrm{p1}}(\kappa, \|\boldsymbol{r}\|) \approx \frac{1}{h^2} \sum_{\boldsymbol{l} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}} \hat{\Theta}^{\mathrm{p1}}(\kappa, h^{-1}\|\boldsymbol{l}\|)\, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l}^\top \boldsymbol{r}/h},$$

where we have to choose $\mathcal{I}_{\tilde{\boldsymbol{M}}} \in 2\mathbb{N}^2$ large enough in order to achieve a good approximation.

Similarly to the 2d-periodic case, cf. Remark 4.6, we remark that the periodization approach is applied in the spectrally accurate Ewald method for singly periodic domains [ST17]. In contrast, the method presented in [MMY+04] relies on the truncation approach, cf. Variant II in Section 3.4.1, where the decrease of the Fourier coefficients is rather slow.

Summary: We summarize the precomputations that are necessary for the P²NFFT method under 1d-periodic boundary conditions in the following algorithm.

Algorithm 4.2 (Precomputations 1d-periodic P^{2}NFFT).
Input: Splitting parameter α, far field cutoff $M_1 \in 2\mathbb{N}$, mesh size $\tilde{\boldsymbol{M}} = (M_2, M_3)^\top \in 2\mathbb{N}^2$, maximum distance $L > 0$, period $h \geq 2L$, periodic direction $\boldsymbol{\ell}_1 = (l_{11}, 0, 0)^\top$.

For all $k = -{}^{M_1}\!/_2, \ldots, {}^{M_1}\!/_2 - 1$:

i) Set $\kappa := l_{11}^{-1} k$.

ii) If $\kappa = 0$ or $\Theta^{\mathrm{p1}}(\kappa, h - L) \geq \varepsilon$:

a) Construct the regularized kernel $K_{\mathrm{R}}(\kappa, \cdot)$, see (4.31).

b) Compute the coefficients $\hat{b}_{\kappa,\boldsymbol{l}}$, $\boldsymbol{l} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}$, via the inverse FFT, see (4.32).

iii) Else, i.e., $\Theta^{\mathrm{p1}}(\kappa, h - L) < \varepsilon$: Set

$$\hat{b}_{\kappa,\boldsymbol{l}} := \frac{\mathrm{e}^{-\pi^2\kappa^2/\alpha^2 - \pi^2 \|\boldsymbol{l}^2\|/(\alpha^2 h^2)}}{h^2 \pi (\kappa^2 + \|\boldsymbol{l}\|^2/h^2)}$$

for all $\boldsymbol{l} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}$.

iv) Set $\hat{\psi}^{\mathrm{p1}}(k, \boldsymbol{l}) := \hat{b}_{\kappa,\boldsymbol{l}}$, $\boldsymbol{l} \in \mathcal{I}_{\tilde{\boldsymbol{M}}}$.

Output: Fourier coefficients $\hat{\psi}^{\mathrm{p1}}(k, \boldsymbol{l})$ for $(k, \boldsymbol{l}) \in \mathcal{I}_{\boldsymbol{M}}$, $\boldsymbol{M} := (M_1, M_2, M_3)^\top$. □

In the current implementation of the algorithm, as summarized in Section 4.2, we apply $\varepsilon := 10^{-16}$ within the second step of the above described precomputation procedure.

With the precomputed Fourier coefficients $\hat{\psi}^{\mathrm{p1}}(k, \boldsymbol{l})$ and the matrix

$$\boldsymbol{L}_{\mathrm{p1}} := \begin{pmatrix} l_{11} & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & h \end{pmatrix}$$

we can write

$$\phi^{\mathrm{F}}(j) \approx \phi_{\mathrm{t}}^{\mathrm{F}}(j) := \sum_{(k,\boldsymbol{l}) \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p1}}(k, \boldsymbol{l}) \left[S_{\mathrm{c}}(k, \boldsymbol{l}) + S_{\mathrm{d}}(k, \boldsymbol{l}) \right] \mathrm{e}^{-2\pi \mathrm{i} (k, \boldsymbol{l}^\top) \boldsymbol{L}_{\mathrm{p1}}^{-1} \boldsymbol{x}_j}, \tag{4.35}$$

where we define the structure factors

$$S_{\mathrm{c}}(k, \boldsymbol{l}) := \sum_{i=1}^{N_{\mathrm{c}}} q_i \, \mathrm{e}^{2\pi \mathrm{i} (k, \boldsymbol{l}^\top) \boldsymbol{L}_{\mathrm{p1}}^{-1} \boldsymbol{x}_i}, \tag{4.36}$$

$$S_{\mathrm{d}}(k, \boldsymbol{l}) := \sum_{i=N_{\mathrm{c}}+1}^{N} \boldsymbol{\mu}_i^\top \nabla_{\boldsymbol{x}_i} \, \mathrm{e}^{2\pi \mathrm{i} (k, \boldsymbol{l}^\top) \boldsymbol{L}_{\mathrm{p1}}^{-1} \boldsymbol{x}_i}. \tag{4.37}$$

Note that the approximation (4.35) has the same structure as the one obtained for 3d-periodic boundary conditions, see (4.4). We just formally replace $(\pi V)^{-1} \hat{\psi}(\boldsymbol{k})$ by $\hat{\psi}^{\mathrm{p1}}(k, \boldsymbol{l})$ and $\boldsymbol{L}$ by $\boldsymbol{L}_{\mathrm{p1}}$.

In an analogue manner we obtain the following approximations for the Fourier space parts of the fields and the negative field gradients.

$$\boldsymbol{E}^{\mathrm{F}}(j) \approx \boldsymbol{E}_{\mathrm{t}}^{\mathrm{F}}(j) := -\nabla_{\boldsymbol{x}_j} \sum_{(k,\boldsymbol{l}) \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p1}}(k, \boldsymbol{l}) \left[S_{\mathrm{c}}(k, \boldsymbol{l}) + S_{\mathrm{d}}(k, \boldsymbol{l}) \right] \mathrm{e}^{-2\pi \mathrm{i} (k, \boldsymbol{l}^\top) \boldsymbol{L}_{\mathrm{p1}}^{-1} \boldsymbol{x}_j},$$

$$\boldsymbol{G}^{\mathrm{F}}(j) \approx \boldsymbol{G}_{\mathrm{t}}^{\mathrm{F}}(j) := -\nabla_{\boldsymbol{x}_j} \nabla_{\boldsymbol{x}_j}^\top \sum_{(k,\boldsymbol{l}) \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p1}}(k, \boldsymbol{l}) \left[S_{\mathrm{c}}(k, \boldsymbol{l}) + S_{\mathrm{d}}(k, \boldsymbol{l}) \right] \mathrm{e}^{-2\pi \mathrm{i} (k, \boldsymbol{l}^\top) \boldsymbol{L}_{\mathrm{p1}}^{-1} \boldsymbol{x}_j}.$$

Since the structure is the same as for 3d-periodic boundary conditions, we can proceed analogously, i.e., the sums $S_c(k, \boldsymbol{l})$ are computed via the adjoint NFFT in three dimensions and $S_d(k, \boldsymbol{l})$ are obtained by applying the adjoint gradient NFFT. In order to approximate $\phi^F(j)$ the outer sums are evaluated via the NFFT.

$$\phi_t^F(j) \approx \left(\sum_{(k,\boldsymbol{l})\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{p1}(k,\boldsymbol{l}) \left[S_{c,\approx}(k,\boldsymbol{l}) + S_{d,\approx}(k,\boldsymbol{l}) \right] \mathrm{e}^{-2\pi \mathrm{i}(k,\boldsymbol{l}^\top)\boldsymbol{L}_{p1}^{-1}\boldsymbol{x}_j} \right)_{\approx} =: \phi_{p2nfft}^F(j). \tag{4.38}$$

For $\boldsymbol{E}^F(j)$ and $\boldsymbol{G}^F(j)$ we use the gradient NFFT and the Hessian NFFT, respectively.

$$\boldsymbol{E}_t^F(j) \approx -\left(\nabla_{\boldsymbol{x}_j} \sum_{(k,\boldsymbol{l})\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{p1}(k,\boldsymbol{l}) \left[S_{c,\approx}(k,\boldsymbol{l}) + S_{d,\approx}(k,\boldsymbol{l}) \right] \mathrm{e}^{-2\pi \mathrm{i}(k,\boldsymbol{l}^\top)\boldsymbol{L}_{p1}^{-1}\boldsymbol{x}_j} \right)_{\approx} =: \boldsymbol{E}_{p2nfft}^F(j), \tag{4.39}$$

$$\boldsymbol{G}_t^F(j) \approx -\left(\nabla_{\boldsymbol{x}_j} \nabla_{\boldsymbol{x}_j}^\top \sum_{(k,\boldsymbol{l})\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{p1}(k,\boldsymbol{l}) \left[S_{c,\approx}(k,\boldsymbol{l}) + S_{d,\approx}(k,\boldsymbol{l}) \right] \mathrm{e}^{-2\pi \mathrm{i}(k,\boldsymbol{l}^\top)\boldsymbol{L}_{p1}^{-1}\boldsymbol{x}_j} \right)_{\approx} =: \boldsymbol{G}_{p2nfft}^F(j). \tag{4.40}$$

Remark 4.10 (Standard basis vectors)**.** Based on Lemma 2.14 we may, without loss of generality, assume that the matrix $\boldsymbol{L}$ has the form

$$\boldsymbol{L} = \begin{pmatrix} l_{11} & 0 & 0 \\ 0 & l_{22} & l_{32} \\ 0 & l_{23} & l_{33} \end{pmatrix},$$

see (2.61). Since the box vectors regarding the nonperiodic dimensions are not relevant for the Ewald summation formulas, see Lemma 2.12, we may replace them by other adequate vectors.

If the particle system is rotated appropriately, we may simply replace the box vectors $\boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$ by standard basis vectors

$$\boldsymbol{\ell}_2 = \begin{pmatrix} 0 \\ L_2 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\ell}_3 = \begin{pmatrix} 0 \\ 0 \\ L_3 \end{pmatrix}, \tag{4.41}$$

see Figure 4.3. Hereby, the volume of the primary simulation box might change but the diameter L does not, if done correctly. Note that only the diameter L is of importance for the above described fast summation approach.

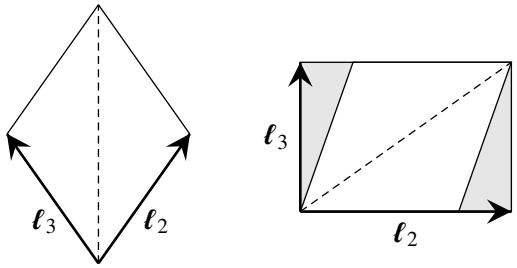


Figure 4.3: In the 1d-periodic case the particle system may be rotated appropriately regarding the nonperiodic dimensions in order to replace the box vectors $\boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$ by standard basis vectors, without changing the diameter of the limiting box.

As an example, in case of the box vectors as given in Example 2.15, we may further replace the vectors

$$\tilde{\boldsymbol{\ell}}_2 = \begin{pmatrix} 0 \\ 5\sqrt{2} \\ 10 \end{pmatrix} \text{ and } \tilde{\boldsymbol{\ell}}_3 = \begin{pmatrix} 0 \\ -5\sqrt{2} \\ 10 \end{pmatrix}$$

by the two orthogonal vectors

$$\tilde{\boldsymbol{\ell}}_2 = \begin{pmatrix} 0 \\ \frac{20\sqrt{2}}{3} \\ \frac{40}{3} \end{pmatrix} \text{ and } \tilde{\boldsymbol{\ell}}_3 = \begin{pmatrix} 0 \\ \frac{-20\sqrt{2}}{3} \\ \frac{20}{3} \end{pmatrix},$$

as illustrated in Figure 4.4. By rotating the particle system appropriately, we finally end up with a primary simulation box spanned by $\boldsymbol{\ell}_1 = (10\sqrt{2}, 0, 0)^\top$, $\boldsymbol{\ell}_2 = (0, {}^{20\sqrt{2}}/_{\sqrt{3}}, 0)^\top$ and $\boldsymbol{\ell}_3 = (0, 0, {}^{20}/_{\sqrt{3}})^\top$.

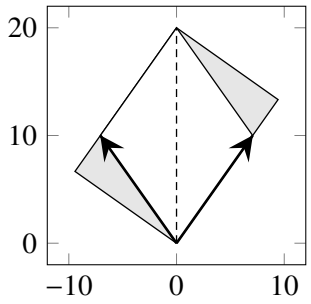


Figure 4.4: In case of Example 2.15 we may replace the two non orthogonal vectors $\tilde{\boldsymbol{\ell}}_2$ and $\tilde{\boldsymbol{\ell}}_3$ by two orthogonal vectors in order to fulfill the requirements of Algorithm 4.3.

In other words, in the 1d-periodic case we can break down the box structure to the case of an orthorhombic box, i.e., $\boldsymbol{L} = \operatorname{diag}(L_1, L_2, L_3)$. Consequently, we obtain $\|\tilde{\boldsymbol{x}}_{ij}\| \leq L$ with

$$L := \sqrt{L_2^2 + L_3^2}.$$

□

A summary of the resulting fast method is given in Algorithm 4.3, where 1d-periodic boundary conditions are obtained for $\rho = 1$.

0d-periodic (open) boundary conditions

For open boundary conditions we have

$$\phi(j) = \phi^{\text{short}}(j) + \phi^{\text{long}}(j) - \phi^{\text{self}}(j),$$

where the single parts are defined in (2.64), (2.65) and (2.24), respectively. We remark that charge neutrality is not required, since all considered sums are finite.

The computation of the self interaction terms $\phi^{\text{self}}(j)$, $\boldsymbol{E}^{\text{self}}(j)$ and $\boldsymbol{G}^{\text{self}}(j)$ is straight forward, see (2.24), (4.8) and (4.15), respectively. Furthermore, the computation of the short ranged interactions $\phi^{\text{short}}(j)$, $\boldsymbol{E}^{\text{short}}(j)$ and $\boldsymbol{G}^{\text{short}}(j)$ can be realized efficiently based on a direct computation, analogously to the previously considered types of periodic boundary conditions.

In order to compute the long range parts efficiently, we apply the NFFT based fast summation approach for radial kernels, see Section 3.4.2. Therefor, we assume that the norms of the particle positions are bounded, i.e., $\|\boldsymbol{x}_j\| \leq {}^{L}/_{2}$, which implies $\|\boldsymbol{x}_{ij}\| \leq L$, and choose $h \geq 2L$.

We define

$$\Theta^{\mathrm{p0}}(r) := \begin{cases} \dfrac{2\alpha}{\sqrt{\pi}} & : r = 0, \\ \dfrac{\operatorname{erf}(\alpha r)}{r} & : \text{else}, \end{cases}$$

and construct the regularized kernel $K_{\mathrm{R}} : [-{}^{h}/_{2}, {}^{h}/_{2}]^3 \to \mathbb{R}$ via

$$K_{\mathrm{R}}(\boldsymbol{x}) := \begin{cases} \Theta^{\mathrm{p0}}(\|\boldsymbol{x}\|) & : \|\boldsymbol{x}\| \leq L, \\ K_{\mathrm{B}}(\|\boldsymbol{x}\|) & : L < \|\boldsymbol{x}\| \leq {}^{h}/_{2}, \\ K_{\mathrm{B}}({}^{h}/_{2}) & : \boldsymbol{x} \in [-{}^{h}/_{2}, {}^{h}/_{2}]^3, \|\boldsymbol{x}\| > {}^{h}/_{2}, \end{cases} \tag{4.42}$$

where the polynomial K_{B} fulfills the $2p-1$ interpolation conditions (3.71), (3.72).

Thus, the periodic function

$$\tilde{K}_{\mathrm{R}}(\boldsymbol{x}) := K_{\mathrm{R}}(\boldsymbol{x} + h\boldsymbol{n}) \quad \text{with } \boldsymbol{n} \in \mathbb{Z}^3 : \boldsymbol{x} + h\boldsymbol{n} \in [-{}^{h}/_{2}, {}^{h}/_{2})^3,$$

is smooth on the three dimensional torus $h\mathbb{T}^3$, or more precisely, $K_{\mathrm{R}} \in \mathrm{C}^{p-1}(h\mathbb{T}^3)$. Consequently, the function can be approximated by a trigonometric polynomial

$$\tilde{K}_{\mathrm{R}}(\boldsymbol{x}) \approx \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p0}}(\boldsymbol{l})\, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l}^{\top} \boldsymbol{x}/h},$$

where $\boldsymbol{M} \in 2\mathbb{N}^3$ should be chosen large enough. The Fourier coefficients $\hat{\psi}^{\mathrm{p0}}(\boldsymbol{l})$ can be obtained by using the FFT via

$$\hat{\psi}^{\mathrm{p0}}(\boldsymbol{l}) := \frac{1}{|\mathcal{I}_{\boldsymbol{M}}|} \sum_{\boldsymbol{j} \in \mathcal{I}_{\boldsymbol{M}}} \tilde{K}_{\mathrm{R}}((h\boldsymbol{j}) \odot \boldsymbol{M}^{-1})\, \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{j}^{\top} (\boldsymbol{l} \odot \boldsymbol{M}^{-1})}. \tag{4.43}$$

Since we have $\|\boldsymbol{x}_{ij}\| \leq L$, we obtain

$$\Theta^{\mathrm{p0}}(\|\boldsymbol{x}_{ij}\|) \approx \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p0}}(\boldsymbol{l})\, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l}^{\top} \boldsymbol{x}_{ij}/h}.$$

The derivatives of the function Θ^{p0} can be expressed via

$$\frac{\mathrm{d}^n}{\mathrm{d}r^n} \Theta^{\mathrm{p0}}(r) = \sum_{k=0}^{n} \binom{n}{k} \frac{\mathrm{d}^k}{\mathrm{d}r^k} \operatorname{erf}(\alpha r) \frac{\mathrm{d}^{n-k}}{\mathrm{d}r^{n-k}} \frac{1}{r}, \tag{4.44}$$

where the computation of the derivatives $\frac{\mathrm{d}^{n-k}}{\mathrm{d}r^{n-k}} \frac{1}{r}$ is straight forward and the derivatives $\frac{\mathrm{d}^k}{\mathrm{d}r^k} \operatorname{erf}(\alpha r)$ can be obtained by considering the well known derivatives of the Gaussian function.

Now, we have

$$\phi^{\mathrm{F}}(j) \approx \phi^{\mathrm{F}}_{\mathrm{t}}(j) := \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p0}}(\boldsymbol{l}) \left[S_{\mathrm{c}}(\boldsymbol{l}) + S_{\mathrm{d}}(\boldsymbol{l}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{l}^{\top} \boldsymbol{L}_{\mathrm{p0}}^{-1} \boldsymbol{x}_j}, \tag{4.45}$$

where we define the matrix

$$\boldsymbol{L}_{\mathrm{p0}} := \begin{pmatrix} h & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & h \end{pmatrix}$$

an the structure factors

$$S_{\mathrm{c}}(\boldsymbol{l}) := \sum_{i=1}^{N_{\mathrm{c}}} q_i \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l}^{\top} \boldsymbol{L}_{\mathrm{p0}}^{-1} \boldsymbol{x}_i}, \tag{4.46}$$

$$S_{\mathrm{d}}(\boldsymbol{l}) := \sum_{i=N_{\mathrm{c}}+1}^{N} \boldsymbol{\mu}_i^{\top} \nabla_{\boldsymbol{x}_i} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{l}^{\top} \boldsymbol{L}_{\mathrm{p0}}^{-1} \boldsymbol{x}_i}. \tag{4.47}$$

Note that the approximation (4.45) has again the same structure as the one obtained for 3d-periodic boundary conditions, see (4.4). We just formally replace $(\pi V)^{-1} \hat{\psi}(\boldsymbol{k})$ by $\hat{\psi}^{\mathrm{p0}}(\boldsymbol{l})$ and $\boldsymbol{L}$ by $\boldsymbol{L}_{\mathrm{p0}}$.

In an analogue manner we obtain the following approximations for the Fourier space parts of the fields and the negative field gradients.

$$\boldsymbol{E}^{\mathrm{F}}(j) \approx \boldsymbol{E}_{\mathrm{t}}^{\mathrm{F}}(j) := -\nabla_{\boldsymbol{x}_j} \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p0}}(\boldsymbol{l}) \left[S_{\mathrm{c}}(\boldsymbol{l}) + S_{\mathrm{d}}(\boldsymbol{l}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{l}^{\top} \boldsymbol{L}_{\mathrm{p0}}^{-1} \boldsymbol{x}_j},$$

$$\boldsymbol{G}^{\mathrm{F}}(j) \approx \boldsymbol{G}_{\mathrm{t}}^{\mathrm{F}}(j) := -\nabla_{\boldsymbol{x}_j} \nabla_{\boldsymbol{x}_j}^{\top} \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p0}}(\boldsymbol{l}) \left[S_{\mathrm{c}}(\boldsymbol{l}) + S_{\mathrm{d}}(\boldsymbol{l}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{l}^{\top} \boldsymbol{L}_{\mathrm{p0}}^{-1} \boldsymbol{x}_j}.$$

Since the structure is the same as for 3d-periodic boundary conditions, we can proceed analogously, i.e., the sums $S_{\mathrm{c}}(\boldsymbol{l})$ are computed via the adjoint NFFT in three dimensions and $S_{\mathrm{d}}(\boldsymbol{l})$ are obtained by applying the adjoint gradient NFFT. In order to approximate $\phi^{\mathrm{F}}(j)$ the outer sums are evaluated via the NFFT.

$$\phi_{\mathrm{t}}^{\mathrm{F}}(j) \approx \left(\sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p0}}(\boldsymbol{l}) \left[S_{\mathrm{c},\approx}(\boldsymbol{l}) + S_{\mathrm{d},\approx}(\boldsymbol{l}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{l}^{\top} \boldsymbol{L}_{\mathrm{p0}}^{-1} \boldsymbol{x}_j} \right)_{\approx} =: \phi_{\mathrm{p2nfft}}^{\mathrm{F}}(j). \tag{4.48}$$

For $\boldsymbol{E}^{\mathrm{F}}(j)$ and $\boldsymbol{G}^{\mathrm{F}}(j)$ we use the gradient NFFT and the Hessian NFFT, respectively.

$$\boldsymbol{E}_{\mathrm{t}}^{\mathrm{F}}(j) \approx -\left(\nabla_{\boldsymbol{x}_j} \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p0}}(\boldsymbol{l}) \left[S_{\mathrm{c},\approx}(\boldsymbol{l}) + S_{\mathrm{d},\approx}(\boldsymbol{l}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{l}^{\top} \boldsymbol{L}_{\mathrm{p0}}^{-1} \boldsymbol{x}_j} \right)_{\approx} =: \boldsymbol{E}_{\mathrm{p2nfft}}^{\mathrm{F}}(j), \tag{4.49}$$

$$\boldsymbol{G}_{\mathrm{t}}^{\mathrm{F}}(j) \approx -\left(\nabla_{\boldsymbol{x}_j} \nabla_{\boldsymbol{x}_j}^{\top} \sum_{\boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}^{\mathrm{p0}}(\boldsymbol{l}) \left[S_{\mathrm{c},\approx}(\boldsymbol{l}) + S_{\mathrm{d},\approx}(\boldsymbol{l}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{l}^{\top} \boldsymbol{L}_{\mathrm{p0}}^{-1} \boldsymbol{x}_j} \right)_{\approx} =: \boldsymbol{G}_{\mathrm{p2nfft}}^{\mathrm{F}}(j). \tag{4.50}$$

Remark 4.11 (Standard basis vectors)**.** Suppose that the given particle positions $\boldsymbol{x}_j$ are located in a box spanned by three linearly independent box vectors $\boldsymbol{\ell}_1$, $\boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$. We remark that only the maximum distance L between the particles is relevant for the fast algorithm, as described above.

Thus, we may replace the box vectors by other appropriate vectors. If the particle system is rotated appropriately, we can simply replace the box vectors by standard basis vectors

$$\boldsymbol{\ell}_1 = \begin{pmatrix} L_1 \\ 0 \\ 0 \end{pmatrix}, \; \boldsymbol{\ell}_2 = \begin{pmatrix} 0 \\ L_2 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\ell}_3 = \begin{pmatrix} 0 \\ 0 \\ L_3 \end{pmatrix} \tag{4.51}$$

without changing the diameter L of the limiting box, see Figure 4.3. Then, we have $\boldsymbol{x}_{ij} \leq L$ with

$$L := \sqrt{L_1^2 + L_2^2 + L_3^2}.$$

□

The P^{2}NFFT method for open boundary conditions is summarized in Algorithm 4.3, where we have to set $\rho := 0$.

For open boundary conditions we may also apply the fast summation approach by regularizing the kernel function $K(r) = r^{-1}$ at the boundary of the interval as well as in a small region $[-\epsilon_\mathrm{I}, \epsilon_\mathrm{I}]$ around the singularity in $r = 0$, see [PS03, FS04, PSN04, PPvR06]. However, the application of the Ewald splitting (2.17) embeds the 0d-periodic case into the unified framework of P^{2}NFFT for 3d-periodic and mixed periodic boundary constraints. At the end of the day, our algorithm follows the same modular structure for all types of periodic boundary conditions, see Algorithm 4.3.

Whereas the above presented approach makes use of the regularization technique in $d = 3$ dimensions, the method introduced in [MT99] simply truncates the non-periodic kernel function, cf. Variant II in Section 3.4.1. Consequently, this method yields a rather low accuracy, since the approximated periodic function is only continuous.

4.2 Implementation

An implementation of the P^{2}NFFT method, as presented above, is part of the ScaFaCoS library [ABD$^+$], which is a massively parallel, open source software library for computing Coulomb interactions. It includes several other methods, such as P^{3}M [HE88, DH98a, DH98b], FMM [GR87, KD11, Kab12], and multigrid based solvers [TOS00, Bol08], for instance. A first implementation of P^{2}NFFT realized the efficient computation of charge-charge interactions subject to fully periodic boundary conditions, see [PP11, PP13, Pip15]. Later on, the generalization to mixed periodic and open boundary conditions was introduced and implemented, as presented in [NP13, NPP15a, NPP15b, Nes17].

Finally, the implementation of the versatile P^{2}NFFT framework for charged particle systems has been extended in order to compute also dipole-dipole [Nes16b] as well as charge-dipole interactions [HNP17], which are present in particle systems containing a mixture of charges and dipoles. Note that the treatment of pure charge as well as pure dipole systems is still included. Therefore, adding support for dipole-dipole and charge-dipole interactions makes the software applicable to an even wider range of possible applications.

In the following we present a brief summary of important details concerning the current implementation.

What ScaFaCoS computes

In order to commit the particle data to the ScaFaCoS test program, the data have to be listed in a specially formatted `xml` file. The ScaFaCoS interface had to be extended appropriately to allow for the presence of dipole particles. To this end, each particle is specified by the following attributes.

- Either `type = "charge"` or `type = "dipole"`
- and correspondingly `charge = "q"` or `moment = "mu1 mu2 mu3"`,
- and `position = "x1 x2 x3"`.

Furthermore, the file should contain the three box vectors $\boldsymbol{\ell}_1$, $\boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$ as well as the periodicity $\boldsymbol{p} \in \{0, 1\}^3$, where $p(j) = 1$ if periodic boundary conditions are required with respect to the dimension j and $p(j) = 0$ otherwise.

Possible structures of the regular matrix $\boldsymbol{L} = [\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3]$ are summarized in Figure 4.5.

$$\boldsymbol{L} = \underset{\substack{\text{3d-periodic}\\(\rho = 3)}}{\begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix}} \quad \underset{\substack{\text{2d-periodic}\\(\rho = 2)}}{\begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & L_3 \end{pmatrix}} \quad \underset{\substack{\text{1d-periodic}\\(\rho = 1)}}{\begin{pmatrix} L_1 & 0 & 0 \\ 0 & L_2 & 0 \\ 0 & 0 & L_3 \end{pmatrix}} \quad \underset{\substack{\text{0d-periodic}\\(\rho = 0)}}{\begin{pmatrix} L_1 & 0 & 0 \\ 0 & L_2 & 0 \\ 0 & 0 & L_3 \end{pmatrix}}$$

Figure 4.5: Possible structures of the matrix $\boldsymbol{L} \in \mathbb{R}^{3\times 3}$ for the P^{2}NFFT method in ScaFaCoS. The columns, i.e., the ordering of periodic and nonperiodic directions, may be permuted arbitrarily.

Note that the restriction to the presented structures for mixed periodic and open boundary conditions does not really pose a restriction since an arbitrary given triclinic structure can be traced back to the more simpler cases, see Lemmas 2.10, 2.14 and Remarks 4.10, 4.11. Furthermore, periodic and nonperiodic dimensions are allowed to be ordered arbitrarily.

All required parameters can be easily specified by the user, as for example the near field cutoff radius $r_{\text{cut}} > 0$, the splitting parameter $\alpha > 0$, the mesh size $\boldsymbol{M} \in 2\mathbb{N}^3$, the oversampled mesh size $\boldsymbol{M}_{\text{o}} \in 2\mathbb{N}^3$, the NFFT window function φ, the support parameter $m \in \mathbb{N}$, the shape parameter $\boldsymbol{b} \in \mathbb{R}^3_+$, the type of the applied differentiation scheme (i$\boldsymbol{k}$ or analytic), the degree of smoothness $p \in \mathbb{N}$ as well as the relative size of the regularization domain represented by

$$\varepsilon_{\text{B}} := \frac{1}{2} - \frac{L}{h}. \tag{4.52}$$

We refer to $\varepsilon_{\text{B}} \in [0, 1/2)$ as the regularization parameter. As usual, we denote by $L > 0$ the maximum distance of two particles regarding the nonperiodic dimensions.

Note that different shape parameters may be applied in the single dimensions and, thus, we write $\boldsymbol{b} \in \mathbb{R}^3_+$. However, in our numerical examples we always apply the same shape parameter b in all three dimensions, as already assumed in Section 3.3.

Finally, the P^{2}NFFT algorithm computes

- the potentials $\phi(j)$ as well as the fields $\boldsymbol{E}(j)$ for all charges ($j = 1, \ldots, N_{\text{c}}$),
- the fields $\boldsymbol{E}(j)$ and the negative field gradients $\boldsymbol{G}(j)$ for all dipoles ($j = N_{\text{c}} + 1, \ldots, N$),

which can be written to an output `xml` file.

Based on this, the energies per particle $U(j)$, the total energy U, the acting forces $\boldsymbol{F}(j)$ as well as the torques of the dipoles $\boldsymbol{\tau}(j)$ can be computed afterward with complexity $\mathcal{O}(N)$, see equations (2.13), (2.16), (2.14) and (2.15).

We give a summary of the described algorithm for various types of periodic boundary conditions in Algorithm 4.3. In order to keep the notation simple, we assume that the box vectors $\boldsymbol{\ell}_1$, $\boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$ are sorted such that all periodic directions are listed first.

Algorithm 4.3 (P^{2}NFFT)**.**
Input: Number $\rho \in \{0, 1, 2, 3\}$ of directions subject to periodic boundary conditions, particle positions $\boldsymbol{x}_j \in \mathbb{R}^3$ located in a box spanned by three vectors $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2$ and $\boldsymbol{\ell}_3$ (resulting in a valid box structure as summarized in Figure 4.5), charges $q_j \in \mathbb{R}$, $j = 1, \ldots, N_{\text{c}}$, and dipole moments $\boldsymbol{\mu}_j \in \mathbb{R}^3$, $j = N_{\text{c}} + 1, \ldots, N$. Further required parameters: Ewald parameter $\alpha > 0$, near field cutoff radius $r_{\text{cut}} > 0$,

mesh size $\boldsymbol{M} \in 2\mathbb{N}^3$, oversampled mesh size $\boldsymbol{M}_\mathrm{o} \in 3\mathbb{N}^3$, NFFT window function φ, support parameter $m \in \mathbb{N}$, shape parameter $\boldsymbol{b} \in \mathbb{R}_+^3$, degree of smoothness $p \in \mathbb{N}$, regularization parameter $\varepsilon_\mathrm{B} \in [0, 1/2)$.

i) Precomputations: Compute the set of Fourier coefficients

$$\{c(\boldsymbol{\kappa}), \boldsymbol{\kappa} \in \mathcal{I}_{\boldsymbol{M}}\} := \begin{cases} \left\{(\pi V)^{-1}\hat{\psi}(\boldsymbol{k}), \boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}\right\} \text{ via (2.31)} & : \rho = 3, \\ \left\{\hat{\psi}^{\mathrm{p}2}(\boldsymbol{k}, l), (\boldsymbol{k}, l) \in \mathcal{I}_{\boldsymbol{M}}\right\} \text{ via Alg. 4.1} & : \rho = 2, \\ \left\{\hat{\psi}^{\mathrm{p}1}(k, \boldsymbol{l}), (k, \boldsymbol{l}) \in \mathcal{I}_{\boldsymbol{M}}\right\} \text{ via Alg. 4.2} & : \rho = 1, \\ \left\{\hat{\psi}^{\mathrm{p}0}(\boldsymbol{l}), \boldsymbol{l} \in \mathcal{I}_{\boldsymbol{M}}\right\} \text{ via (4.42), (4.43)} & : \rho = 0. \end{cases}$$

ii) Set $\mathcal{S} := \mathbb{Z}^\rho \times \{0\}^{3-\rho}$ and compute the short range parts $\phi^{\mathrm{short}}(j) \approx \phi_\mathrm{t}^{\mathrm{short}}(j)$, $\boldsymbol{E}^{\mathrm{short}}(j) \approx \boldsymbol{E}_\mathrm{t}^{\mathrm{short}}(j)$ and $\boldsymbol{G}^{\mathrm{short}}(j) \approx \boldsymbol{G}_\mathrm{t}^{\mathrm{short}}(j)$ directly via truncation at radius r_cut, i.e., evaluate (4.1), (4.6) and (4.14), respectively, and take only terms $\|\boldsymbol{x}_{ij} + \boldsymbol{L}\boldsymbol{n}\| \leq r_\mathrm{cut}$ into account.

iii) If $\rho \in \{0, 1, 2\}$: Set

$$L := \begin{cases} |L_3| & : \rho = 2, \\ \sqrt{L_2^2 + L_3^2} & : \rho = 1, \\ \sqrt{L_1^2 + L_2^2 + L_3^2} & : \rho = 0, \end{cases}$$

and $h := (1/2 - \varepsilon_\mathrm{B})^{-1} L$, cf. (4.52).

iv) Define the matrix

$$\boldsymbol{\Lambda} := \begin{cases} \boldsymbol{L} = (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3) & : \rho = 3, \\ \boldsymbol{L}_{\mathrm{p}2} = \mathrm{diag}(\tilde{\boldsymbol{L}}, h) & : \rho = 2, \\ \boldsymbol{L}_{\mathrm{p}1} = \mathrm{diag}(L_1, h, h) & : \rho = 1, \\ \boldsymbol{L}_{\mathrm{p}0} = h\boldsymbol{I}_{3,3} & : \rho = 0. \end{cases}$$

v) Approximate the structure factors

$$S_\mathrm{c}(\boldsymbol{\kappa}) := \sum_{i=1}^{N_\mathrm{c}} q_i \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{\kappa}^\top \boldsymbol{\Lambda}^{-1} \boldsymbol{x}_i} \approx S_{\mathrm{c},\approx}(\boldsymbol{\kappa}), \quad \boldsymbol{\kappa} \in \mathcal{I}_{\boldsymbol{M}},$$

via the adjoint NFFT.

vi) Approximate the dipole structure factors

$$S_\mathrm{d}(\boldsymbol{\kappa}) := \sum_{i=N_\mathrm{c}+1}^{N} \boldsymbol{\mu}_i^\top \nabla_{\boldsymbol{x}_i} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{\kappa}^\top \boldsymbol{\Lambda}^{-1} \boldsymbol{x}_i} \approx S_{\mathrm{d},\approx}(\boldsymbol{\kappa}), \quad \boldsymbol{\kappa} \in \mathcal{I}_{\boldsymbol{M}},$$

via the adjoint gradient NFFT.

vii) Approximate the Fourier space parts of the potentials via the NFFT.

$$\phi^\mathrm{F}(j) \approx \left(\sum_{\boldsymbol{\kappa} \in \mathcal{I}_{\boldsymbol{M}}} c(\boldsymbol{\kappa}) \left[S_{\mathrm{c},\approx}(\boldsymbol{\kappa}) + S_{\mathrm{d},\approx}(\boldsymbol{\kappa}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{\kappa}^\top \boldsymbol{\Lambda}^{-1} \boldsymbol{x}_j} \right)_\approx =: \phi^\mathrm{F}_{\mathrm{p2nfft}}(j)$$

viii) Approximate the Fourier space parts of the fields via the gradient NFFT.

$$\boldsymbol{E}^{\mathrm{F}}(j) \approx -\left(\nabla_{\boldsymbol{x}_j} \sum_{\boldsymbol{\kappa} \in I_{\boldsymbol{M}}} c(\boldsymbol{\kappa}) \left[S_{\mathrm{c},\approx}(\boldsymbol{\kappa}) + S_{\mathrm{d},\approx}(\boldsymbol{\kappa}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{\kappa}^{\top} \boldsymbol{\Lambda}^{-1} \boldsymbol{x}_j} \right)_{\approx} =: \boldsymbol{E}^{\mathrm{F}}_{\mathrm{p2nfft}}(j)$$

ix) Approximate the Fourier space parts of the negative field gradients via the Hessian NFFT.

$$\boldsymbol{G}^{\mathrm{F}}(j) \approx -\left(\nabla_{\boldsymbol{x}_j} \nabla_{\boldsymbol{x}_j}^{\top} \sum_{\boldsymbol{\kappa} \in I_{\boldsymbol{M}}} c(\boldsymbol{\kappa}) \left[S_{\mathrm{c},\approx}(\boldsymbol{\kappa}) + S_{\mathrm{d},\approx}(\boldsymbol{\kappa}) \right] \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{\kappa}^{\top} \boldsymbol{\Lambda}^{-1} \boldsymbol{x}_j} \right)_{\approx} =: \boldsymbol{G}^{\mathrm{F}}_{\mathrm{p2nfft}}(j)$$

x) If $\rho = 3$: Compute the $\boldsymbol{k} = \boldsymbol{0}$ parts $\phi^{\boldsymbol{0}}(j)$ via (2.36) and $\boldsymbol{E}^{\boldsymbol{0}}(j)$ via (4.7).
If non-vacuum boundary conditions are desired, replace the term $3V$ in (2.36) and (4.7) by $3V(1 + \epsilon)$, where ϵ denotes the dielectric constant of the surrounding medium.

xi) Compute the self interactions $\phi^{\mathrm{self}}(j)$ via (2.24), $\boldsymbol{E}^{\mathrm{self}}(j)$ via (4.8) and $\boldsymbol{G}^{\mathrm{self}}(j)$ via (4.15).

Output: Approximate potentials, fields and field gradients, i.e.,

$$\begin{aligned}
\phi_{\mathrm{p2nfft}}(j) &:= \phi^{\mathrm{short}}_{\mathrm{t}}(j) + \phi^{\mathrm{F}}_{\mathrm{p2nfft}}(j) - \phi^{\mathrm{self}}(j) + \delta_{\rho,3}\, \phi^{\boldsymbol{0}}(j), \\
\boldsymbol{E}_{\mathrm{p2nfft}}(j) &:= \boldsymbol{E}^{\mathrm{short}}_{\mathrm{t}}(j) + \boldsymbol{E}^{\mathrm{F}}_{\mathrm{p2nfft}}(j) - \boldsymbol{E}^{\mathrm{self}}(j) + \delta_{\rho,3}\, \boldsymbol{E}^{\boldsymbol{0}}(j), \\
\boldsymbol{G}_{\mathrm{p2nfft}}(j) &:= \boldsymbol{G}^{\mathrm{short}}_{\mathrm{t}}(j) + \boldsymbol{G}^{\mathrm{F}}_{\mathrm{p2nfft}}(j) - \boldsymbol{G}^{\mathrm{self}}(j).
\end{aligned}$$

Afterward, the forces $\boldsymbol{F}(j)$ are obtained based on (2.14) and the torques of the dipole particles $\boldsymbol{\tau}(j)$ are computed via (2.15). The single energies $U(j)$ and the overall energy U can be obtained based on (2.13) and (2.16), respectively. □

We emphasize that the first step of the algorithm, which we refer to as the precomputation step, has to be executed only once, even if the algorithm is applied several times during a simulation involving a large number of time steps. This is due to the fact that these computations do not depend on the positions of the particles, but only on the size of the primary simulation box. A rerun of the precomputation step becomes necessary if particles leave the simulation box in a direction subject to nonperiodic boundary conditions, which requires an extension of the simulation box. We remark that the P^{2}NFFT method has already been used in connection with the simulation software ESPResSo [LAM+06] for simulations in dipolar systems with open boundary conditions, see [WNW+18].

The near field module

The computation of the near field interactions, as described in Step 2 of Algorithm 4.3, is delegated to the near field module, which is shared among several solvers in ScaFaCoS. Arbitrary triclinic box shapes and types of periodic boundary conditions are supported. The module computes particle interactions efficiently based on a linked cell scheme and evaluates a solver-specific short range function, e.g., by applying the Ewald splitting as needed for our method. The massively parallel implementation is based on a sorting library providing efficient data distribution methods, see [DHR07, HR09, HR13].

Far field computations

The computation of the Fourier space parts is done as described in Sections 4.1.1 and 4.1.2 or rather Algorithm 4.3. The massively parallel computation is realized with the PNFFT software [PP13, Pipb], where the Hessian NFFT and the adjoint gradient NFFT have been added as new modules, needed in order to treat systems containing dipoles. Note that the PNFFT algorithm itself is based on the PFFT library [Pip13, Pipa, Pip15] for computing massively parallel FFTs.

The software enables the usage of various NFFT window functions, as for example B-splines, the Bessel and the Gaussian window, as introduced in Section 3.3. Furthermore, also the so called Kaiser-Bessel as well as the Sinc window are available, see [Pip15] for instance. The user of the software may easily switch between the different window functions. In addition, the shape parameter, if present, may be changed. By default, the repeated evaluation of the window function is realized based on a third order interpolation scheme, i.e., a list of function values is computed in advance and saved in an interpolation table. This enables an efficient and very accurate evaluation, which is independent from the used window function. Thus, only a few arithmetical operations are required, which is in terms of computational efficiency comparable to the fast Gaussian gridding [GL04].

The computation of the Fourier space parts follows the same modular process for all types of boundary conditions, which is the computation of the structure factors via the adjoint (gradient) NFFT, followed by the multiplication with the Fourier coefficients and a (gradient or Hessian) NFFT for the approximation of the outer sums. A change of periodicity can be performed easily by replacing the Fourier coefficients $\hat{\psi}(\boldsymbol{k})$ from the 3d-periodic case with some appropriately precomputed coefficients, as described before. In addition, the nature of the used Fourier coefficients is not relevant for the presented modifications for dipole-dipole or charge-dipole systems. All formerly derived features, as for example the suitability for massively parallel architectures, cf. [PP13, Pip15], and the application of interlacing, are also available in combination with the presented extensions. For a detailed discussion of the interlacing technique within the scope of P^{3}M and P^{2}NFFT we refer to [NH10, CBH11, Pip15].

In the precomputation step for mixed periodic and open boundary conditions we compute the necessary regularized kernel functions, cf. (4.18), (4.31) and (4.42). The required Fourier coefficients are obtained by sampling the regularized functions at equispaced nodes and applying the FFT regarding the relevant dimensions. Therefor, the evaluation of the kernel functions as well as of their derivatives up to a certain order is required. These derivatives can be computed by the presented analytic representations, see (4.20), (4.21), Lemmas 4.7 and 4.8 as well as (4.44). In order to save time within the precomputation step we used the presented recursions in combination with a computer algebra system in order to precompute the analytic expressions of the considered derivatives up to a sufficient order.

We remark that we face several numerical problems concerning the computation of the required derivatives in case of 2d-periodic and 1d-periodic boundary conditions. Details on overcoming these challenges are discussed in [NPP15a]. Especially, the evaluation of the modified incomplete Bessel function in case of 1d-periodic constraints is difficult and there are only a few publicly available implementations for its evaluation. Within the P^{2}NFFT precomputation step we evaluated the function based on the iterative algorithms presented in [SS10], with some modifications in order to overcome numerical problems, see [NPP15a]. Later on, the implementation could be further improved based on [SW15].

P^{2}NDFT Ewald

In addition, the PNFFT library also allows the direct evaluation of the various trigonometric sums. We refer to the direct computation of the sums (3.6) as discrete Fourier transform for nonequispaced data (NDFT). The NDFT ends up with an undesirable arithmetic complexity of $O(N|\mathcal{I}_{\boldsymbol{M}}|)$. Analogously, the sums (3.21), (3.24), (3.29), and (3.31) may be obtained via a direct evaluation, which we denote as the adjoint NDFT, the gradient NDFT, the Hessian NDFT and the adjoint gradient NDFT, respectively. The P^{2}NFFT interface in ScaFaCoS enables the user to easily switch between the fast computation and the direct evaluation (via flag `pnfft_direct`).

Remark 4.12 (P^{2}NDFT Ewald)**.** By applying the direct evaluation methods (NDFT) instead of the fast modules (NFFT), the described method is in case of 3d-periodic boundary conditions equivalent to a pure Ewald summation, where only the Ewald truncation errors, see Section 2.3, are present. Thus, we refer to this method as the P^{2}NDFT Ewald algorithm. This method may serve as a reference method for numerical experiments since its approximation error is well controllable. □

4.3 Approximation errors

In the field of molecular dynamics simulations the root mean square (rms) error in the forces is a common and adequate measure of accuracy. The rms errors in the forces arising from the truncation of the Ewald summation formulas have been considered in Section 2.3.

The NFFT based approximation of the truncated Fourier space sums introduce further approximation errors, which are considered in the following.

We split the Fourier space part of the force into two parts and truncate the infinite sums, i.e.,

$$\boldsymbol{F}^{\mathrm{F}}(j) = \begin{cases} \boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.c.}}(j) + \boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.d.}}(j) \approx \boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.c.}}(j)_{\mathrm{t}} + \boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.d.}}(j)_{\mathrm{t}} & j = 1, \ldots, N_{\mathrm{c}}, \\ \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j) + \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j) \approx \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j)_{\mathrm{t}} + \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j)_{\mathrm{t}} & j = N_{\mathrm{c}}+1, \ldots, N, \end{cases}$$

see Section 2.3.3.

In the following we derive estimates for the rms errors resulting from approximating the truncated forces via the different NFFT modules. In order to keep the notation short, we define for $n \in \{0, 2, 4, 6\}$ the sums

$$Y_n(\boldsymbol{k}) := \sum_{\boldsymbol{r} \in \mathbb{Z}^3} \|\boldsymbol{L}^{-\top}(\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}})\|^n c^2_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}}),$$

as well as

$$\tilde{Y}_n(\boldsymbol{k}) := \sum_{\boldsymbol{r} \in \mathbb{Z}^3 \setminus \{\boldsymbol{0}\}} \|\boldsymbol{L}^{-\top}(\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}})\|^n c^2_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}}) = Y_n(\boldsymbol{k}) - \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^n c^2_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}}).$$

Again, we rewrite the approximation errors via

$$\begin{aligned}
\boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.c.}}(j)_{\mathrm{t}} - \boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.c.}}(j)_{\approx} &= q_j \sum_{i=1}^{N_{\mathrm{c}}} q_i \boldsymbol{\chi}^{\mathrm{c.c.}}_{ij},\\
\boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.d.}}(j)_{\mathrm{t}} - \boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.d.}}(j)_{\approx} &= q_j \sum_{i=N_{\mathrm{c}}+1}^{N} \|\boldsymbol{\mu}_i\| \boldsymbol{\chi}^{\mathrm{c.d.}}_{ij},\\
\boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j)_{\mathrm{t}} - \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j)_{\approx} &= \|\boldsymbol{\mu}_j\| \sum_{i=1}^{N_{\mathrm{c}}} q_i \boldsymbol{\chi}^{\mathrm{d.c.}}_{ij},\\
\boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j)_{\mathrm{t}} - \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j)_{\approx} &= \|\boldsymbol{\mu}_j\| \sum_{i=N_{\mathrm{c}}+1}^{N} \|\boldsymbol{\mu}_i\| \boldsymbol{\chi}^{\mathrm{d.d.}}_{ij},
\end{aligned} \tag{4.53}$$

where $F^{\mathrm{F}}_{\mathrm{type}}(j)_{\approx} = F^{\mathrm{F}}_{\mathrm{type}}(j)_{\mathrm{ik}}$ in case of $\mathrm{i}\boldsymbol{k}$-differentiation and $F^{\mathrm{F}}_{\mathrm{type}}(j)_{\approx} = F^{\mathrm{F}}_{\mathrm{type}}(j)_{\mathrm{ad}}$ for the analytic differentiation approach, where the place marker type stands for the different types of errors (c.c., c.d., d.c., d.d.).

In order to estimate the expected force errors, we compute the quadratic means

$$\chi^2_{\mathrm{type}} := \frac{1}{V^2} \int_{\boldsymbol{L}\mathbb{T}} \int_{\boldsymbol{L}\mathbb{T}} \|\boldsymbol{\chi}^{\mathrm{type}}_{ij}\|^2 \mathrm{d}\boldsymbol{x}_i \mathrm{d}\boldsymbol{x}_j \tag{4.54}$$

for the four different types (c.c., c.d., d.c., d.d.).

In order to keep the notation short, we furthermore define

$$Z(\boldsymbol{k}, \boldsymbol{x}) := \sum_{\boldsymbol{r} \in \mathbb{Z}^3} c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\, \mathrm{e}^{2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}} \tag{4.55}$$

in the following. The $\mathrm{i}\boldsymbol{k}$-differentiation approach is considered in Section 4.3.1 and the analytic differentiation approach is discussed in Section 4.3.2.

4.3.1 Differentiation in Fourier space

We recall that the Fourier space parts of the forces acting on the charges are approximated by applying the gradient NFFT. Based on definition (4.55) we can rewrite the approximated forces as follows, cf. (3.25).

$$\begin{aligned}
\boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.c.}}(j)_{\mathrm{t}} &\approx \frac{2\pi\mathrm{i}q_j}{\pi V} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \sum_{\boldsymbol{r}\in\mathbb{Z}^3} (\boldsymbol{L}^{-\top}\boldsymbol{k}) \hat{d}_{\boldsymbol{k}} \hat{\psi}(\boldsymbol{k}) S_{\mathrm{c},\approx}(\boldsymbol{k}) c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\, \mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_j}\\
&= \frac{2\mathrm{i}q_j}{V} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} (\boldsymbol{L}^{-\top}\boldsymbol{k}) \hat{d}_{\boldsymbol{k}} \hat{\psi}(\boldsymbol{k}) S_{\mathrm{c},\approx}(\boldsymbol{k}) Z(\boldsymbol{k}, -\boldsymbol{x}_j) =: \boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.c.}}(j)_{\mathrm{ik}}
\end{aligned} \tag{4.56}$$

and

$$\begin{aligned}
\boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.d.}}(j)_{\mathrm{t}} &\approx \frac{2\pi\mathrm{i}q_j}{\pi V} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \sum_{\boldsymbol{r}\in\mathbb{Z}^3} (\boldsymbol{L}^{-\top}\boldsymbol{k}) \hat{d}_{\boldsymbol{k}} \hat{\psi}(\boldsymbol{k}) S_{\mathrm{d,ik}}(\boldsymbol{k}) c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\, \mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_j}\\
&= \frac{2\mathrm{i}q_j}{V} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} (\boldsymbol{L}^{-\top}\boldsymbol{k}) \hat{d}_{\boldsymbol{k}} \hat{\psi}(\boldsymbol{k}) S_{\mathrm{d,ik}}(\boldsymbol{k}) Z(\boldsymbol{k}, -\boldsymbol{x}_j) =: \boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.d.}}(j)_{\mathrm{ik}}.
\end{aligned}$$

The Fourier space parts of the forces for the dipole particles are obtained analogously by using the Hessian NFFT. By making use of (3.30) we get

$$\begin{aligned}\boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j)_{\mathrm{t}} &\approx \frac{4\pi^2}{\pi V}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}\sum_{\boldsymbol{r}\in\mathbb{Z}^3}(\boldsymbol{L}^{-\top}\boldsymbol{k})\boldsymbol{\mu}_j^{\top}(\boldsymbol{L}^{-\top}\boldsymbol{k})\hat{d}_{\boldsymbol{k}}\hat{\psi}(\boldsymbol{k})S_{\mathrm{c},\approx}(\boldsymbol{k})c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\,\mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_j}\\ &= \frac{4\pi}{V}\|\boldsymbol{\mu}_j\|\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}(\boldsymbol{L}^{-\top}\boldsymbol{k})\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|\cos\gamma_{\boldsymbol{k},j}\hat{d}_{\boldsymbol{k}}\hat{\psi}(\boldsymbol{k})S_{\mathrm{c},\approx}(\boldsymbol{k})Z(\boldsymbol{k},-\boldsymbol{x}_j) =: \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j)_{\mathrm{ik}}\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j)_{\mathrm{t}} &\approx \frac{4\pi^2}{\pi V}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}\sum_{\boldsymbol{r}\in\mathbb{Z}^3}(\boldsymbol{L}^{-\top}\boldsymbol{k})\boldsymbol{\mu}_j^{\top}(\boldsymbol{L}^{-\top}\boldsymbol{k})\hat{d}_{\boldsymbol{k}}\hat{\psi}(\boldsymbol{k})S_{\mathrm{d,ik}}(\boldsymbol{k})c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\,\mathrm{e}^{-2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_j}\\ &= \frac{4\pi}{V}\|\boldsymbol{\mu}_j\|\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}(\boldsymbol{L}^{-\top}\boldsymbol{k})\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|\cos\gamma_{\boldsymbol{k},j}\hat{d}_{\boldsymbol{k}}\hat{\psi}(\boldsymbol{k})S_{\mathrm{d,ik}}(\boldsymbol{k})Z(\boldsymbol{k},-\boldsymbol{x}_j) =: \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j)_{\mathrm{ik}}.\end{aligned}\tag{4.57}$$

In the following theorem we present estimates of the quantities $\chi^2_{\mathrm{c.c.}}$, $\chi^2_{\mathrm{c.d.}}$, $\chi^2_{\mathrm{d.c.}}$ and $\chi^2_{\mathrm{d.d.}}$, as defined via (4.53) and (4.54).

Theorem 4.13. *Let a charge neutral system* (2.5) *with N_{c} charges $q_j \in \mathbb{R}$, $j = 1,\dots,N_{\mathrm{c}}$, and N_{d} dipoles with dipole moments $\boldsymbol{\mu}_j \in \mathbb{R}^3$, $j = N_{\mathrm{c}}+1,\dots,N$, be given. Assume that the truncated Fourier space parts of the forces $\boldsymbol{F}^{\mathrm{F}}_{\mathrm{t}}(j)$ are approximated via the P²NFFT method with oversampled mesh size $\boldsymbol{M}_{\mathrm{o}}$ and a symmetric window function φ. In case of $\mathrm{i}\boldsymbol{k}$-differentiation we obtain*

$$\chi^2_{\mathrm{c.c.}} = \frac{4}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2\hat{\psi}(\boldsymbol{k})^2\left(1-2\hat{d}^2_{\boldsymbol{k}}c^2_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})+\hat{d}^4_{\boldsymbol{k}}Y_0(\boldsymbol{k})^2\right),\tag{4.58}$$

$$\chi^2_{\mathrm{c.d.}} = \chi^2_{\mathrm{d.c.}} \le \frac{16\pi^2}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^4\hat{\psi}(\boldsymbol{k})^2\left(1-2\hat{d}^2_{\boldsymbol{k}}c^2_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})+\hat{d}^4_{\boldsymbol{k}}Y_0(\boldsymbol{k})^2\right),\tag{4.59}$$

$$\chi^2_{\mathrm{d.d.}} \le \frac{64\pi^4}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^6\hat{\psi}(\boldsymbol{k})^2\left(1-2\hat{d}^2_{\boldsymbol{k}}c^2_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})+\hat{d}^4_{\boldsymbol{k}}Y_0(\boldsymbol{k})^2\right).\tag{4.60}$$

Proof. By utilizing (3.23) as well as (3.32) we obtain the following representations of the approximations $S_{\mathrm{c},\approx}(\boldsymbol{k})$ and $S_{\mathrm{d,ik}}(\boldsymbol{k})$, respectively.

$$S_{\mathrm{c},\approx}(\boldsymbol{k}) = \hat{d}_{\boldsymbol{k}}\sum_{i=1}^{N_{\mathrm{c}}}q_i\sum_{\boldsymbol{r}\in\mathbb{Z}^3}c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\,\mathrm{e}^{2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_i} = \hat{d}_{\boldsymbol{k}}\sum_{i=1}^{N_{\mathrm{c}}}q_iZ(\boldsymbol{k},\boldsymbol{x}_i)\tag{4.61}$$

and

$$\begin{aligned}S_{\mathrm{d,ik}}(\boldsymbol{k}) &= 2\pi\mathrm{i}\hat{d}_{\boldsymbol{k}}\sum_{i=N_{\mathrm{c}}+1}^{N}\boldsymbol{\mu}_i^{\top}(\boldsymbol{L}^{-\top}\boldsymbol{k})\sum_{\boldsymbol{r}\in\mathbb{Z}^3}c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\,\mathrm{e}^{2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}_i}\\ &= 2\pi\mathrm{i}\hat{d}_{\boldsymbol{k}}\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|\sum_{i=N_{\mathrm{c}}+1}^{N}\|\boldsymbol{\mu}_i\|\cos\gamma_{\boldsymbol{k},i}Z(\boldsymbol{k},\boldsymbol{x}_i).\end{aligned}\tag{4.62}$$

Thus, we obtain

$$F^{\mathrm{F}}_{\mathrm{c.c.}}(j)_{\mathrm{ik}} = \frac{2\mathrm{i}q_j}{V}\sum_{i=1}^{N_{\mathrm{c}}}q_i\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}}(\boldsymbol{L}^{-\top}\boldsymbol{k})\hat{d}^2_{\boldsymbol{k}}\hat{\psi}(\boldsymbol{k})Z(\boldsymbol{k},\boldsymbol{x}_i)Z(\boldsymbol{k},-\boldsymbol{x}_j),$$

i.e.,

$$\chi_{ij}^{\text{c.c.}} = \frac{2\mathrm{i}}{V} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} (\boldsymbol{L}^{-\top}\boldsymbol{k}) \hat{\psi}(\boldsymbol{k}) \left(\mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{L}^{-1} \boldsymbol{x}_{ij}} - \hat{d}_{\boldsymbol{k}}^2 Z(\boldsymbol{k}, \boldsymbol{x}_i) Z(\boldsymbol{k}, -\boldsymbol{x}_j) \right).$$

Integration with respect to $\boldsymbol{x}_i$ yields

$$\begin{aligned} \frac{1}{V} \int_{\boldsymbol{L}\mathbb{T}^3} \|\chi_{ij}^{\text{c.c.}}\|^2 \mathrm{d}\boldsymbol{x}_i = \frac{4}{V^2} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 \hat{\psi}(\boldsymbol{k})^2 \left| \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{L}^{-1} \boldsymbol{x}_j} - \hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}}) Z(\boldsymbol{k}, -\boldsymbol{x}_j) \right|^2 \\ + \frac{4}{V^2} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 \hat{\psi}(\boldsymbol{k})^2 \hat{d}_{\boldsymbol{k}}^4 \tilde{Y}_0(\boldsymbol{k}) |Z(\boldsymbol{k}, -\boldsymbol{x}_j)|^2, \end{aligned}$$

where we simply use Parseval's identity. Applying the identity once again yields

$$\chi_{\text{c.c.}}^2 = \frac{4}{V^2} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 \hat{\psi}(\boldsymbol{k})^2 \left((1 - \hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}^2(\tilde{\varphi}_{\mathrm{t}}))^2 + \hat{d}_{\boldsymbol{k}}^4 c_{\boldsymbol{k}}^2(\tilde{\varphi}_{\mathrm{t}}) \tilde{Y}_0(\boldsymbol{k}) + \hat{d}_{\boldsymbol{k}}^4 \tilde{Y}_0(\boldsymbol{k}) Y_0(\boldsymbol{k}) \right).$$

Replacing $\tilde{Y}_0(\boldsymbol{k})$ by $Y_0(\boldsymbol{k}) - c_{\boldsymbol{k}}^2(\tilde{\varphi}_{\mathrm{t}})$ and simplifying further gives (4.58).

In exactly the same manner we obtain

$$\chi_{ij}^{\text{c.d.}} = -\frac{4\pi}{V} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} (\boldsymbol{L}^{-\top}\boldsymbol{k}) \hat{\psi}(\boldsymbol{k}) \|\boldsymbol{L}^{-\top}\boldsymbol{k}\| \cos \gamma_{\boldsymbol{k},i} \left(\mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{L}^{-1} \boldsymbol{x}_{ij}} - \hat{d}_{\boldsymbol{k}}^2 Z(\boldsymbol{k}, \boldsymbol{x}_i) Z(\boldsymbol{k}, -\boldsymbol{x}_j) \right).$$

Integration with respect to $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ and estimating $\cos^2 \gamma_{i,\boldsymbol{k}} \leq 1$ gives (4.59). The same is obtained for the dipole-charge term, where

$$\chi_{ij}^{\text{d.c.}} = -\frac{4\pi}{V} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} (\boldsymbol{L}^{-\top}\boldsymbol{k}) \hat{\psi}(\boldsymbol{k}) \|\boldsymbol{L}^{-\top}\boldsymbol{k}\| \cos \gamma_{\boldsymbol{k},j} \left(\mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{L}^{-1} \boldsymbol{x}_{ij}} - \hat{d}_{\boldsymbol{k}}^2 Z(\boldsymbol{k}, \boldsymbol{x}_i) Z(\boldsymbol{k}, -\boldsymbol{x}_j) \right).$$

For the dipole-dipole interactions we easily see

$$\chi_{ij}^{\text{d.d.}} = \frac{8\pi^2 \mathrm{i}}{V} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} (\boldsymbol{L}^{-\top}\boldsymbol{k}) \hat{\psi}(\boldsymbol{k}) \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 \cos \gamma_{\boldsymbol{k},i} \cos \gamma_{\boldsymbol{k},j} \left(\mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}^{\top} \boldsymbol{L}^{-1} \boldsymbol{x}_{ij}} - \hat{d}_{\boldsymbol{k}}^2 Z(\boldsymbol{k}, \boldsymbol{x}_i) Z(\boldsymbol{k}, -\boldsymbol{x}_j) \right).$$

Integration and estimating the quadratic cosine terms from above by 1 gives (4.60). ■

Remark 4.14 (Undesired self interaction)**.** In Remark 4.1 we discussed the introduced systematic errors regarding the approximations $\boldsymbol{F}_{\text{c.c.}}^{\text{F}}(j) \approx \boldsymbol{F}_{\text{c.c.}}^{\text{F}}(j)_{\mathrm{t}}$ and $\boldsymbol{F}_{\text{d.d.}}^{\text{F}}(j) \approx \boldsymbol{F}_{\text{d.d.}}^{\text{F}}(j)_{\mathrm{t}}$.

Similarly, we compute undesired self interaction terms in (4.56) and (4.57), which read as

$$\boldsymbol{F}_{\text{c.c.}}^{\text{F},j}(j)_{\text{ik}} := \frac{2\mathrm{i}q_j^2}{V} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}} \setminus \mathcal{I}_{\boldsymbol{M}}^*} (\boldsymbol{L}^{-\top}\boldsymbol{k}) \hat{d}_{\boldsymbol{k}}^2 \hat{\psi}(\boldsymbol{k}) Z(\boldsymbol{k}, \boldsymbol{x}_j) Z(\boldsymbol{k}, -\boldsymbol{x}_j)$$

and

$$\boldsymbol{F}_{\text{d.d.}}^{\text{F},j}(j)_{\text{ik}} := \frac{8\pi \mathrm{i}}{V} \|\boldsymbol{\mu}_j\|^2 \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}} \setminus \mathcal{I}_{\boldsymbol{M}}^*} (\boldsymbol{L}^{-\top}\boldsymbol{k}) \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 \cos^2 \gamma_{\boldsymbol{k},j} \hat{d}_{\boldsymbol{k}}^2 \hat{\psi}(\boldsymbol{k}) Z(\boldsymbol{k}, \boldsymbol{x}_j) Z(\boldsymbol{k}, -\boldsymbol{x}_j),$$

respectively. The result is obtained by extracting the term with $\boldsymbol{x}_i = \boldsymbol{x}_j$ in $S_{\mathrm{c},\approx}(\boldsymbol{k})$ or $S_{\mathrm{d,ik}}(\boldsymbol{k})$ and applying (4.61) or (4.10), respectively. Since the NFFT window is assumed to be symmetric, we obtain

$$Z(\boldsymbol{k}, -\boldsymbol{x}_j) = Z(-\boldsymbol{k}, \boldsymbol{x}_j)$$

and, as a result of the underlying symmetry, the terms for $\pm\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}^*$ have opposite signs and cancel out each other.

Note that the resulting errors are not completely independent from the particle positions and, thus, we do not speak about systematic errors in this context.

However, when using symmetric index sets, the i$\boldsymbol{k}$-differentiation approach is by construction momentum conserving, since the self forces are zero and the forces regarding the interaction between two particles are equal with opposite signs, cf. [HE88, Section 5.3.3]. In contrast, the i$\boldsymbol{k}$-differentiation approach is not energy conserving, see [HE88, Sections 5.5 and 7.6]. □

We denote the resulting rms errors by

$$\begin{aligned}
\Delta F_{\mathrm{c.c.}}^{\mathrm{F,ik}} &\approx \frac{Q\chi_{\mathrm{c.c.}}}{\sqrt{N_{\mathrm{c}}}} && \text{with } \chi_{\mathrm{c.c.}} \text{ as given in (4.58),}\\
\Delta F_{\mathrm{c.d.}}^{\mathrm{F,ik}} &\approx \sqrt{\frac{Q\mathcal{M}}{N_{\mathrm{c}}}}\chi_{\mathrm{c.d.}} && \text{with } \chi_{\mathrm{c.d.}} \text{ as given in (4.59),}\\
\Delta F_{\mathrm{d.c.}}^{\mathrm{F,ik}} &\approx \sqrt{\frac{Q\mathcal{M}}{N_{\mathrm{d}}}}\chi_{\mathrm{d.c.}} && \text{with } \chi_{\mathrm{d.c.}} \text{ as given in (4.59),}\\
\Delta F_{\mathrm{d.d.}}^{\mathrm{F,ik}} &\approx \frac{\mathcal{M}\chi_{\mathrm{d.d.}}}{\sqrt{N_{\mathrm{d}}}} && \text{with } \chi_{\mathrm{d.d.}} \text{ as given in (4.60).}
\end{aligned} \tag{4.63}$$

We obtain approximations of the actual errors based on Theorem 4.13, provided that the diagonal contributions with $i = j$, see Remark 4.14, are negligible. In this context, we remark that the influence of the present self interactions on the overall approximation errors has not been studied so far. Hence, it is widely unknown how much the usage of non symmetric index sets $\mathcal{I}_{\boldsymbol{M}}$ might influence the accuracy of the P^{2}NFFT method in general.

Corollary 4.15. *For the* i$\boldsymbol{k}$*-differentiation approach the rms force errors become minimal for the deconvolution coefficients* (3.17).

Proof. The result is obtained easily by differentiating the expressions

$$1 - 2\hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}^2(\tilde{\varphi}_{\mathrm{t}}) + \hat{d}_{\boldsymbol{k}}^4 Y_0(\boldsymbol{k})^2$$

with respect to $\hat{d}_{\boldsymbol{k}}$ and setting the results equal to zero, which gives

$$\begin{aligned}
&-4\hat{d}_{\boldsymbol{k}}\, c_{\boldsymbol{k}}^2(\tilde{\varphi}_{\mathrm{t}}) + 4\hat{d}_{\boldsymbol{k}}^3 Y_0(\boldsymbol{k})^2 = 0\\
\iff\ &\hat{d}_{\boldsymbol{k}} \left(c_{\boldsymbol{k}}^2(\tilde{\varphi}_{\mathrm{t}}) - \hat{d}_{\boldsymbol{k}}^2 Y_0(\boldsymbol{k})^2 \right) = 0.
\end{aligned}$$

By taking also the second derivative into consideration we see that the minimum is indeed obtained for

$$\hat{d}_{\boldsymbol{k}} = \pm c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}}) Y_0(\boldsymbol{k})^{-1}.$$

Finally, we obtain

$$1 - 2\hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}^2(\tilde{\varphi}_\mathrm{t}) + \hat{d}_{\boldsymbol{k}}^4 Y_0(\boldsymbol{k})^2 = \begin{cases} \left(\dfrac{Y_0(\boldsymbol{k})}{c_{\boldsymbol{k}}^2(\tilde{\varphi}_\mathrm{t})}\right)^2 - 1 & : \hat{d}_{\boldsymbol{k}} = c_{\boldsymbol{k}}^{-1}(\tilde{\varphi}_\mathrm{t}), \\ 1 - \left(\dfrac{Y_0(\boldsymbol{k})}{c_{\boldsymbol{k}}^2(\tilde{\varphi}_\mathrm{t})}\right)^{-2} & : \hat{d}_{\boldsymbol{k}} = c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t}) Y_0(\boldsymbol{k})^{-1}. \end{cases}$$

■

Efficient computation of the resulting rms errors

We consider the computation of

$$\begin{aligned} &\sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^n \hat{\psi}(\boldsymbol{k})^2 \left(1 - 2\hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}^2(\tilde{\varphi}_\mathrm{t}) + \hat{d}_{\boldsymbol{k}}^4 Y_0(\boldsymbol{k})^2\right) \\ &= \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}} \setminus \{\boldsymbol{0}\}} \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^{n-4} \, \mathrm{e}^{-2\pi^2 \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2/\alpha^2} \left(1 - 2\hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}^2(\tilde{\varphi}_\mathrm{t}) + \hat{d}_{\boldsymbol{k}}^4 Y_0(\boldsymbol{k})^2\right), \end{aligned}$$

cf. (4.58), (4.59) and (4.60).

For simplicity we restrict our considerations to the standard deconvolution approach, where we set $\hat{d}_{\boldsymbol{k}} := c_{\boldsymbol{k}}^{-1}(\tilde{\varphi}_\mathrm{t})$. We remark that the optimized deconvolution approach has been discussed in [Nes16c]. A comparison between the two variants yielded that the optimized deconvolution does not promise significant improvements in terms of accuracy, which is mainly due to the rapid decrease of the Fourier coefficients (2.31). The same has been observed in Section 3.3.1 and [Nes16a].

In case of the standard deconvolution approach we obtain [Nes16c]

$$\begin{aligned} &\sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}} \setminus \{\boldsymbol{0}\}} \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^{n-4} \, \mathrm{e}^{-2\pi^2 \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2/\alpha^2} \left(1 - 2\hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}^2(\tilde{\varphi}_\mathrm{t}) + \hat{d}_{\boldsymbol{k}}^4 Y_0(\boldsymbol{k})^2\right) \\ &= \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}} \setminus \{\boldsymbol{0}\}} \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^{n-4} \, \mathrm{e}^{-2\pi^2 \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2/\alpha^2} \left(\left(\frac{Y_0(\boldsymbol{k})}{c_{\boldsymbol{k}}^2(\tilde{\varphi}_\mathrm{t})}\right)^2 - 1\right) \\ &\leq \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}} \setminus \{\boldsymbol{0}\}} \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^{n-4} \, \mathrm{e}^{-2\pi^2 \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2/\alpha^2} \left(\prod_{j=1}^{3} \big(1 + s_j(k_j)\big)^2 - 1\right) \\ &= \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}} \setminus \{\boldsymbol{0}\}} \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^{n-4} \, \mathrm{e}^{-2\pi^2 \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2/\alpha^2} \left(\prod_{j=1}^{3} \big(1 + \underbrace{2s_j(k_j) + s_j^2(k_j)}_{=: \, \tilde{s}_j(k_j)}\big) - 1\right), \end{aligned} \tag{4.64}$$

where we make use of the upper limits $s_j(k_j)$ in the single dimensions, cf. (3.43).

The sums (4.64) may be evaluated directly with $O(|\mathcal{I}_{\boldsymbol{M}}|)$ arithmetic operations. In the case that the box vectors are orthogonal to each other, see Remark 4.16, we may use a more efficient evaluation scheme.

Remark 4.16. If $\boldsymbol{L}$ is a diagonal matrix $\boldsymbol{L} = \operatorname{diag}(L_1, L_2, L_3)$, i.e., the particles are located in a orthorhombic box, the computation of (4.64) is possible in $O(M_1 + M_2 + M_3)$ for $n \in \{4, 6\}$.

This can be seen as follows. We have $\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 = {}^{k_1^2}/_{L_1^2} + {}^{k_2^2}/_{L_2^2} + {}^{k_3^2}/_{L_3^2}$ and define

$$R_j^{(p)} := \sum_{k=-M_j/2}^{M_j/1-1} \left(\frac{k}{L_j}\right)^p \mathrm{e}^{-2\pi^2 k^2/(\alpha^2 L_j^2)}$$

$$S_j^{(p)} := \sum_{k=-M_j/2}^{M_j/1-1} \tilde{s}_j(k) \left(\frac{k}{L_j}\right)^p \mathrm{e}^{-2\pi^2 k^2/(\alpha^2 L_j^2)}.$$

Using this notation, we get

$$\chi^2_{\mathrm{c.d.}} = \chi^2_{\mathrm{d.c.}} \le \frac{16\pi^2}{V^2}\left(\sum_{j_1,j_2,j_3}\left(R_{j_1}^{(0)}R_{j_2}^{(0)}S_{j_3}^{(0)} + R_{j_1}^{(0)}S_{j_2}^{(0)}S_{j_3}^{(0)}\right) + S_1^{(0)}S_2^{(0)}S_3^{(0)} - \prod_{j=1}^{3}(1+\tilde{s}_j(0)) + 1\right)$$

$$\chi^{(2)}_{\mathrm{d.d.}} \le \frac{64\pi^4}{V^2}\left(\sum_{j_1,j_2,j_3}\left(R_{j_1}^{(2)}R_{j_2}^{(2)}S_{j_3}^{(2)} + R_{j_1}^{(2)}S_{j_2}^{(2)}S_{j_3}^{(2)}\right) + S_1^{(2)}S_2^{(2)}S_3^{(2)}\right).$$

In order to estimate $\chi^2_{\mathrm{c.c.}}$ we need to evaluate (4.64) with $n = 2$. Due to the presence of the term $\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^{-2}$ the evaluation is not readily possible with only $O(M_1 + M_2 + M_3)$ arithmetic operations.

Suppose that we have given an approximation of the form

$$\frac{1}{x} \approx \sum_{i=1}^{n} r_i \,\mathrm{e}^{-\omega_i x}, \tag{4.65}$$

which is valid for all x in a sufficiently large interval. In our numerical tests we apply an approximation presented in [Hac05] with only $n = 11$ exponential terms, which are required in order to keep the relative approximation errors below 1 for all $x \in [1, 2 \cdot 10^5)$. The maximum absolute approximation error is $\epsilon_{\mathrm{abs}} \approx 6.8 \cdot 10^{-6}$. The relative errors are small for small values of x and adopt values of approximately 1 for large x.

We obtain

$$1 \le x := L_{\max}^2 \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 = L_{\max}^2\left(\frac{k_1^2}{L_1^2} + \frac{k_2^2}{L_2^2} + \frac{k_3^2}{L_3^2}\right) \quad \forall\, \boldsymbol{k} \neq \boldsymbol{0},$$

where $L_{\max} := \max\{L_1, L_2, L_3\}$. Based on that, we get

$$\begin{aligned}\chi^2_{\mathrm{c.c.}} &\le \frac{4}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}\setminus\{\boldsymbol{0}\}} \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^{-2}\,\mathrm{e}^{-2\pi^2\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2/\alpha^2}\left(\prod_{j=1}^{3}\left(1+\tilde{s}_j(k_j)\right) - 1\right)\\ &\approx \frac{4L_{\max}^2}{V^2}\sum_{i=1}^{n} r_i\left(\sum_{j_1,j_2,j_3}\left(R_{i,j_1}R_{i,j_2}S_{i,j_3} + R_{i,j_1}S_{i,j_2}S_{i,j_3}\right) + S_{i,1}S_{i,2}S_{i,3} - \prod_{j=1}^{3}(1+\tilde{s}_j(0)) + 1\right),\end{aligned}$$

where we set

$$R_{i,j} := \sum_{k=-M_j/2}^{M_j/1-1} \mathrm{e}^{-2\pi^2k^2/(\alpha^2L_j^2) - \omega_i k^2 L_{\max}^2/L_j^2}$$

$$S_{i,j} := \sum_{k=-M_j/2}^{M_j/1-1} \tilde{s}_j(k)\,\mathrm{e}^{-2\pi^2k^2/(\alpha^2L_j^2) - \omega_i k^2 L_{\max}^2/L_j^2}$$

for all $i = 1, \ldots, n$ and $j \in \{1, 2, 3\}$. □

Based on the presented error estimates we may also tune the NFFT parameters in order to achieve a required accuracy of the P²NFFT method. In case of orthorhombic box shapes the tuning is possible with only $O(M_1 + M_2 + M_3)$ arithmetic operations, as discussed in Remark 4.16.

Remark 4.17 (Shape parameter tuning P²NFFT)**.** The shape parameter of the Bessel window function, for instance, may be optimized similarly as presented in Algorithm 3.4 via

$$\min_b \max \left\{ \Delta F^{\mathrm{F,ik}}_{\mathrm{c.c.}}, \Delta F^{\mathrm{F,ik}}_{\mathrm{c.d.}}, \Delta F^{\mathrm{F,ik}}_{\mathrm{d.c.}}, \Delta F^{\mathrm{F,ik}}_{\mathrm{d.d.}} \right\}.$$

□

In the following Example we demonstrate how the choice of the shape parameter influences the approximation error of the P²NFFT method. In the same breath, we test the above described automated tuning procedure.

Example 4.18. We set $N_{\mathrm{c}} = N_{\mathrm{d}} = Q = \mathcal{M} := 150$, assume a cubic primary simulation box of size $10 \times 10 \times 10$ and compute the error estimates (4.63) based on Theorem 4.13.

For the NFFT approximations we assume a support parameter $m := 4$ and an oversampling factor $\sigma := 1$. In a first test we computed the resulting errors for different shape parameters b, where we fixed the Ewald parameter $\alpha := 1$, see left hand side of Figure 4.6. Again, the errors are highly influenced by the choice of the shape parameter b. For comparison, with the B-spline window we obtain

$$\begin{aligned} \Delta F^{\mathrm{F,ik}}_{\mathrm{c.c.}} &\approx 3.5372 \cdot 10^{-7}, \\ \Delta F^{\mathrm{F,ik}}_{\mathrm{c.d.}} = \Delta F^{\mathrm{F,ik}}_{\mathrm{d.c.}} &\approx 1.6398 \cdot 10^{-6}, \\ \Delta F^{\mathrm{F,ik}}_{\mathrm{d.d.}} &\approx 7.7628 \cdot 10^{-6}, \end{aligned}$$

where we applied the same parameters m, σ and α.

In a second test the shape parameter b has been optimized for different Ewald parameters α. Fundamentally different optimal shape parameters are obtained among the different values of α, see right hand side of Figure 4.6. In the case $\alpha = 1$ we obtain $b_{\mathrm{opt}} \approx 4.75$, as expected based on the first test. □

4.3.2 Analytic differentiation

In order to keep the notation short we define the sums

$$\begin{aligned} z(\boldsymbol{k}, \boldsymbol{x}) &:= \sum_{\boldsymbol{r} \in \mathbb{Z}^3} \boldsymbol{L}^{-\top}(\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}) c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\, \mathrm{e}^{2\pi \mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}}, \\ W(\boldsymbol{k}, \boldsymbol{x}, \boldsymbol{\mu}) &:= \sum_{\boldsymbol{r} \in \mathbb{Z}^3} \cos \gamma_{\boldsymbol{\mu},\boldsymbol{k},\boldsymbol{r}} \, \|\boldsymbol{L}^{-\top}(\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}})\| c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\, \mathrm{e}^{2\pi \mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}}, \end{aligned}$$

as well as

$$\boldsymbol{w}(\boldsymbol{k}, \boldsymbol{x}, \boldsymbol{\mu}) := \sum_{\boldsymbol{r} \in \mathbb{Z}^3} \boldsymbol{L}^{-\top}(\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}}) \cos \gamma_{\boldsymbol{\mu},\boldsymbol{k},\boldsymbol{r}} \, \|\boldsymbol{L}^{-\top}(\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}})\| c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\, \mathrm{e}^{2\pi \mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^{\top}\boldsymbol{L}^{-1}\boldsymbol{x}},$$

where $\gamma_{\boldsymbol{\mu},\boldsymbol{k},\boldsymbol{r}}$ denotes the angle between a vector $\boldsymbol{\mu} \in \mathbb{R}^3$ and $\boldsymbol{L}^{-\top}(\boldsymbol{k} + \boldsymbol{r} \odot \boldsymbol{M}_{\mathrm{o}})$.

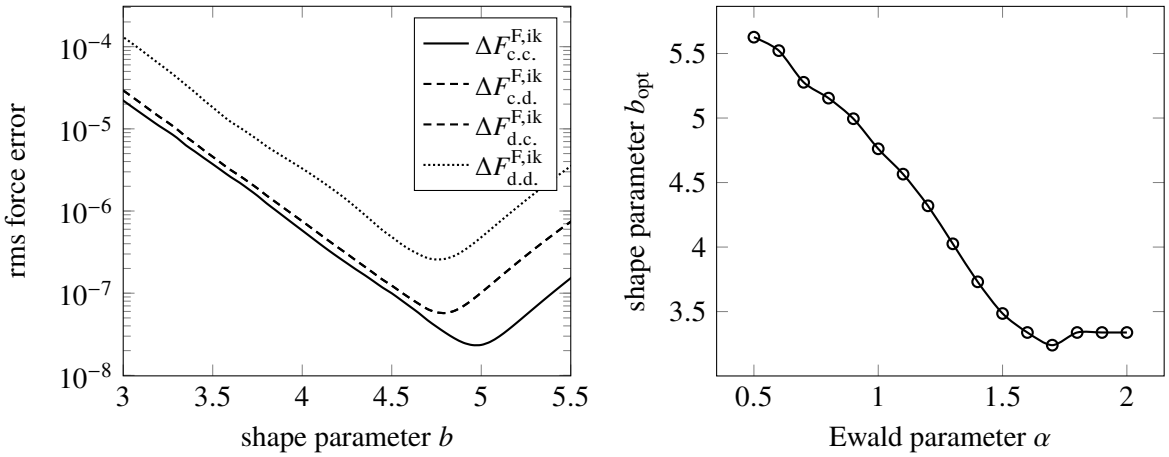


Figure 4.6: Predicted rms force errors (4.63) for the Bessel window. The errors are plotted with respect to the shape parameter b, where we applied the Ewald parameter $\alpha := 1$ (left). The tuned optimal shape parameters for different values of α are plotted on the right hand side.

Therewith we can rewrite the approximated forces as follows

$$\boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.c.}}(j)_{\mathrm{t}} \approx \frac{2\mathrm{i}q_j}{V} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{d}_{\boldsymbol{k}}\hat{\psi}(\boldsymbol{k}) S_{\mathrm{c},\approx}(\boldsymbol{k}) z(\boldsymbol{k}, -\boldsymbol{x}_j) =: \boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.c.}}(j)_{\mathrm{ad}}, \tag{4.66}$$

$$\boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.d.}}(j)_{\mathrm{t}} \approx \frac{2\mathrm{i}q_j}{V} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{d}_{\boldsymbol{k}}\hat{\psi}(\boldsymbol{k}) S_{\mathrm{d,ad}}(\boldsymbol{k}) z(\boldsymbol{k}, -\boldsymbol{x}_j) =: \boldsymbol{F}^{\mathrm{F}}_{\mathrm{c.d.}}(j)_{\mathrm{ad}},$$

$$\boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j)_{\mathrm{t}} \approx \frac{4\pi\|\boldsymbol{\mu}_j\|}{V} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{d}_{\boldsymbol{k}}\hat{\psi}(\boldsymbol{k}) S_{\mathrm{c},\approx}(\boldsymbol{k}) \boldsymbol{w}(\boldsymbol{k}, -\boldsymbol{x}_j, \boldsymbol{\mu}_j) =: \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.c.}}(j)_{\mathrm{ad}},$$

$$\boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j)_{\mathrm{t}} \approx \frac{4\pi\|\boldsymbol{\mu}_j\|}{V} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{d}_{\boldsymbol{k}}\hat{\psi}(\boldsymbol{k}) S_{\mathrm{d,ad}}(\boldsymbol{k}) \boldsymbol{w}(\boldsymbol{k}, -\boldsymbol{x}_j, \boldsymbol{\mu}_j) =: \boldsymbol{F}^{\mathrm{F}}_{\mathrm{d.d.}}(j)_{\mathrm{ad}}. \tag{4.67}$$

The approximations $S_{\mathrm{d,ad}}(\boldsymbol{k})$, obtained via the adjoint gradient NFFT with analytic differentiation, can be rewritten as

$$S_{\mathrm{d,ad}}(\boldsymbol{k}) = 2\pi\mathrm{i}\hat{d}_{\boldsymbol{k}} \sum_{i=N_{\mathrm{c}}+1}^{N} \|\boldsymbol{\mu}_i\| W(\boldsymbol{k}, \boldsymbol{x}_i, \boldsymbol{\mu}_i),$$

based on (3.33).

Remark 4.19 (Undesired self interaction)**.** In Remark 4.14 we discussed the presence of undesired diagonal contributions to the approximation errors caused by applying the NFFT algorithms with i$\boldsymbol{k}$-differentiation. Of course, this applies also in case of analytic differentiation, i.e., we compute undesired self interaction terms in (4.66) and (4.67). In contrast to the i$\boldsymbol{k}$-differentiation approach, it is widely known, cf. [HE88, CDDL09, BCH11], that the analytic differentiation approach is not momentum conserving by construction, even not if symmetric index sets $\mathcal{I}^*_{\boldsymbol{M}}$ are used. In other words, the computed self interactions may influence the accuracy significantly. On the other hand, the analytic differentiation scheme is energy conserving, provided that the truncated NFFT window function is differentiable, cf. [HE88, Sections 5.3.3, 5.5 and 7.6].

The included self forces in (4.66) and (4.67) are given by

$$\boldsymbol{F}^{\mathrm{F},j}_{\mathrm{c.c.}}(j)_{\mathrm{ad}} := \frac{2\mathrm{i}q_j^2}{V} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{d}^2_{\boldsymbol{k}}\hat{\psi}(\boldsymbol{k})Z(\boldsymbol{k},\boldsymbol{x}_j)z(\boldsymbol{k},-\boldsymbol{x}_j)$$

and

$$\boldsymbol{F}^{\mathrm{F},j}_{\mathrm{d.d.}}(j)_{\mathrm{ad}} := \frac{8\pi\mathrm{i}}{V}\|\boldsymbol{\mu}_j\|^2 \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{d}^2_{\boldsymbol{k}}\hat{\psi}(\boldsymbol{k})W(\boldsymbol{k},\boldsymbol{x}_j,\boldsymbol{\mu}_j)\boldsymbol{w}(\boldsymbol{k},-\boldsymbol{x}_j,\boldsymbol{\mu}_j),$$

respectively. The terms for $\pm\boldsymbol{k} \in \mathcal{I}^*_{\boldsymbol{M}}$ do not cancel out by construction, as it is the case for the $\mathrm{i}\boldsymbol{k}$-differentiation approach. We remark that the arising self-interactions in the P³M method have already been studied in [BCH11], where the authors also derive an analytic formula in order to subtract the self forces regarding charge-charge interactions. Furthermore, it is shown that subtracting the self forces may increase the accuracy considerably. □

As in the case of $\mathrm{i}\boldsymbol{k}$-differentiation we initially assume that the diagonal contributions to the overall error are negligible and compute the quantities $\chi^2_{\mathrm{c.c.}}, \chi^2_{\mathrm{c.d.}}, \chi^2_{\mathrm{d.c.}}$ and $\chi^2_{\mathrm{d.d.}}$ as defined via (4.53) and (4.54).

Theorem 4.20. *Let a charge neutral system* (2.5) *with N_{c} charges $q_j \in \mathbb{R}$, $j = 1,\dots,N_{\mathrm{c}}$, and N_{d} dipoles with dipole moments $\boldsymbol{\mu}_j \in \mathbb{R}^3$, $j = N_{\mathrm{c}}+1,\dots,N$, be given. Assume that the truncated Fourier space parts of the forces $\boldsymbol{F}^{\mathrm{F}}_{\mathrm{t}}(j)$ are approximated via the P²NFFT method with oversampled mesh size $\boldsymbol{M}_{\mathrm{o}}$ and a symmetric window function φ. In case of analytic differentiation we obtain*

$$\begin{aligned}
\chi^2_{\mathrm{c.c.}} &= \frac{4}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2\left(\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2(1-2\hat{d}^2_{\boldsymbol{k}}c^2_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})) + \hat{d}^4_{\boldsymbol{k}}Y_0(\boldsymbol{k})Y_2(\boldsymbol{k})\right),\\
\chi^2_{\mathrm{c.d.}} &\le \frac{16\pi^2}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2\left(\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^4(1-2\hat{d}^2_{\boldsymbol{k}}c^2_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})) + \hat{d}^4_{\boldsymbol{k}}Y_2(\boldsymbol{k})Y_2(\boldsymbol{k})\right),\\
\chi^2_{\mathrm{d.c.}} &\le \frac{16\pi^2}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2\left(\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^4(1-2\hat{d}^2_{\boldsymbol{k}}c^2_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})) + \hat{d}^4_{\boldsymbol{k}}Y_0(\boldsymbol{k})Y_4(\boldsymbol{k})\right),\\
\chi^2_{\mathrm{d.d.}} &\le \frac{64\pi^4}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2\left(\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^6(1-2\hat{d}^2_{\boldsymbol{k}}c^2_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})) + \hat{d}^4_{\boldsymbol{k}}Y_2(\boldsymbol{k})Y_4(\boldsymbol{k})\right).
\end{aligned}$$

Proof. With the representation (4.61) of the approximations $S_{\mathrm{c},\approx}(\boldsymbol{k})$ we obtain

$$\begin{aligned}
\chi^{\mathrm{c.c.}}_{ij} &= \frac{2\mathrm{i}}{V}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})\left((\boldsymbol{L}^{-\top}\boldsymbol{k})\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}} - \hat{d}^2_{\boldsymbol{k}}Z(\boldsymbol{k},\boldsymbol{x}_i)z(\boldsymbol{k},-\boldsymbol{x}_j)\right)\\
&= \frac{2\mathrm{i}}{V}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})\left((\boldsymbol{L}^{-\top}\boldsymbol{k})\,\mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_j} - \hat{d}^2_{\boldsymbol{k}}c_{\boldsymbol{k}}(\tilde{\varphi}_{\mathrm{t}})z(\boldsymbol{k},-\boldsymbol{x}_j)\right)\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_i}\\
&\quad - \frac{2\mathrm{i}}{V}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})\hat{d}^2_{\boldsymbol{k}}z(\boldsymbol{k},-\boldsymbol{x}_j)\sum_{\boldsymbol{r}\in\mathbb{Z}^3\setminus\{\boldsymbol{0}\}} c_{\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}}}(\tilde{\varphi}_{\mathrm{t}})\,\mathrm{e}^{2\pi\mathrm{i}(\boldsymbol{k}+\boldsymbol{r}\odot\boldsymbol{M}_{\mathrm{o}})^\top\boldsymbol{L}^{-1}\boldsymbol{x}_i}.
\end{aligned}$$

Integration with respect to $\boldsymbol{x}_i$ yields

$$\frac{1}{V}\int_{\boldsymbol{L}\mathbb{T}} \|\chi_{ij}^{\text{c.c.}}\|^2 \mathrm{d}\boldsymbol{x}_i = \frac{4}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2 \left\| (\boldsymbol{L}^{-\top}\boldsymbol{k})\, \mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_j} - \hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t}) z(\boldsymbol{k}, -\boldsymbol{x}_j)\right\|^2 + \frac{4}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2 \hat{d}_{\boldsymbol{k}}^4 \|z(\boldsymbol{k}, -\boldsymbol{x}_j)\|^2 \tilde{Y}_0(\boldsymbol{k}),$$

where we use Parseval's identity. Integration with respect to $\boldsymbol{x}_j$ gives

$$\begin{aligned}\chi_{\text{c.c.}}^2 &= \frac{4}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2 \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 (1 - \hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}^2(\tilde{\varphi}_\mathrm{t}))^2 + \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2 \hat{d}_{\boldsymbol{k}}^4 \left(c_{\boldsymbol{k}}^2(\tilde{\varphi}_\mathrm{t}) \tilde{Y}_2(\boldsymbol{k}) + \tilde{Y}_0(\boldsymbol{k}) Y_2(\boldsymbol{k})\right) \\ &= \frac{4}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2 \left(\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 (1 - 2\hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}^2(\tilde{\varphi}_\mathrm{t})) + \hat{d}_{\boldsymbol{k}}^4 Y_0(\boldsymbol{k}) Y_2(\boldsymbol{k})\right).\end{aligned}$$

In the other cases we obtain

$$\begin{aligned}\chi_{ij}^{\text{c.d.}} &= -\frac{4\pi}{V}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k}) \left((\boldsymbol{L}^{-\top}\boldsymbol{k}) \|\boldsymbol{L}^{-\top}\boldsymbol{k}\| \cos\gamma_{i,\boldsymbol{k}}\, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}} - \hat{d}_{\boldsymbol{k}}^2 W(\boldsymbol{k}, \boldsymbol{x}_i, \boldsymbol{\mu}_i) z(\boldsymbol{k}, -\boldsymbol{x}_j)\right), \\ \chi_{ij}^{\text{d.c.}} &= \frac{4\pi}{V}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k}) \left((\boldsymbol{L}^{-\top}\boldsymbol{k}) \|\boldsymbol{L}^{-\top}\boldsymbol{k}\| \cos\gamma_{j,\boldsymbol{k}}\, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}} - \hat{d}_{\boldsymbol{k}}^2 Z(\boldsymbol{k}, \boldsymbol{x}_i) \boldsymbol{w}(\boldsymbol{k}, -\boldsymbol{x}_j, \boldsymbol{\mu}_j)\right), \\ \chi_{ij}^{\text{d.d.}} &= \frac{8\pi^2\mathrm{i}}{V^2}\sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k}) \Big((\boldsymbol{L}^{-\top}\boldsymbol{k}) \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 \cos\gamma_{i,\boldsymbol{k}} \cos\gamma_{j,\boldsymbol{k}}\, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}^\top\boldsymbol{L}^{-1}\boldsymbol{x}_{ij}} \\ &\qquad - \hat{d}_{\boldsymbol{k}}^2 W(\boldsymbol{k}, \boldsymbol{x}_i, \boldsymbol{\mu}_i) \boldsymbol{w}(\boldsymbol{k}, -\boldsymbol{x}_j, \boldsymbol{\mu}_j)\Big).\end{aligned}$$

Using Parseval's identity two times, analogously to the charge-charge case, and estimating the squared cosine terms from above by 1 completes the proof. ■

Optimizing the above stated estimates with respect to the deconvolution coefficients $\hat{d}_{\boldsymbol{k}}$ yields four different results. As an example, in case of charge-charge interactions we obtain the optimal coefficients

$$\hat{d}_{\boldsymbol{k}} = \pm\frac{\|\boldsymbol{L}^{-\top}\boldsymbol{k}\| c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t})}{\sqrt{Y_0(\boldsymbol{k}) Y_2(\boldsymbol{k})}}.$$

If the Fourier coefficients $c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t})$ tend to zero sufficiently fast for growing $\boldsymbol{k}$, we have $Y_0(\boldsymbol{k}) \approx c_{\boldsymbol{k}}^2(\tilde{\varphi}_\mathrm{t})$, $Y_2(\boldsymbol{k}) \approx \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 c_{\boldsymbol{k}}^2(\tilde{\varphi}_\mathrm{t})$ and thus

$$\frac{\|\boldsymbol{L}^{-\top}\boldsymbol{k}\| c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t})}{\sqrt{Y_0(\boldsymbol{k}) Y_2(\boldsymbol{k})}} \approx \frac{1}{c_{\boldsymbol{k}}(\tilde{\varphi}_\mathrm{t})},$$

which are the standard deconvolution coefficients. This applies also in case of charge-dipole, dipole-charge and dipole-dipole interactions.

Efficient computation of the resulting rms errors

We sketch an approach how to efficiently estimate the rms force errors in case of analytic differentiation. Since the presented estimates are only valid for differentiable window functions, we restrict our

considerations to the B-spline window as well as, for simplicity, to orthorhombic box shapes with $\boldsymbol{L} = \operatorname{diag}(L_1, L_2, L_3)$.

In the following we set

$$y_n(k) := \sum_{r\in\mathbb{Z}} (k + r\sigma M)^n \frac{c^2_{k+r\sigma \boldsymbol{M}}(\tilde{\varphi})}{c^2_k(\tilde{\varphi})},$$

where we have given upper bounds of the form

$$y_n(k) \le k^n + s_n(k).$$

For the B-spline window we have

$$s_0(k) = \frac{8m}{4m-1}\left(\frac{|k|}{|k| - \sigma M}\right)^{4m},$$

see (3.38), as well as

$$s_2(k) = \frac{8m-4}{4m-3}\frac{k^{4m}}{(|k| - \sigma M)^{4m-2}},$$

as presented in Example 3.19, and

$$s_4(k) = \frac{8m-8}{4m-5}\frac{k^{4m}}{(|k| - \sigma M)^{4m-4}},$$

which is obtained analogously.

By making use of the terms $y_0(k)$ and $y_2(k)$ we can write

$$\frac{Y_0(\boldsymbol{k})}{c^2_{\boldsymbol{k}}(\tilde{\varphi})} = y_0(k_1)y_0(k_2)y_0(k_3)$$

and

$$\frac{Y_2(\boldsymbol{k})}{c^2_{\boldsymbol{k}}(\tilde{\varphi})} = \frac{y_2(k_1)y_0(k_2)y_0(k_3)}{L_1^2} + \frac{y_0(k_1)y_2(k_2)y_0(k_3)}{L_2^2} + \frac{y_0(k_1)y_0(k_2)y_2(k_3)}{L_3^2}.$$

In case of the standard deconvolution approach we set $\hat{d}_{\boldsymbol{k}} := c^{-1}_{\boldsymbol{k}}(\tilde{\varphi})$ and obtain

$$\chi^2_{\text{c.c.}} = \frac{4}{V^2}\sum_{\boldsymbol{k}\in I_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2 \left(\frac{Y_0(\boldsymbol{k})Y_2(\boldsymbol{k})}{c^4_{\boldsymbol{k}}(\tilde{\varphi})} - \frac{k_1^2}{L_1^2} - \frac{k_2^2}{L_2^2} - \frac{k_3^2}{L_3^2}\right).$$

Note that we have $Y_0(\boldsymbol{k}) = c^2_{\boldsymbol{k}}(\tilde{\varphi}) + \tilde{Y}_0(\boldsymbol{k})$ as well as $Y_2(\boldsymbol{k}) = \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 c^2_{\boldsymbol{k}}(\tilde{\varphi}) + \tilde{Y}_2(\boldsymbol{k})$ and, thus, the term $\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 = {}^{k_1^2}/_{L_1^2} + {}^{k_2^2}/_{L_2^2} + {}^{k_3^2}/_{L_3^2}$ finally cancels out.

Similarly to the i$\boldsymbol{k}$-differentiation approach we use the approximation (4.65) with $n = 11$ terms and obtain, via differentiation with respect to x,

$$\frac{1}{x^2} \approx \sum_{i=1}^{n} r_i\omega_i\, \mathrm{e}^{-\omega_i x}.$$

Again, we set $x := L_{\max}^2 \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2$ in each summand, which now gives

$$\hat{\psi}(\boldsymbol{k})^2 = \frac{\mathrm{e}^{-2\pi^2\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2/\alpha^2}}{\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^4} \approx L_{\max}^4\,\mathrm{e}^{-2\pi^2\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2/\alpha^2} \sum_{i=1}^{n} r_i\omega_i\,\mathrm{e}^{-\omega_i L_{\max}^2\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2},$$

where the three dimensions are separated due to the assumed orthorhombic box shape. Note that a similar idea has been applied in Example 3.17.

By defining the sums

$$R_{i,j}^{(p)} := \sum_{k=-M/2}^{M/2-1} \left(\frac{k}{L_j}\right)^p \mathrm{e}^{-2\pi^2k^2/(\alpha^2L_j^2)-\omega_i k^2 L_{\max}^2/L_j^2},$$

$$S_{i,j}^{(p,q)} := \frac{1}{L_j^{p+q}} \sum_{k=-M/2}^{M/2-1} \left(k^p s_q(k) + k^q s_p(k) + s_p(k)s_q(k)\right) \mathrm{e}^{-2\pi^2k^2/(\alpha^2L_j^2)-\omega_i k^2 L_{\max}^2/L_j^2},$$

and

$$S_{i,j}^{(p)} := S_{i,j}^{(p,p)},$$

we end up with an approximation of the form

$$\chi_{\mathrm{c.c.}}^2 \approx \frac{4L_{\max}^2}{V^2} \sum_{i=1}^{n} r_i\omega_i O_i, \tag{4.68}$$

where

$$\begin{aligned} O_i := {} & \left(S_{i,1}^{(0)} + R_{i,1}^{(0)}\right)\left(S_{i,2}^{(0)} + R_{i,2}^{(0)}\right)S_{i,3}^{(0,2)} + \left(S_{i,1}^{(0)}S_{i,2}^{(0)} + R_{i,1}^{(0)}S_{i,2}^{(0)} + S_{i,1}^{(0)}R_{i,2}^{(0)}\right)R_{i,3}^{(2)} \\ & + \left(S_{i,1}^{(0)} + R_{i,1}^{(0)}\right)\left(S_{i,3}^{(0)} + R_{i,3}^{(0)}\right)S_{i,2}^{(0,2)} + \left(S_{i,1}^{(0)}S_{i,3}^{(0)} + R_{i,1}^{(0)}S_{i,3}^{(0)} + S_{i,1}^{(0)}R_{i,3}^{(0)}\right)R_{i,2}^{(2)} \\ & + \left(S_{i,2}^{(0)} + R_{i,2}^{(0)}\right)\left(S_{i,3}^{(0)} + R_{i,3}^{(0)}\right)S_{i,1}^{(0,2)} + \left(S_{i,2}^{(0)}S_{i,3}^{(0)} + R_{i,2}^{(0)}S_{i,3}^{(0)} + S_{i,2}^{(0)}R_{i,3}^{(0)}\right)R_{i,1}^{(2)}. \end{aligned}$$

This is verified by only a few calculation steps.

In order to keep the notation short we introduce the following abbreviated form, where we set

$$\left[a_{j_1}b_{j_2}c_{j_3}\right]_{1,2,3} := a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2.$$

Based on that, we can rewrite O_i in (4.68) as

$$O_i = \left[S_{i,j_1}^{(0,2)}\left(S_{i,j_2}^{(0)} + R_{i,j_2}^{(0)}\right)\left(S_{i,j_3}^{(0)} + R_{i,j_3}^{(0)}\right) + R_{i,j_1}^{(2)}\left(S_{i,j_2}^{(0)}S_{i,j_3}^{(0)} + R_{i,j_2}^{(0)}S_{i,j_3}^{(0)} + S_{i,j_2}^{(0)}R_{i,j_3}^{(0)}\right)\right]_{1,2,3},$$

where the terms $R_{i,j_1}^{(2)}R_{i,j_2}^{(0)}R_{i,j_3}^{(0)}$ have been canceled out by $-\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2$.

Concerning the interactions from charges with dipoles, we obtain in case of applying the standard deconvolution approach

$$\chi_{\mathrm{c.d.}}^2 \le \frac{16\pi^2}{V^2} \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2 \left(\frac{Y_2^2(\boldsymbol{k})}{c_{\boldsymbol{k}}^4(\tilde{\varphi})} - \left(\frac{k_1^2}{L_1^2} - \frac{k_2^2}{L_2^2} - \frac{k_3^2}{L_3^2}\right)^2\right).$$

In an analogue manner, we obtain the following approximation.

$$\chi^2_{\text{c.d.}} \approx \frac{16\pi^2 L_{\max}^2}{V^2} \sum_{i=1}^{n} r_i \omega_i O_i,$$

where

$$\begin{aligned} O_i := & \left[\left(R^{(4)}_{i,j_1} + S^{(2)}_{i,j_1} \right) \left(R^{(0)}_{i,j_2} + S^{(0)}_{i,j_2} \right) \left(R^{(0)}_{i,j_3} + S^{(0)}_{i,j_3} \right) - R^{(4)}_{i,j_1} R^{(0)}_{i,j_2} R^{(0)}_{i,j_3} \right]_{1,2,3} \\ & + 2 \left[\left(R^{(0)}_{i,j_1} + S^{(0)}_{i,j_1} \right) \left(R^{(2)}_{i,j_2} + S^{(0,2)}_{i,j_2} \right) \left(R^{(2)}_{i,j_3} + S^{(0,2)}_{i,j_3} \right) - R^{(0)}_{i,j_1} R^{(2)}_{i,j_2} R^{(2)}_{i,j_3} \right]_{1,2,3}. \end{aligned}$$

The terms $R^{(4)}_{i,j_1} R^{(0)}_{i,j_2} R^{(0)}_{i,j_3}$ and $R^{(0)}_{i,j_1} R^{(2)}_{i,j_2} R^{(2)}_{i,j_3}$ represent the values $-\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2$, which are excluded. Note that we skip expanding the bracket terms, for overview purposes.

Furthermore, we have

$$\chi^2_{\text{d.c.}} \leq \frac{16\pi^2}{V^2} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2 \left(\frac{Y_0(\boldsymbol{k}) Y_4(\boldsymbol{k})}{c^4_{\boldsymbol{k}}(\tilde{\varphi})} - \left(\frac{k_1^2}{L_1^2} - \frac{k_2^2}{L_2^2} - \frac{k_3^2}{L_3^2} \right)^2 \right),$$

where

$$\begin{aligned} \frac{Y_4(\boldsymbol{k})}{c^2_{\boldsymbol{k}}(\tilde{\varphi})} = & \frac{y_4(k_1) y_0(k_1) y_0(k_1)}{L_1^4} + \frac{y_0(k_1) y_4(k_1) y_0(k_1)}{L_2^4} + \frac{y_0(k_1) y_0(k_1) y_4(k_1)}{L_3^4} \\ & + \frac{2 y_0(k_1) y_2(k_2) y_2(k_3)}{L_2^2 L_3^2} + \frac{2 y_2(k_1) y_0(k_2) y_2(k_3)}{L_1^2 L_3^2} + \frac{2 y_2(k_1) y_2(k_2) y_0(k_3)}{L_1^2 L_2^2}. \end{aligned}$$

Thus, we obtain

$$\chi^2_{\text{d.c.}} \approx \frac{16\pi^2 L_{\max}^2}{V^2} \sum_{i=1}^{n} r_i \omega_i O_i$$

with

$$\begin{aligned} O_i := & \left[\left(R^{(4)}_{i,j_1} + S^{(0,4)}_{i,j_1} \right) \left(R^{(0)}_{i,j_2} + S^{(0)}_{i,j_2} \right) \left(R^{(0)}_{i,j_3} + S^{(0)}_{i,j_3} \right) - R^{(4)}_{i,j_1} R^{(0)}_{i,j_2} R^{(0)}_{i,j_3} \right]_{1,2,3} \\ & + 2 \left[\left(R^{(0)}_{i,j_1} + S^{(0)}_{i,j_1} \right) \left(R^{(2)}_{i,j_2} + S^{(0,2)}_{i,j_2} \right) \left(R^{(2)}_{i,j_3} + S^{(0,2)}_{i,j_3} \right) - R^{(0)}_{i,j_1} R^{(2)}_{i,j_2} R^{(2)}_{i,j_3} \right]_{1,2,3}. \end{aligned}$$

For the dipole-dipole part we have

$$\chi^2_{\text{d.d.}} \leq \frac{64\pi^4}{V^2} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \hat{\psi}(\boldsymbol{k})^2 \left(\frac{Y_2(\boldsymbol{k}) Y_4(\boldsymbol{k})}{c^4_{\boldsymbol{k}}(\tilde{\varphi})} - \left(\frac{k_1^2}{L_1^2} - \frac{k_2^2}{L_2^2} - \frac{k_3^2}{L_3^2} \right)^3 \right).$$

Now, we end up with,

$$\chi^2_{\text{d.d.}} \approx \frac{64\pi^4 L_{\max}^2}{V^2} \sum_{i=1}^{n} r_i \omega_i O_i$$

where

$$
\begin{aligned}
O_i = & \left[\left(R^{(6)}_{i,j_1} + S^{(2,4)}_{i,j_1}\right)\left(R^{(0)}_{i,j_2} + S^{(0)}_{i,j_2}\right)\left(R^{(0)}_{i,j_3} + S^{(0)}_{i,j_3}\right) - R^{(6)}_{i,j_1} R^{(0)}_{i,j_2} R^{(0)}_{i,j_3}\right]_{1,2,3} \\
& + \left[\left(R^{(2)}_{i,j_1} + S^{(0,2)}_{i,j_1}\right)\left(R^{(4)}_{i,j_2} + S^{(0,4)}_{i,j_2}\right)\left(R^{(0)}_{i,j_3} + S^{(0)}_{i,j_3}\right) - R^{(2)}_{i,j_1} R^{(4)}_{i,j_2} R^{(0)}_{i,j_3}\right]_{1,2,3} \\
& + \left[\left(R^{(2)}_{i,j_1} + S^{(0,2)}_{i,j_1}\right)\left(R^{(0)}_{i,j_2} + S^{(0)}_{i,j_2}\right)\left(R^{(4)}_{i,j_3} + S^{(0,4)}_{i,j_3}\right) - R^{(2)}_{i,j_1} R^{(0)}_{i,j_2} R^{(4)}_{i,j_3}\right]_{1,2,3} \\
& + 2\left[\left(R^{(2)}_{i,j_1} + S^{(0,2)}_{i,j_1}\right)\left(R^{(2)}_{i,j_2} + S^{(0,2)}_{i,j_2}\right)\left(R^{(2)}_{i,j_3} + S^{(0,2)}_{i,j_3}\right) - R^{(2)}_{i,j_1} R^{(2)}_{i,j_2} R^{(2)}_{i,j_3}\right]_{1,2,3} \\
& + 2\left[\left(R^{(4)}_{i,j_1} + S^{(2)}_{i,j_1}\right)\left(R^{(2)}_{i,j_2} + S^{(0,2)}_{i,j_2}\right)\left(R^{(0)}_{i,j_3} + S^{(0)}_{i,j_3}\right) - R^{(4)}_{i,j_1} R^{(2)}_{i,j_2} R^{(0)}_{i,j_3}\right]_{1,2,3} \\
& + 2\left[\left(R^{(4)}_{i,j_1} + S^{(2)}_{i,j_1}\right)\left(R^{(0)}_{i,j_2} + S^{(0)}_{i,j_2}\right)\left(R^{(2)}_{i,j_3} + S^{(0,2)}_{i,j_3}\right) - R^{(4)}_{i,j_1} R^{(0)}_{i,j_2} R^{(2)}_{i,j_3}\right]_{1,2,3}.
\end{aligned}
$$

Again, the subtracted terms represent the values $-\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^6$, which are excluded.

5 Numerical results

In this chapter we present several numerical results demonstrating the functional suitability as well as the efficiency of the P^{2}NFFT method for mixed charge-dipole systems, see Algorithm 4.3. Furthermore, the presented error estimates are verified based on numerical examples.

As already discussed in the previous sections, there are two types of errors arising in our P^{2}NFFT method. Fist, truncating the Ewald sums introduces cutoff or rather truncation errors, see Section 2.3. The discussed error estimates are going to be validated based on numerical examples in Section 5.1, where we apply the P^{2}NDFT Ewald method, see Remark 4.12, to different particle systems.

Secondly, the application of the different NFFT modules introduces further approximation errors, see Section 4.3. Numerical results regarding the fast and approximative P^{2}NFFT method are going to be considered in Section 5.2, where we verify the correctness of the presented error estimates for 3d-periodic boundary conditions, see Section 5.2.1, and discuss how all involved parameters may be tuned automatically, see Section 5.2.2. Some results for mixed periodic and open boundary conditions are presented in Section 5.3.

According to Section 2.3.3, the mesh size $\boldsymbol{M} = (M_1, M_2, M_3)$ in Fourier space is chosen such that (2.82) is approximately fulfilled for some cutoff parameter $\beta > 0$. Since we are restricted to $\boldsymbol{M} \in 2\mathbb{N}^3$, we set

$$M_j := 2\left\lceil \frac{\beta \|\boldsymbol{\ell}_j\|}{2} \right\rceil, \quad j \in \{1, 2, 3\}, \tag{5.1}$$

i.e., we round up to the next even integer, if necessary.

5.1 Ewald truncation errors

In order to verify the estimates discussed in Section 2.3, we applied the P^{2}NDFT Ewald method to different particle systems and compared the computed estimates with the measured errors.

As indicated by Remark 4.1, the present implementation of the P^{2}NDFT Ewald method causes the computation of small, undesired self interaction terms in the Fourier space parts of the forces in case of charge-charge as well as dipole-dipole interactions. This is due to the usage of non symmetric index sets $\mathcal{I}_{\boldsymbol{M}}$, as defined in (3.2). In order to get an impression of how much the approximation errors are influenced by these computed self interactions, we applied the P^{2}NDFT method to the following two particle systems. The results are presented in Example 5.1.

Particle system 1.

- Charges: $N_\mathrm{c} = 300$, $q_j = (-1)^j$.
- Positions: $\boldsymbol{x}_j$ are uniformly distributed in a box spanned by $\boldsymbol{\ell}_1 := (10, 0, 0)^\top$, $\boldsymbol{\ell}_2 := (0, 10, 0)^\top$ and $\boldsymbol{\ell}_3 := (0, 0, 10)^\top$. The volume of the primary box is $V = 10^3$.

Particle system 2.

- Dipoles: $N_\mathrm{d} = 300$, randomly oriented dipole moments with $\|\boldsymbol{\mu}_j\| = 1$.
- Positions: $\boldsymbol{x}_j$ are uniformly distributed in a box spanned by $\boldsymbol{\ell}_1 := (10, 0, 0)^\top$, $\boldsymbol{\ell}_2 := (0, 10, 0)^\top$ and $\boldsymbol{\ell}_3 := (0, 0, 10)^\top$. The volume of the primary box is $V = 10^3$.

Example 5.1 (Undesired self interactions)**.** We consider the Particle systems 1 and 2 and pay attention to the self interactions in the forces originating from using non symmetric index sets $\mathcal{I}_{\boldsymbol{M}}$, cf. Remark 4.1. Particle system 1 is a pure charge system, whereas only dipole-dipole interactions are present in Particle system 2. We suppose the self interaction terms to be negligible, since the supporting index set $\mathcal{I}_{-\boldsymbol{M}} \setminus \mathcal{I}_{\boldsymbol{M}}^*$ is rather small, see (4.12) and (4.13).

In order to verify our supposition, we investigated whether the actual truncation errors are predicted properly based on the estimates (2.88) and (2.92). For this purpose, we applied the P^{2}NDFT Ewald method with different parameter sets to the two considered particle systems. For the computation of reference data we set

$$\beta := 6.4 \quad \Longleftrightarrow \quad \boldsymbol{M} := (64, 64, 64)$$

and applied the P^{2}NDFT method, where we used the symmetric index set $\mathcal{I}_{(64,64,64)}^*$ and various splitting parameters α to approximate the Fourier space parts.

In order to achieve a lower accuracy, we applied the P^{2}NDFT method by using smaller cutoff parameters $\beta \in \{0.8, 1.6, 3.2\}$ for varying splitting parameter α. First, we applied the resulting non symmetric index sets $\mathcal{I}_{\boldsymbol{M}}$. Then, each test run was repeated, whereby the non symmetric index set $\mathcal{I}_{\boldsymbol{M}}$ was replaced by the corresponding symmetric index set $\mathcal{I}_{\boldsymbol{M}}^*$. The measured approximation errors are plotted in Figure 5.1, where also the corresponding estimates (2.88) and (2.92) are depicted. It can be seen that the usage of an index set $\mathcal{I}_{\boldsymbol{M}}$ does not introduce recognizable self interactions. The measured errors are even somewhat smaller compared to the errors obtained with the reduced symmetric index sets $\mathcal{I}_{\boldsymbol{M}}^*$. Furthermore, the actual errors are properly estimated from above. In case of relatively small cutoff parameters β or large splitting parameters α the estimates might not be valid, as analyzed in the proofs of Theorems 2.23 and 2.24. Here, the actual errors exceed the estimates in case of Particle system 2 for $\beta = 0.8$ and relatively large α.

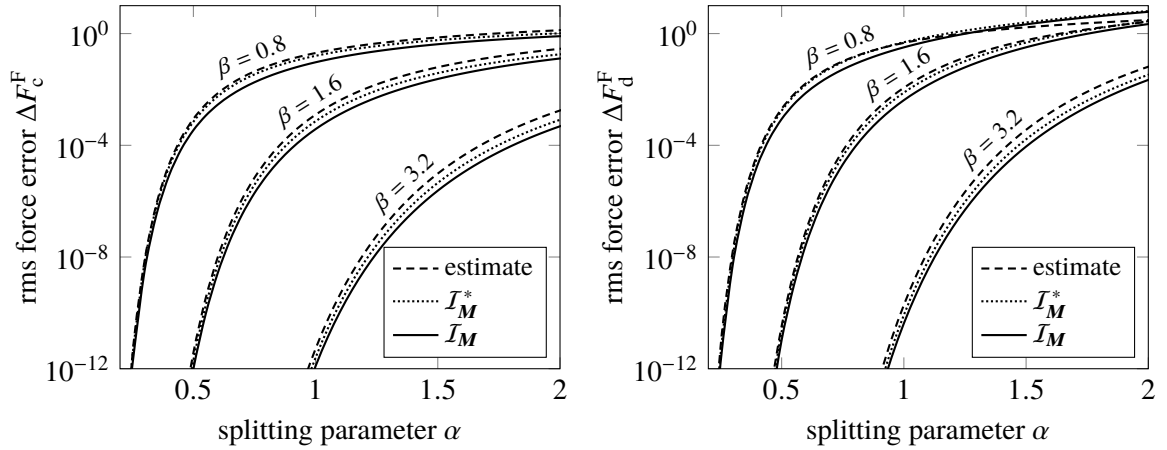


Figure 5.1: Estimated rms force errors in the Fourier space parts of the forces, see (2.88) and (2.92). We also plot the measured errors obtained by using the index sets $\mathcal{I}_{\boldsymbol{M}}$ as well as $\mathcal{I}_{\boldsymbol{M}}^*$. We consider the Particle systems 1 (left) and 2 (right).

□

The following examples are dedicated to the verification of the Ewald truncation errors for mixed charge-dipole systems. For this purpose, we introduce the Particle systems 3 and 4, each containing a mixture of charges and dipoles. Note that the particles are distributed in an orthorhombic box in Particle system 3, whereas a triclinic box shape is given in Particle system 4.

Particle system 3.

- Charges: $N_c = 300$, $q_j = (-1)^j$.
- Dipoles: $N_d = 300$, randomly oriented dipole moments with $\|\boldsymbol{\mu}_j\| = 1$.
- Positions: $\boldsymbol{x}_j$ are uniformly distributed in a box spanned by $\boldsymbol{\ell}_1 := (20, 0, 0)^\top$, $\boldsymbol{\ell}_2 := (0, 10, 0)^\top$ and $\boldsymbol{\ell}_3 := (0, 0, 10)^\top$. The volume of the primary box is $V = 2 \cdot 10^3$.

Particle system 4.

- Charges: $N_c = 300$, $q_j = (-1)^j$.
- Dipoles: $N_d = 300$, randomly oriented dipole moments with $\|\boldsymbol{\mu}_j\| = 1$.
- Positions: $\boldsymbol{x}_j$ are uniformly distributed in a box spanned by $\boldsymbol{\ell}_1 := (10, 5, 0)^\top$, $\boldsymbol{\ell}_2 := (0, 20, 10)^\top$ and $\boldsymbol{\ell}_3 := (20, 0, 5)^\top$. The volume of the primary box is $V = 2 \cdot 10^3$.

In order to estimate the overall rms force errors in the short range and the Fourier space parts, respectively, we assume for simplicity that the contributions originating from interactions with charges and those resulting from interactions with dipoles are independent.

In other words, we estimate the rms force errors regarding the short range parts via

$$\Delta F^{\text{short}}_{\text{c,est}} := \sqrt{\left(\Delta F^{\text{short}}_{\text{c.c.}}\right)^2 + \left(\Delta F^{\text{short}}_{\text{c.d.}}\right)^2},$$
$$\Delta F^{\text{short}}_{\text{d,est}} := \sqrt{\left(\Delta F^{\text{short}}_{\text{d.c.}}\right)^2 + \left(\Delta F^{\text{short}}_{\text{d.d.}}\right)^2},$$

and in the Fourier space parts via

$$\Delta F^{\text{F}}_{\text{c,est}} := \sqrt{\left(\Delta F^{\text{F}}_{\text{c.c.}}\right)^2 + \left(\Delta F^{\text{F}}_{\text{c.d.}}\right)^2},$$
$$\Delta F^{\text{F}}_{\text{d,est}} := \sqrt{\left(\Delta F^{\text{F}}_{\text{d.c.}}\right)^2 + \left(\Delta F^{\text{F}}_{\text{d.d.}}\right)^2}.$$

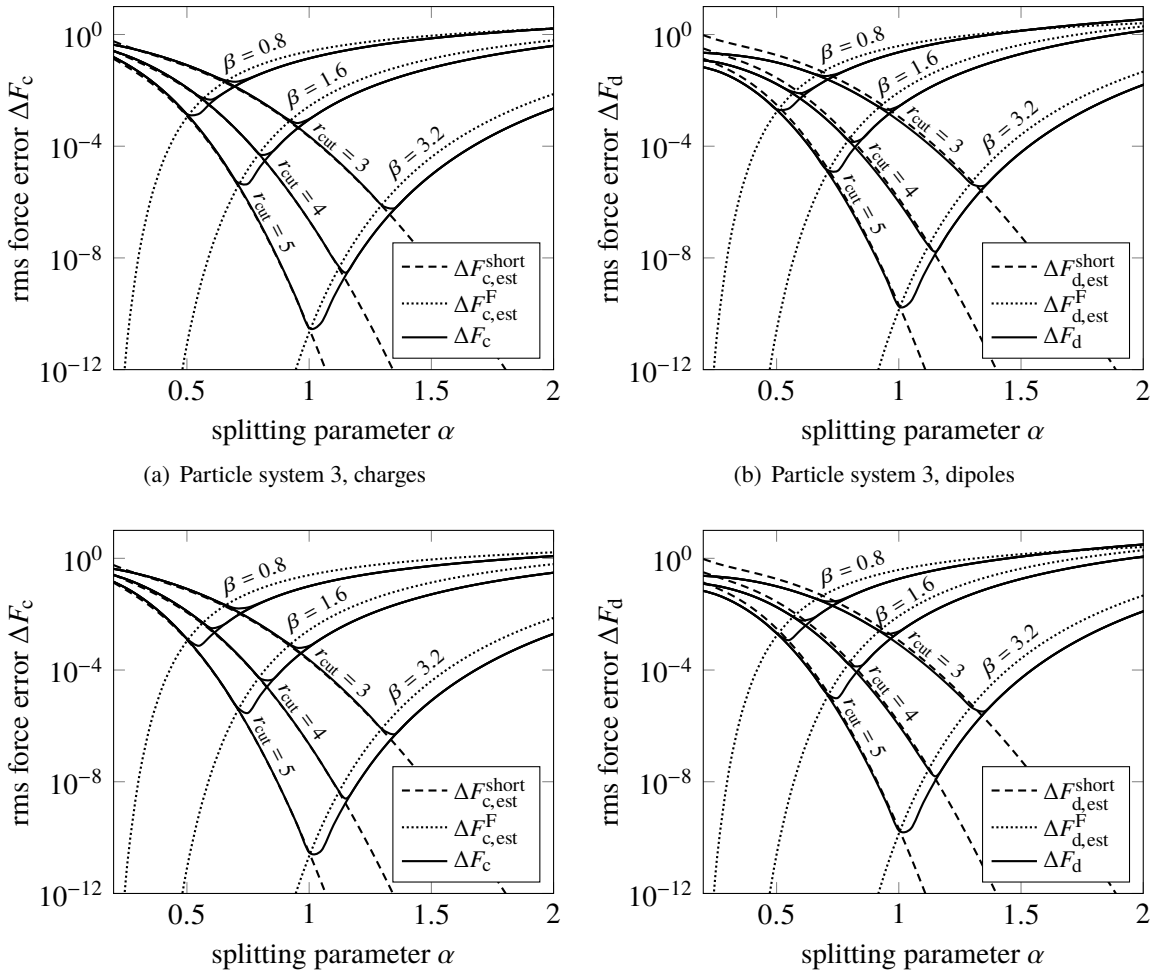
The single contributions to the overall errors have been studied in Sections 2.3.2 and 2.3.3.

Example 5.2. We consider the Particle systems 3 and 4. In order to compute highly accurate reference data we fixed the near field cutoff radius $r_{\text{cut}} := 5$ and tuned the splitting parameter α as well as the far field cutoff β as described in Section 2.3.4, where we applied the basic accuracy $\varepsilon := 10^{-16}$. Note that the same parameter constellation has been considered in Example 2.26. Based on the obtained parameters, highly accurate reference data have been computed via the P²NDFT Ewald method.

In order to verify the derived error estimates we applied the P²NDFT Ewald method by using different near field cutoff radii $r_{\text{cut}} \in \{3, 4, 5\}$ in combination with far field cutoffs $\beta \in \{0.8, 1.6, 3.2\}$. The resulting overall approximation errors have been computed based on the reference data. In Figure 5.2 we summarize the obtained measured errors and the presented error estimates, see Section 2.3, with respect to the splitting parameter α. We achieve a very good concordance between the estimates and the measured errors. □

5.2 P²NFFT under 3d-periodic constraints

In this section we set the focus on the resulting approximation errors caused by approximating the truncated Fourier space parts based on the introduced NFFT modules. We approximated the acting forces in different particle systems via the P²NFFT method, where we applied various parameter sets.



(a) Particle system 3, charges

(b) Particle system 3, dipoles

(c) Particle system 4, charges

(d) Particle system 4, dipoles

Figure 5.2: Estimated errors $\Delta F^{\text{short}}_{\text{c/d,est}}$ and $\Delta F^{\text{F}}_{\text{c/d,est}}$ as well as measured errors $\Delta F_{\text{c/d}}$ for the Particle systems 3 and 4. We applied different near field as well as far field cutoffs (see labels).

In order to determine the actual approximation errors, we also applied the P^{2}NDFT Ewald method for each considered parameter combination. The results of the described experiments are summarized in Section 5.2.1.

Note that the PNFFT sotware allows the usage of different oversampling factors σ_j and shape parameters b_j in the single dimensions. However, in our numerical tests we always used the same oversampling factor σ as well as the same shape parameter b in all three dimensions. The oversampled mesh size $\boldsymbol{M}_{\text{o}} \in 2\mathbb{N}^3$ is defined via

$$M_{\text{o},j} := 2\left\lceil \frac{\sigma M_j}{2} \right\rceil, \quad j \in \{1,2,3\},$$

i.e., we round up to the next even integer, if necessary.

5.2.1 Verification of the error estimates

The validity of the error estimates, presented in Section 4.3, has been verified based on several numerical experiments, which are summarized in the following.

We start with the i$\boldsymbol{k}$-differentiation approach, as considered in Section 4.3.1. Assuming that the contributions originating from interactions with charges and those resulting from interactions with dipoles are independent, we may approximate the overall rms errors in the Fourier space parts of the forces via

$$\Delta F_{\mathrm{c}}^{\mathrm{F,ik}} \approx \sqrt{\left(\Delta F_{\mathrm{c.c.}}^{\mathrm{F,ik}}\right)^2 + \left(\Delta F_{\mathrm{c.d.}}^{\mathrm{F,ik}}\right)^2}$$

and

$$\Delta F_{\mathrm{d}}^{\mathrm{F,ik}} \approx \sqrt{\left(\Delta F_{\mathrm{d.c.}}^{\mathrm{F,ik}}\right)^2 + \left(\Delta F_{\mathrm{d.d.}}^{\mathrm{F,ik}}\right)^2}. \tag{5.2}$$

For the definition of the estimates $\Delta F_{\mathrm{c.c.}}^{\mathrm{F,ik}}$, $\Delta F_{\mathrm{c.d.}}^{\mathrm{F,ik}}$, $\Delta F_{\mathrm{d.c.}}^{\mathrm{F,ik}}$ and $\Delta F_{\mathrm{d.d.}}^{\mathrm{F,ik}}$ we refer to Theorem 4.13 and (4.63).

In the following examples we compare the predicted errors, as denoted above, with the measured errors, which are determined by applying the P²NDFT Ewald and the P²NFFT methods to the same particle systems.

Example 5.3. We consider Particle system 3 and the mesh sizes

$$\boldsymbol{M} = (32, 16, 16) \iff \beta = 1.6$$

and

$$\boldsymbol{M} = (64, 32, 32) \iff \beta = 3.2.$$

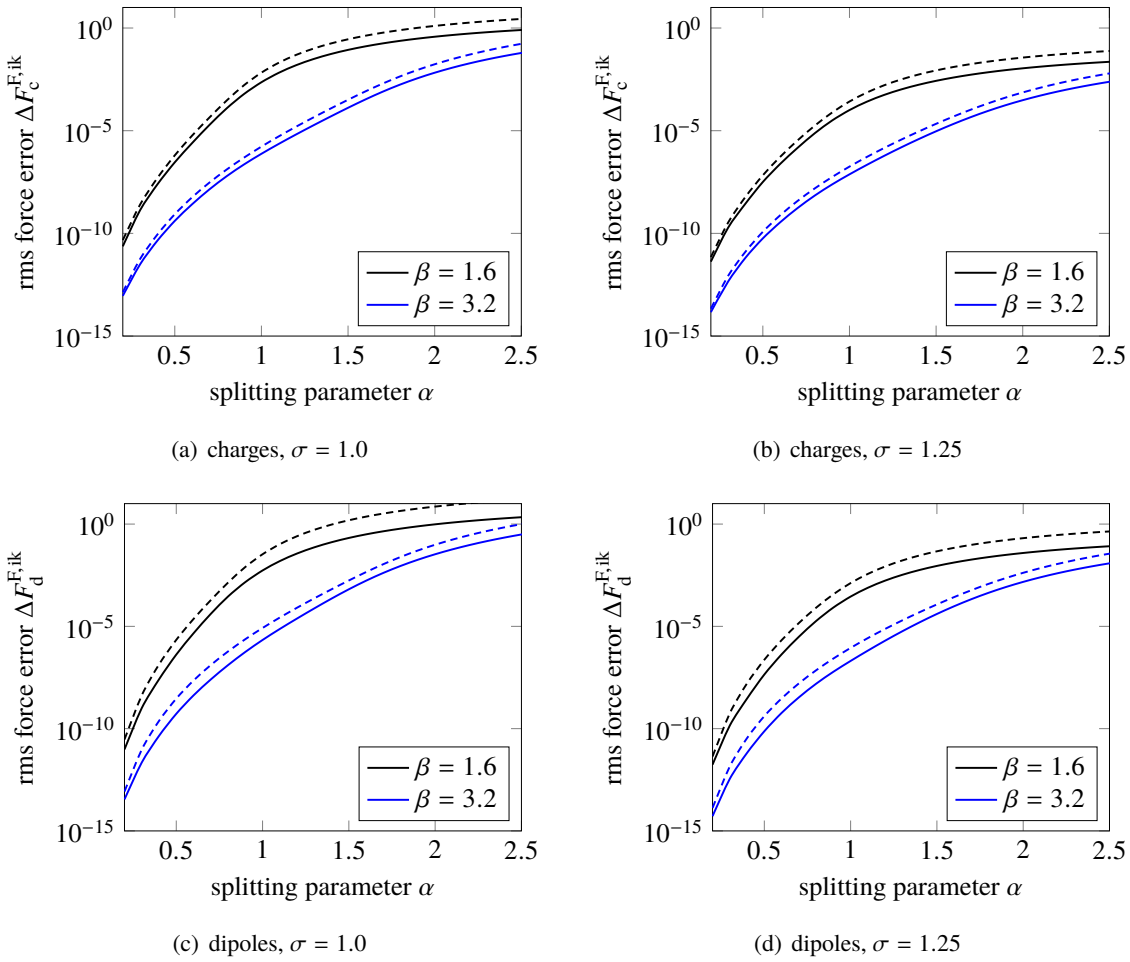
The truncated Fourier space parts of the forces have been approximated for varying splitting parameter α via the P²NFFT method, where we used the B-spline window function. Corresponding reference data have been computed based on the P²NDFT Ewald method, in each case. Exploiting the underlying orthorhombic box shape, the resulting rms errors have been predicted as described in Remark 4.16, requiring $O(M_1 + M_2 + M_3)$ arithmetic operations.

We applied the support parameter $m = 4$ and oversampling factors $\sigma \in \{1, 1.25\}$ to obtain the results depicted in Figure 5.3. It can be seen that the estimated and the measured errors do not perfectly match, but the actual errors are properly estimated from above. □

Example 5.4. We consider Particle system 4, where the primary simulation box has a triclinic shape. Again, we considered different parameter combinations and determined the predicted as well as the actual approximation errors, as summarized in Figure 5.4. The predicted errors have been obtained by evaluating (4.64) directly, which requires $O(|\mathcal{I}_{\boldsymbol{M}}|)$ arithmetic operations. Note that the applied far field cutoffs $\beta \in \{1.6, 3.2\}$ result in the following mesh sizes, cf. (5.1).

$$\beta = 1.6 \Rightarrow \boldsymbol{M} = (18, 36, 34) \quad \text{and} \quad \beta = 3.2 \Rightarrow \boldsymbol{M} = (36, 72, 66).$$

It can be seen that the actual errors are also estimated properly from above in case of the triclinic particle system. The estimates are somewhat more pessimistic than in case of an orthorhombic structure, as considered in Example 5.3. □



(a) charges, $\sigma = 1.0$

(b) charges, $\sigma = 1.25$

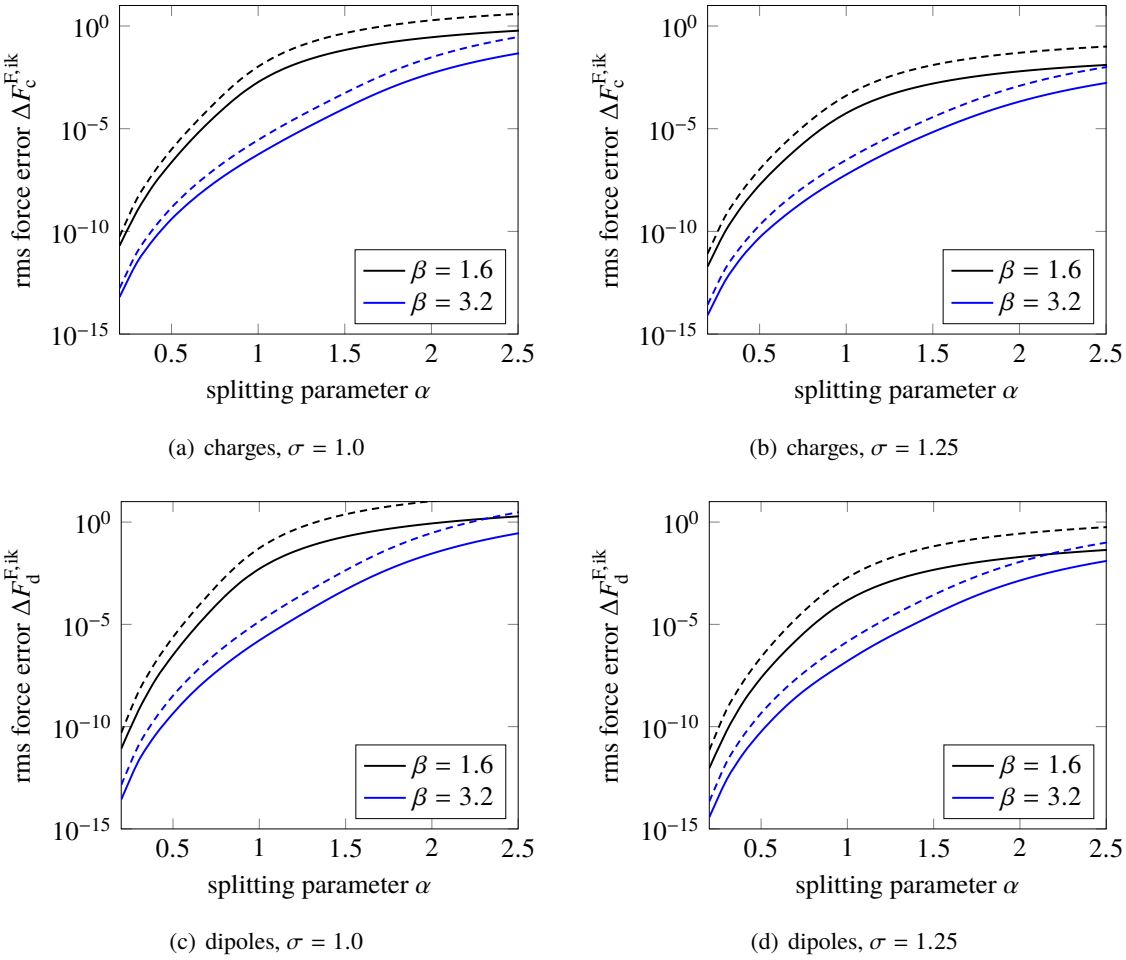
(c) dipoles, $\sigma = 1.0$

(d) dipoles, $\sigma = 1.25$

Figure 5.3: Estimated (dashed) and measured (solid) rms force errors for Particle system 3. We considered different mesh sizes $\boldsymbol{M}$ or rather far field cutoffs β (see legend). The B-spline window function has been applied with $i\boldsymbol{k}$-differentiation, support parameter $m = 4$ and oversampling factors $\sigma \in \{1, 1.25\}$ (see labels).

Example 5.5. We repeated the numerical tests concerning Particle system 3, as described in Example 5.3, and applied the Bessel instead of the B-spline window. For this purpose, the shape parameter b has been tuned in advance based on Algorithm 3.4 and Remark 4.17. Again, due to the orthorhombic box structure we predicted the resulting approximation errors based on Remark 4.16. The results are depicted in Figure 5.5, showing that the actual approximation errors are estimated adequately. Again, we remark that the optimal shape parameter b_{opt} strongly depends on the combination of the parameters β, α, σ and m. Thus, b_{opt} has been determined separately for all considered parameter sets. For overview purposes we abstain from listing the obtained optimal values.

In order to give a first comparison between the Bessel and the B-spline windows, we reconsider some error curves, as presented above, and compare the results obtained for the two different window functions in Figure 5.6. Obviously, smaller approximation errors are obtained with the Bessel window function in case of relatively large splitting parameters α. The B-spline window function performs better only for small values of α. □



(a) charges, $\sigma = 1.0$

(b) charges, $\sigma = 1.25$

(c) dipoles, $\sigma = 1.0$

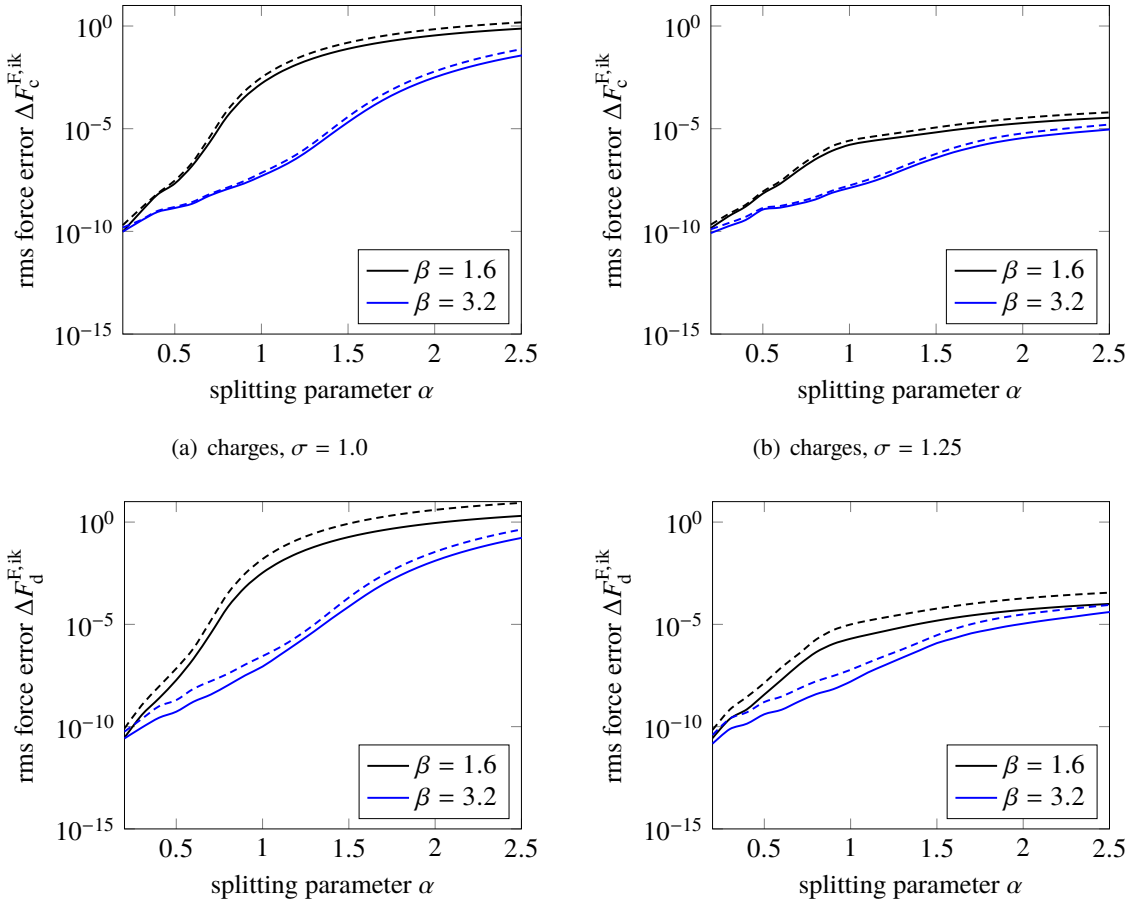
(d) dipoles, $\sigma = 1.25$

Figure 5.4: Estimated (dashed) and measured (solid) rms force errors (5.2) for Particle system 4. We considered different mesh sizes $\boldsymbol{M}$ or rather far field cutoffs β (see legend). The B-spline window function has been applied with i$\boldsymbol{k}$-differentiation, support parameter $m = 4$ and oversampling factors $\sigma \in \{1, 1.25\}$ (see labels).

Example 5.6. We repeated the numerical tests concerning Particle system 4, as described in Example 5.4, and replaced the B-spline by the Bessel window function. The shape parameter b has been tuned automatically for the different parameter sets based on Algorithm 3.4 and Remark 4.17. For this purpose, the error sums (4.64) have been evaluated directly due to the underlying triclinic geometry. The resulting actual approximation errors as well as the corresponding estimates are depicted in Figure 5.7. □

In the following example we consider the analytic differentiation approach for the B-spline window. The resulting rms force errors in the Fourier space parts have been derived in Section 4.3.2. Analogously to the i$\boldsymbol{k}$-differentiation approach we set

$$\Delta F_c^{\mathrm{F,ad}} \approx \sqrt{\left(\Delta F_{\mathrm{c.c.}}^{\mathrm{F,ad}}\right)^2 + \left(\Delta F_{\mathrm{c.d.}}^{\mathrm{F,ad}}\right)^2} \tag{5.3}$$



(a) charges, $\sigma = 1.0$

(b) charges, $\sigma = 1.25$

(c) dipoles, $\sigma = 1.0$

(d) dipoles, $\sigma = 1.25$

Figure 5.5: Estimated (dashed) and measured (solid) rms force errors (5.2) for Particle system 3. We considered different mesh sizes $\boldsymbol{M}$ or rather far field cutoffs β (see legend). The Bessel window function has been applied with $i\boldsymbol{k}$-differentiation, support parameter $m = 4$ and oversampling factors $\sigma \in \{1, 1.25\}$ (see labels).

and

$$\Delta F_{\mathrm{d}}^{\mathrm{F,ad}} \approx \sqrt{\left(\Delta F_{\mathrm{d.c.}}^{\mathrm{F,ad}}\right)^2 + \left(\Delta F_{\mathrm{d.d.}}^{\mathrm{F,ad}}\right)^2}. \tag{5.4}$$

Note that, in contrast to the $i\boldsymbol{k}$-differentiation approach, we expect that the computed self interactions might represent a significant portion of the overall force computed. Thus, the derived error estimates might not describe the actual error behavior to the full extent.

Example 5.7. We consider the Particle systems 1 and 2, where only charge-charge and dipole-dipole interactions are present, respectively. As already explained in the previous examples, we considered different sets of parameters and determined the resulting actual approximation errors as well as the corresponding estimates in case of analytic differentiation, see Section 4.3.2. The results are summarized in Figure 5.8.

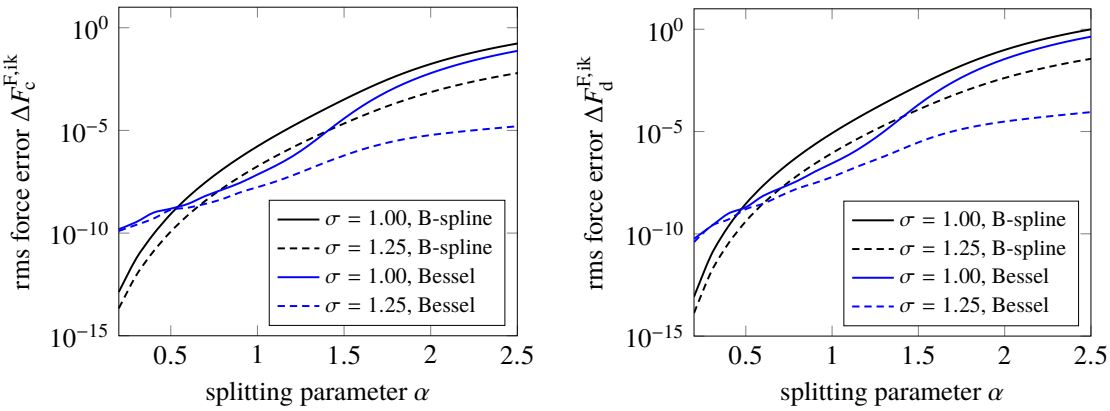


Figure 5.6: Comparison of the B-spline and the Bessel window. We plot the errors for the charges (left) as well as for the dipoles (right) with repsect to the splitting parameter α, where we fixed the far field cutoff $\beta = 3.2$. The P²NFFT method was applied to Particle system 3 with i$\boldsymbol{k}$-differentiation, support parameter $m = 4$ and oversampling factors $\sigma \in \{1, 1.25\}$ (see legend).

It can be seen that the error estimates do represent the actual errors quite good. However, in some cases the measured errors exceed the computed estimates. This is due to the presence of significant self interaction terms, which are ignored in the presented error analysis. □

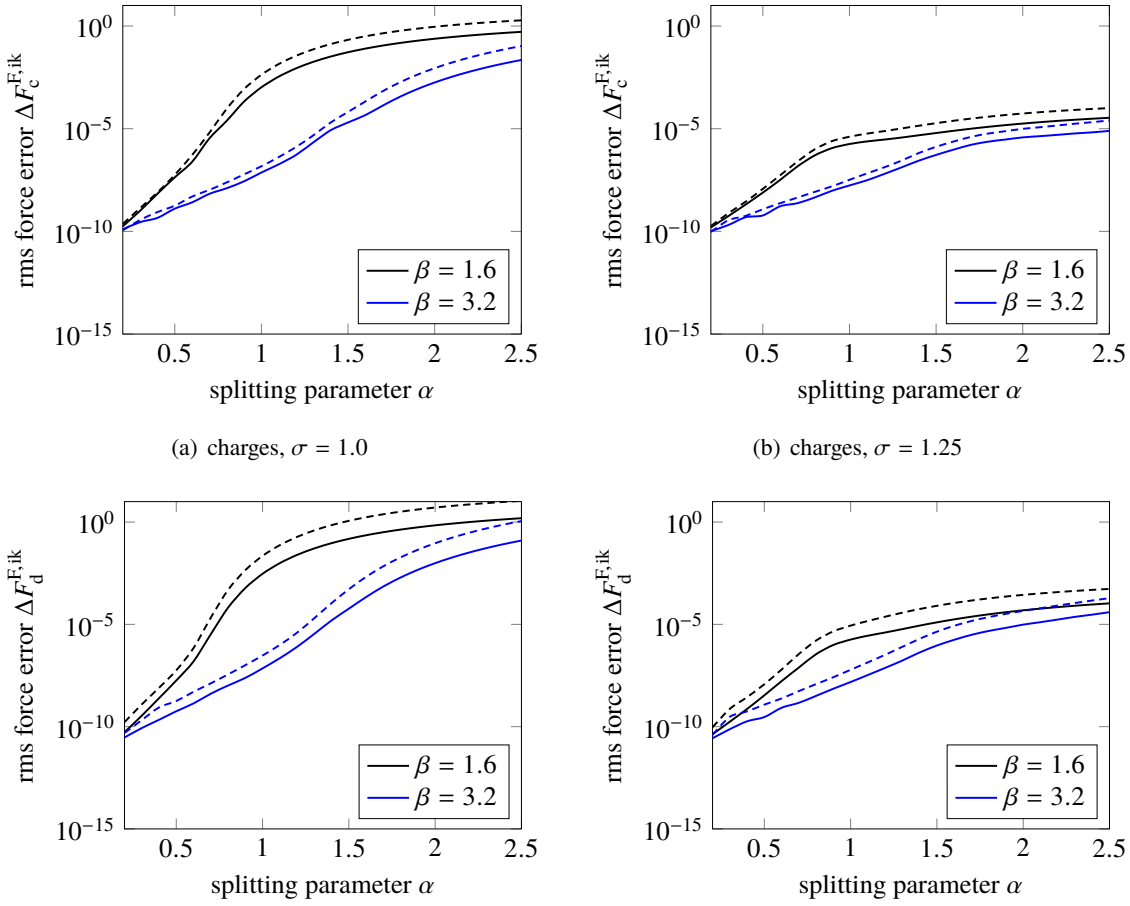
Example 5.8. Again, we consider Particle system 3, containing a mixture of $N_c = 300$ charges and $N_d = 300$ dipoles. Analogously to Example 5.7, we applied the P²NFFT method with the B-spline window function and analytic differentiation. We considered different parameter combinations and compared the measured errors with the predicted errors, see Figure 5.9. As already observed in the last example, the actual errors are represented by the computed estimates fairly well. For large splitting parameters α the measured errors exceed the estimates, due to the presence of significant self interactions.

In order to apply the Bessel window, the shape parameter b has to be chosen appropriately. Since no error estimates are available so far in case of the Bessel window function with analytic differentiation, the shape parameter cannot be tuned automatically in advance. For the purpose of demonstrating the applicability of the Bessel window in case of analytic differentiation, we simply optimized the shape parameter b based on measuring the resulting rms force errors and applying a simple bisection method, in a similar fashion as described in Algorithm 3.4.

The achieved rms force errors are represented by the dotted lines in Figure 5.9. It can be seen that we can achieve a high accuracy also for the Bessel window function. As already seen in case of i$\boldsymbol{k}$-differentiation, the errors obtained with the Bessel window are smaller than those obtained with the B-spline for large enough splitting parameters α. For small splitting parameters better results are achieved by applying the B-spline window. □

5.2.2 Parameter tuning

In this section we discuss how the parameters involved in the presented P²NFFT method for mixed charge-dipole systems may be tuned based on the previously considered and validated estimates. Given a required basic accuracy ε and a near field cutoff radius r_{cut}, the required splitting parameter α as



(a) charges, $\sigma = 1.0$

(b) charges, $\sigma = 1.25$

(c) dipoles, $\sigma = 1.0$

(d) dipoles, $\sigma = 1.25$

Figure 5.7: Estimated (dashed) and measured (solid) rms force errors (5.2) for Particle system 4. We considered different mesh sizes $\boldsymbol{M}$ or rather far field cutoffs β (see legend). The Bessel window function has been applied with $\mathrm{i}\boldsymbol{k}$-differentiation, support parameter $m = 4$ and oversampling factors $\sigma \in \{1, 1.25\}$ (see labels).

well as the far field cutoff β can be determined in order to achieve the desired accuracy regarding the truncation of the Ewald sums, as described in Section 2.3.4. Based on the error estimates presented in Section 4.3 we may compute viable parameter sets in order to reach the given accuracy ε also with respect to the NFFT approximations. For a given near field cutoff radius r_{cut} we suggest the following parameter tuning.

Algorithm 5.1 ($\mathrm{P^2NFFT}$ accuracy tuning)**.**
Input: Data describing the particle system (number of charges N_c, number of dipoles N_d, $Q = \sum_j q_j^2$, $\mathcal{M} = \sum_j \|\boldsymbol{\mu}_j\|^2$, box vectors $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3$), real space cutoff radius r_{cut}, basic accuracy $\varepsilon > 0$.

i) Tune the splitting parameter $\alpha > 0$ and the required far field cutoff $\beta > 0$ as described in Section 2.3.4.

ii) Define a set of support parameters m to be considered. For each m find the smallest oversampling

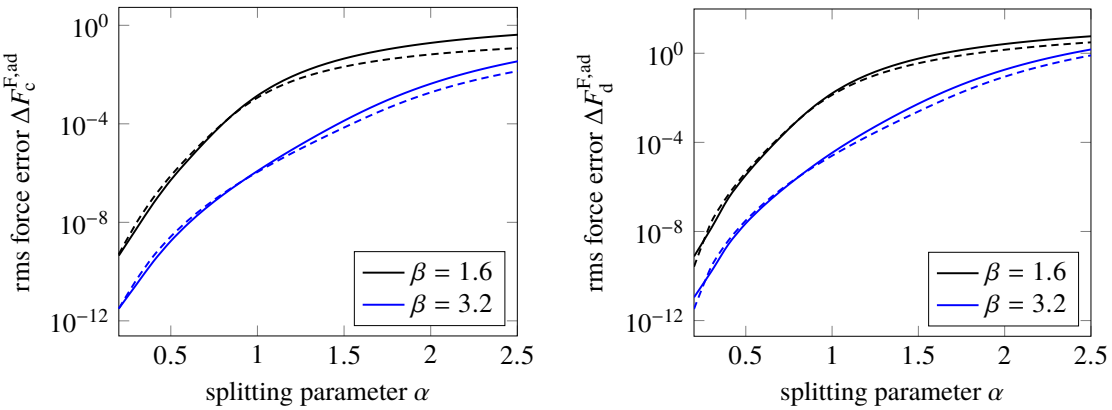


Figure 5.8: Estimated (dashed) as well as measured (solid) rms force errors (5.3) and (5.4) for the Particle systems 1 (left) and respectively 2 (right). We considered different mesh sizes $\boldsymbol{M}$ or rather far field cutoffs β (see legend). The B-spline window function has been applied with support parameter $m = 4$ and without oversampling, i.e., $\sigma = 1$. The partial derivatives have been computed based on the analytic differentiation approach.

factor $\sigma \geq 1$ such that

$$\max\left\{\Delta F_{\text{c.c.}}^{\text{F,ik}}, \Delta F_{\text{c.d.}}^{\text{F,ik}}, \Delta F_{\text{d.c.}}^{\text{F,ik}}, \Delta F_{\text{d.d.}}^{\text{F,ik}}\right\} \leq \tilde{\varepsilon} \leq \varepsilon \tag{5.5}$$

or

$$\max\left\{\Delta F_{\text{c.c.}}^{\text{F,ad}}, \Delta F_{\text{c.d.}}^{\text{F,ad}}, \Delta F_{\text{d.c.}}^{\text{F,ad}}, \Delta F_{\text{d.d.}}^{\text{F,ad}}\right\} \leq \tilde{\varepsilon} \leq \varepsilon \tag{5.6}$$

is fulfilled, depending on the desired differentiation approach (i$\boldsymbol{k}$ or analytic).

When limiting the range of possible oversampling factors to an interval, for example to $[1, 2]$, the required σ may be determined via a simple bisection method.

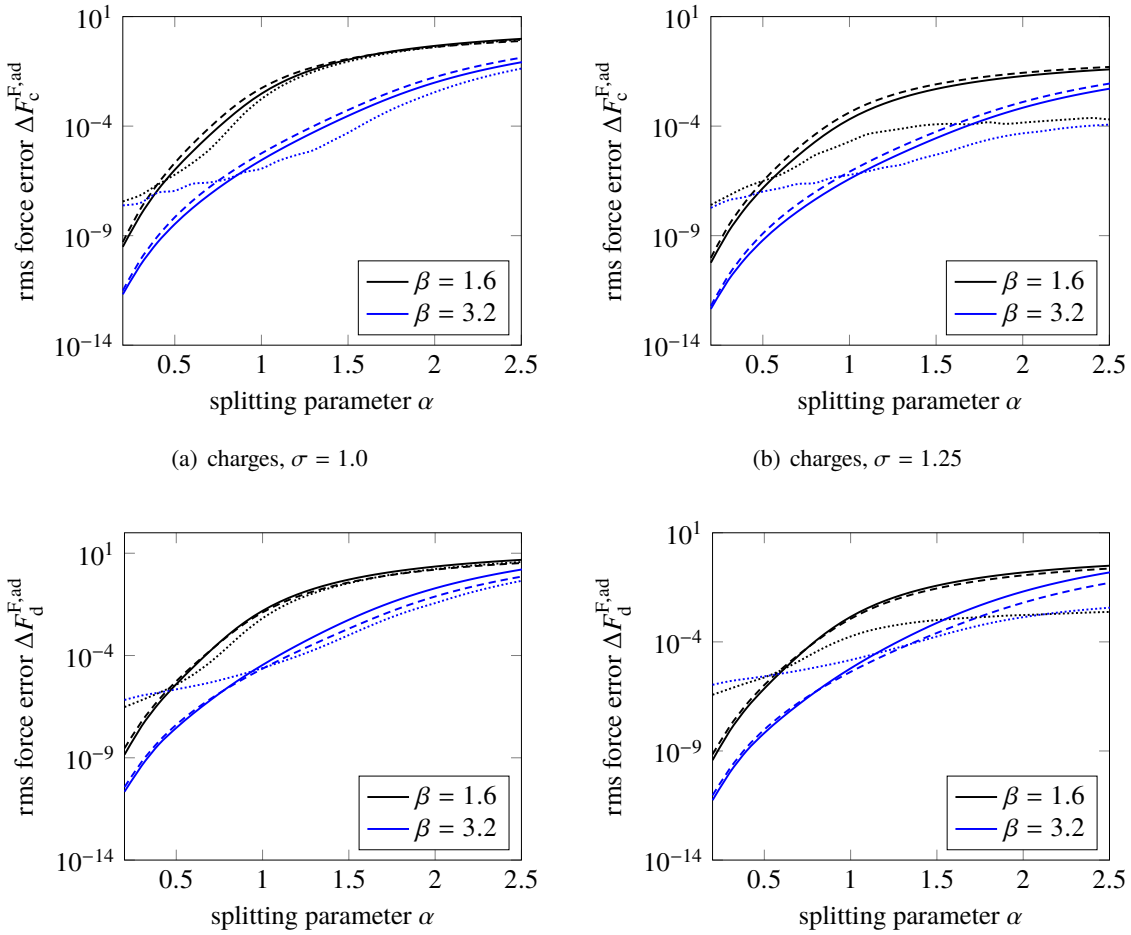
In case of the Bessel window function with i$\boldsymbol{k}$-differentiation, the optimal shape parameter b_{opt} can be determined based on Algorithm 3.4 and Remark 4.17.

iii) One combination of m and σ will be optimal regarding computational cost. This will strongly depend on the used hardware and software. Determine the optimal combination by measuring the required runtimes, for instance.

□

Remark 5.9 (Basic accuracy)**.** In our notation the value ε represents a certain basic accuracy, which is desired to be achieved in the single components of the P²NFFT algorithm, namely the truncation of the short range parts, the truncation of the Fourier space parts as well as the resulting NFFT approximation errors. The resulting overall rms force error might be slightly larger than ε, but nevertheless in the same order of magnitude, in case that the different parts of the error accumulate unfavorably. □

So far, we discussed how the parameters may be tuned in order to reach a certain basic accuracy ε. Algorithm 5.1 serves an approach, where only the near field cutoff radius r_{cut} has to be specified. In other words, the same accuracy is finally achieved for various near field cutoff radii r_{cut} and we still have one degree of freedom concerning the parameter choice. However, the choice of r_{cut} will



(a) charges, $\sigma = 1.0$

(b) charges, $\sigma = 1.25$

(c) dipoles, $\sigma = 1.0$

(d) dipoles, $\sigma = 1.25$

Figure 5.9: Estimated (dashed) as well as the measured (solid) rms force errors (5.4) for Particle system 3, where we applied the B-spline window with analytic differentiation. The obtained errors for the Bessel window are represented by the dotted lines. We considered different far field cutoffs β as well as oversampling factors σ (see labels) and applied the support parameter $m = 4$.

influence the required runtime a lot. Enlarging the near field cutoff radius r_{cut} will make the near field computations more expensive, whereas a smaller far field cutoff β will be required in order to reach the same accuracy, and vice versa. Thus, the computational cost required for the near field computations and the computations in Fourier space may be balanced out by choosing the near field cutoff radius r_{cut} appropriately.

Remark 5.10 (Applied architecture)**.** In the following numerical examples we present computation times of the P^{2}NFFT algorithm. All computations were performed on an Intel i5-3470 processor that runs on 3.20 GHz with 8 GB main memory. Furthermore, we remark that the runtimes include the computation of the potentials and the fields for the charges as well as the computation of the fields and the field gradients for the dipole particles. □

Example 5.11. We consider Particle system 3. In order to apply Algorithm 5.1, we fixed the near field cutoff radius $r_{\text{cut}} := 5$ and the basic accuracy $\varepsilon := 10^{-5}$. Consequently, we obtained the splitting parameter $\alpha \approx 0.72$ and the far field cutoff $\beta \approx 1.64$, i.e., $\boldsymbol{M} = (34, 18, 18)$, in Step 1 of Algorithm 5.1.

In Step 2 of Algorithm 5.1 we considered support parameters $m \in \{2, \ldots, 7\}$ and oversampling factors $\sigma \in [1, 2]$, where the latter have been tuned with a precision of $\frac{1}{128} = 0.0078125$. Furthermore, we set $\tilde{\varepsilon} := \varepsilon = 10^{-5}$ in (5.5), i.e., we applied the i$\boldsymbol{k}$-differentiation approach.

In Tables 5.1 and 5.2 we list the tuned oversampling factors $\sigma \in [1, 2]$ for different support parameters m, the required runtimes t as well as the achieved rms force errors ΔF_{c} and ΔF_{d} for the B-spline and also for the Bessel window. In case of the Bessel window we also list the corresponding tuned optimal shape parameters b_{opt}.

The achieved rms force errors are all below the applied basic accuracy $\varepsilon = 10^{-5}$. In general, less oversampling is required in case of the Bessel window function. In this concrete example, $m = 5$ with $\sigma \approx 1.08$ is optimal for the B-spline window and $m = 4$ with $\sigma = 1$ is optimal in case of the Bessel window. Consequently, we achieve an improvement concerning computational costs by 20% when using the Bessel window function. The presented runtimes have been measured on the architecture specified in Remark 5.10, using 1 of 4 available cores. □

m	σ	t	ΔF_{c}	ΔF_{d}
3	1.72	1.25e–01	2.75e–06	7.94e–06
4	1.23	7.65e–02	2.78e–06	8.06e–06
5	1.08	6.91e–02	2.72e–06	7.86e–06
6	1.02	7.39e–02	2.71e–06	7.89e–06
7	1.00	7.86e–02	2.78e–06	8.14e–06

Table 5.1: Possible combinations of m and σ for the B-spline window. Algorithm 5.1 has been applied for Particle system 3 by setting $r_{\text{cut}} := 5$ and $\varepsilon = \tilde{\varepsilon} := 10^{-5}$. We also list the resulting runtimes t and the achieved rms force errors.

m	σ	b_{opt}	t	ΔF_{c}	ΔF_{d}
3	1.28	4.98	7.35e–02	3.65e–06	8.08e–06
4	1.00	3.93	5.59e–02	2.84e–06	8.05e–06
5	1.00	3.44	6.21e–02	2.71e–06	7.90e–06
6	1.00	3.34	6.88e–02	2.70e–06	7.90e–06
7	1.00	3.34	7.83e–02	2.70e–06	7.90e–06

Table 5.2: Possible combinations of m and σ for the Bessel window. Algorithm 5.1 has been applied for Particle system 3 by setting $r_{\text{cut}} := 5$ and $\varepsilon = \tilde{\varepsilon} := 10^{-5}$. We also list the tuned optimal shape parameters b_{opt}, the measured runtimes t and the achieved rms force errors.

Example 5.12. We reran the numerical test, as described Example 5.11, for the B-spline window function and applied the analytic differentiation instead of the i$\boldsymbol{k}$-differentiation approach. As discussed in the previous sections, the error control is more challenging in case of analytic differentiation, which is mainly due to the computation of significant self forces concerning charge-charge and dipole-dipole interactions, see Example 5.7 for instance.

In Table 5.3 we list the tuned oversampling factors σ for different support parameters m, the measured runtimes t as well as the achieved rms force errors. With $\tilde{\varepsilon} := \varepsilon = 10^{-5}$ in (5.6) we achieve

slightly larger errors compared to the i$\boldsymbol{k}$-differentiation approach. By setting $\tilde{\varepsilon} := 5 \cdot 10^{-6}$ the errors are comparable to those obtained via i$\boldsymbol{k}$-differentiation. In this concrete example, we could save approximately 50% of runtime by applying the analytic differentiation approach. □

m	σ	t	ΔF_c	ΔF_d
3	>2			
4	1.36	3.62e–02	2.79e–06	1.02e–05
5	1.08	3.33e–02	2.91e–06	1.00e–05
6	1.00	3.60e–02	3.23e–06	1.21e–05
7	1.00	4.24e–02	2.83e–06	9.15e–06

m	σ	t	ΔF_c	ΔF_d
3	>2			
4	1.48	3.92e–02	2.71e–06	8.77e–06
5	1.14	3.50e–02	2.72e–06	8.37e–06
6	1.03	3.78e–02	2.72e–06	8.33e–06
7	1.00	4.24e–02	2.83e–06	9.15e–06

Table 5.3: Possible combinations of m and σ for the B-spline window with analytic differentiation, where we applied $r_\mathrm{cut} := 5$, $\tilde{\varepsilon} := \varepsilon = 10^{-5}$ (left) and $\tilde{\varepsilon} := \frac{\varepsilon}{2} = 5 \cdot 10^{-6}$ (right), respectively, in Algorithm 5.1. We also list the measured runtimes t and the achieved rms force errors.

In the following we illustrate the dependence of the required runtime on the near field cutoff radius r_cut. For this purpose, we applied Algorithm 5.1 in order to tune the parameters automatically for various cutoff radii r_cut.

Example 5.13. We consider again Particle system 3. In order to study the dependence of the required runtime on the near field cutoff radius r_cut, we applied Algorithm 5.1 for varying r_cut. For this purpose, we considered the i$\boldsymbol{k}$-differentiation approach with $\varepsilon = \tilde{\varepsilon} := 10^{-5}$. For each near field cutoff radius r_cut we kept the parameter constellation yielding the smallest runtime, see Step 3 of Algorithm 5.1.

The computations have been executed in parallel using 4 cores on the architecture specified in Remark 5.10. The required runtimes t and the achieved rms force errors ΔF_c, ΔF_d are summarized in Figure 5.10. In addition, we plot the tuned splitting parameters α, the far field cutoffs β, the determined optimal support parameters m and the corresponding required oversampling factors σ.

For both the B-spline as well as the Bessel window the errors are smaller than the given basic accuracy $\varepsilon = 10^{-5}$. It can be seen that we require less oversampling or a smaller support parameter m when applying the Bessel window function, resulting in lower runtimes. The lowest runtimes could be achieved for $r_\mathrm{cut} = 5.5$ in case of the Bessel window and for $r_\mathrm{cut} = 6.5$ in case of the B-spline window. In this concrete example, we achieved an improvement concerning computational costs by approximately 15% with the Bessel window, compared to the B-spline window.

We remark that the time needed for the computation of the FFTs strongly depends on whether the oversampled mesh size $\boldsymbol{M}_\mathrm{o}$ yields an advantageous decomposition into prime factors. Thus, single plots may yield unexpected behavior or jumps. □

Remark 5.14 (Scale up to larger particle systems)**.** As already mentioned in Remark 2.25, the Ewald truncation errors are expected to be almost constant for particle systems possessing the same density of charges and dipoles, provided that the same splitting parameter α as well as the same cutoff parameters r_cut and β are applied.

Now, the question arises whether this also applies to the approximation errors, as considered here. For orthorhombic box structures it can be shown easily that the parameters, once tuned for a small particle system, can be applied to larger particle systems with the same charge and dipole densities to yield approximately the same approximation error.

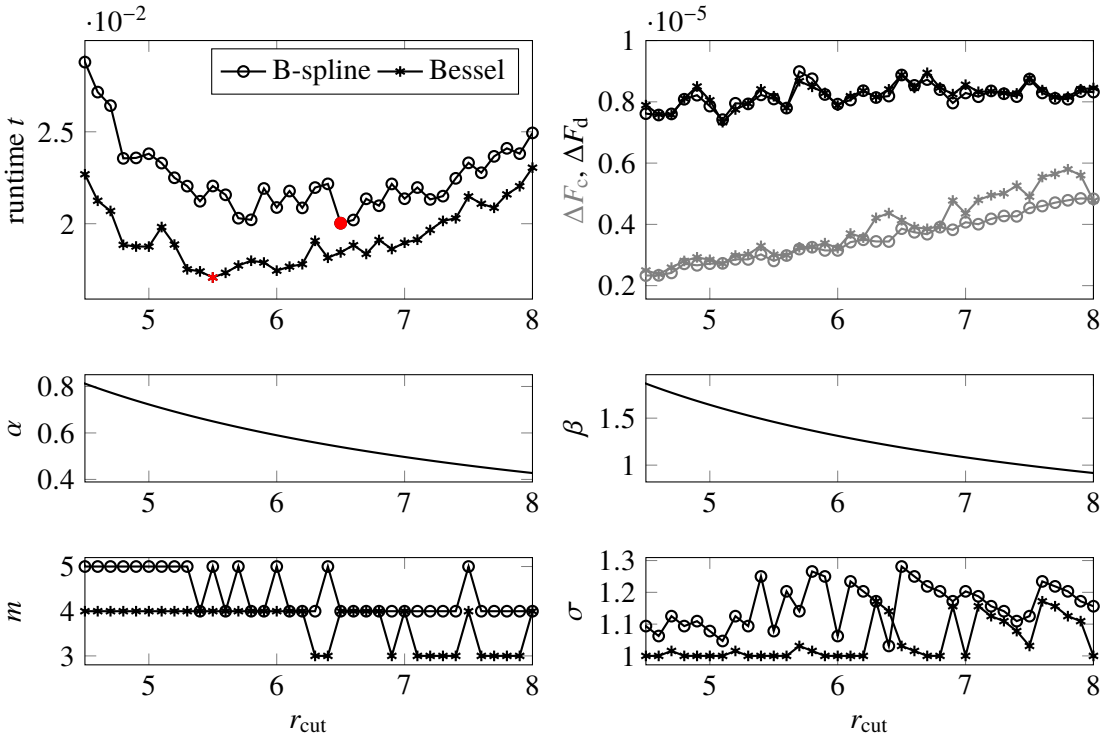


Figure 5.10: Achieved runtimes (in parallel with 4 cores) as well as the resulting rms force errors for Particle system 3. We also plot the tuned parameters α, β, m and σ with respect to r_{cut}.

For orthorhombic box structures we have $\boldsymbol{\ell}_i^\top \boldsymbol{\ell}_j = 0$, $i \neq j$, and thus the dual lattice vectors, which are the columns of the matrix $\boldsymbol{L}^{-\top}$, are given by $\hat{\boldsymbol{\ell}}_j = \|\boldsymbol{\ell}_j\|^{-2}\boldsymbol{\ell}_j$ for $j = 1, 2, 3$. Hence, we obtain

$$\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 = \frac{k_1^2}{\|\boldsymbol{\ell}_1\|^2} + \frac{k_2^2}{\|\boldsymbol{\ell}_2\|^2} + \frac{k_3^2}{\|\boldsymbol{\ell}_3\|^2} \approx \beta^2 \left(\frac{k_1^2}{M_1^2} + \frac{k_2^2}{M_2^2} + \frac{k_3^2}{M_3^2} \right)$$

as well as

$$V = \|\boldsymbol{\ell}_1\|\|\boldsymbol{\ell}_2\|\|\boldsymbol{\ell}_3\| \approx \frac{M_1 M_2 M_3}{\beta^3} = \frac{|\mathcal{I}_{\boldsymbol{M}}|}{\beta^3}.$$

In case of charge-charge interactions we get, cf. (4.58) and (4.63),

$$\begin{aligned}\left(\Delta F_{\text{c.c.}}^{\text{F,ik}}\right)^2 &\sim \frac{Q^2}{N_{\text{c}}} \chi_{\text{c.c.}}^2 \sim \frac{Q}{N_{\text{c}}} \cdot \frac{Q}{V} \cdot \frac{1}{V} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 \hat{\psi}(\boldsymbol{k})^2 \left(1 - 2\hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}^2(\tilde{\varphi}_{\text{t}}) + \hat{d}_{\boldsymbol{k}}^4 Y_0(\boldsymbol{k})^2\right) \\ &\sim \frac{Q}{N_{\text{c}}} \cdot \frac{Q}{V} \cdot \frac{\beta}{|\mathcal{I}_{\boldsymbol{M}}|} \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{M}}} \frac{\mathrm{e}^{-2\pi^2\beta^2 \|\boldsymbol{k} \odot \boldsymbol{M}^{-1}\|^2/\alpha^2}}{\|\boldsymbol{k} \odot \boldsymbol{M}^{-1}\|^2} \left(1 - 2\hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}^2(\tilde{\varphi}_{\text{t}}) + \hat{d}_{\boldsymbol{k}}^4 Y_0(\boldsymbol{k})^2\right).\end{aligned}$$

Furthermore, it is seen easily that the expression in brackets can be interpreted as a function in $\boldsymbol{k} \odot \boldsymbol{M}^{-1}$ for fixed support parameter m, oversampling factor σ and shape parameter b, i.e.,

$$1 - 2\hat{d}_{\boldsymbol{k}}^2 c_{\boldsymbol{k}}^2(\tilde{\varphi}_{\text{t}}) + \hat{d}_{\boldsymbol{k}}^4 Y_0(\boldsymbol{k})^2 =: f_{m,\sigma,b}(\boldsymbol{k} \odot \boldsymbol{M}^{-1}),$$

where $\boldsymbol{k} \odot \boldsymbol{M}^{-1} \in [1/2, 1/2)^3$. In other words, the sum can be interpreted as a discretization of an integral, i.e.,

$$\left(\Delta F_{\mathrm{c.c.}}^{\mathrm{F,ik}}\right)^2 \sim \frac{Q^2\beta}{N_{\mathrm{c}}V} \int_{[1/2,1/2)^3} \frac{\mathrm{e}^{-2\pi^2\beta^2\|\boldsymbol{x}\|^2/\alpha^2}}{\|\boldsymbol{x}\|^2} f_{m,\sigma,b}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Thus, if the densities ${N_{\mathrm{c}}}^{-1}Q$ and $V^{-1}Q$ as well as the parameters α, β, m, σ and b are kept constant, the errors are expected to be of a comparable size. Enlarging the grid size $\boldsymbol{M}$ with growing system size via (2.82) simply results in a finer resolution of the applied quadrature rule. In a similar fashion, the same argumentation applies to the errors $\Delta F_{\mathrm{c.d.}}^{\mathrm{F,ik}}$, $\Delta F_{\mathrm{d.c.}}^{\mathrm{F,ik}}$ and $\Delta F_{\mathrm{d.d.}}^{\mathrm{F,ik}}$.

More generally, it can be shown that the approximation errors are almost the same if additionally the angles between the box vectors $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3$ are kept constant. We obtain

$$\begin{aligned} V = \|\boldsymbol{\ell}_1^\top(\boldsymbol{\ell}_2 \times \boldsymbol{\ell}_3)\| &= \cos \sphericalangle_{\boldsymbol{\ell}_1,\boldsymbol{\ell}_2\times\boldsymbol{\ell}_3} \sin \sphericalangle_{\boldsymbol{\ell}_2,\boldsymbol{\ell}_3} \|\boldsymbol{\ell}_1\|\|\boldsymbol{\ell}_2\|\|\boldsymbol{\ell}_3\| \\ &\approx \cos \sphericalangle_{\boldsymbol{\ell}_1,\boldsymbol{\ell}_2\times\boldsymbol{\ell}_3} \sin \sphericalangle_{\boldsymbol{\ell}_2,\boldsymbol{\ell}_3} \frac{M_1M_2M_3}{\beta^3}. \end{aligned}$$

Furthermore, we get

$$\begin{aligned} \|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2 &= \boldsymbol{k}^\top \boldsymbol{L}^{-1}\boldsymbol{L}^{-\top}\boldsymbol{k} \\ &= (\boldsymbol{k} \odot \boldsymbol{M}^{-1})^\top \mathrm{diag}(M_1, M_2, M_3)\boldsymbol{L}^{-1}\boldsymbol{L}^{-\top}\mathrm{diag}(M_1, M_2, M_3)(\boldsymbol{k} \odot \boldsymbol{M}^{-1}) \\ &\approx \beta^2(\boldsymbol{k} \odot \boldsymbol{M}^{-1})^\top \begin{pmatrix} \|\boldsymbol{\ell}_1\|^2\hat{\boldsymbol{\ell}}_1^\top\hat{\boldsymbol{\ell}}_1 & \|\boldsymbol{\ell}_1\|\|\boldsymbol{\ell}_2\|\hat{\boldsymbol{\ell}}_1^\top\hat{\boldsymbol{\ell}}_2 & \|\boldsymbol{\ell}_1\|\|\boldsymbol{\ell}_3\|\hat{\boldsymbol{\ell}}_1^\top\hat{\boldsymbol{\ell}}_3 \\ \|\boldsymbol{\ell}_2\|\|\boldsymbol{\ell}_1\|\hat{\boldsymbol{\ell}}_2^\top\hat{\boldsymbol{\ell}}_1 & \|\boldsymbol{\ell}_2\|^2\hat{\boldsymbol{\ell}}_2^\top\hat{\boldsymbol{\ell}}_2 & \|\boldsymbol{\ell}_2\|\|\boldsymbol{\ell}_3\|\hat{\boldsymbol{\ell}}_2^\top\hat{\boldsymbol{\ell}}_3 \\ \|\boldsymbol{\ell}_3\|\|\boldsymbol{\ell}_1\|\hat{\boldsymbol{\ell}}_3^\top\hat{\boldsymbol{\ell}}_1 & \|\boldsymbol{\ell}_3\|\|\boldsymbol{\ell}_2\|\hat{\boldsymbol{\ell}}_3^\top\hat{\boldsymbol{\ell}}_2 & \|\boldsymbol{\ell}_3\|^2\hat{\boldsymbol{\ell}}_3^\top\hat{\boldsymbol{\ell}}_3 \end{pmatrix} (\boldsymbol{k} \odot \boldsymbol{M}^{-1})^\top \\ &= \beta^2(\boldsymbol{k} \odot \boldsymbol{M}^{-1})^\top \underbrace{\frac{\|\boldsymbol{\ell}_1\|^2\|\boldsymbol{\ell}_2\|^2\|\boldsymbol{\ell}_3\|^2}{V^2} \begin{pmatrix} s_{23}^2 & s_{13}s_{23}\hat{c}_{12} & s_{12}s_{23}\hat{c}_{13} \\ s_{13}s_{23}\hat{c}_{12} & s_{13}^2 & s_{12}s_{13}\hat{c}_{23} \\ s_{12}s_{23}\hat{c}_{13} & s_{12}s_{13}\hat{c}_{23} & s_{12}^2 \end{pmatrix}}_{=:\,C} (\boldsymbol{k} \odot \boldsymbol{M}^{-1}), \end{aligned}$$

where we define

$$s_{ij} := \sin \sphericalangle_{\boldsymbol{\ell}_i,\boldsymbol{\ell}_j} \quad \text{and} \quad \hat{c}_{ij} := \cos \sphericalangle_{\hat{\boldsymbol{\ell}}_i,\hat{\boldsymbol{\ell}}_j}.$$

The result can be obtained by utilizing the representations of the dual lattice vectors via

$$\hat{\boldsymbol{\ell}}_1 = \frac{\boldsymbol{\ell}_2 \times \boldsymbol{\ell}_3}{\boldsymbol{\ell}_1^\top(\boldsymbol{\ell}_2 \times \boldsymbol{\ell}_3)},$$

where $|\boldsymbol{\ell}_1^\top(\boldsymbol{\ell}_2 \times \boldsymbol{\ell}_3)| = V$. The matrix C remains the same, provided that the angles between the box vectors do not change. Thus, the term $\|\boldsymbol{L}^{-\top}\boldsymbol{k}\|^2$ can again be interpreted as a function of $\boldsymbol{k} \odot \boldsymbol{M}^{-1}$, for fixed β. Now, the argumentation is the same as for orthorhombic geometries.

This remark, together with Remark 2.25, confirms that the parameters α, β, r_{cut}, m, σ and b (if required) may be tuned for a small particle system and then be applied to larger systems, while the approximation errors are kept almost constant. Of course, this is not optimal regarding runtime. Since the near field cutoff radius r_{cut} is kept constant, the near field computations scale like $O(N)$. The computation of the Fourier space parts yields an arithmetic complexity of $O(N \log N)$, since the number of mesh points $\boldsymbol{M}_{\mathrm{o}}$ scales linearly with the size of the box or rather with the number of particles N. We expect that an optimization with respect to runtime, which takes into account the number of particles N, results in an arithmetic complexity of $O(N\sqrt{\log N}(\log\log N)^{3/2})$, cf. [Pip15, Section 4.8]. □

Example 5.15. We consider particle systems possessing the same charge and dipole densities as Particle system 3, see Table 5.4. Each system contains $N_c = N/2$ charges $q_j = (-1)^j$ and $N_d = N/2$ dipoles with randomly oriented dipole moments $\boldsymbol{\mu}_j \in \{\boldsymbol{y} \in \mathbb{R}^3, \|\boldsymbol{y}\| = 1\}$, uniformly distributed in an orthorhombic box of size $L_1 \times L_2 \times L_3$, respectively.

N	L_1	L_2	L_3
600	20	10	10
2400	20	20	20
9600	40	40	20
38400	80	40	40
153600	80	80	80

Table 5.4: Considered particle systems with $N_c = N_d = Q = \mathcal{M} = N/2$. The particles are randomly distributed in orthorhombic boxes of size $L_1 \times L_2 \times L_3$, respectively.

In order to tune the parameters for the introduced particle systems we proceeded as follows. For each near field cutoff radius r_{cut}, the parameters α and β as well as possible combinations of m and σ were tuned for the small system with $N = 600$ particles via Algorithm 5.1, Steps 1 and 2. Then, the obtained sets of parameters have been applied also to the larger particle systems. For each r_{cut} we kept the combination of m and σ minimizing the required computation time, individually for each single particle systems. As an example, for $r_{cut} = 6.3$ the obtained optimal parameters for the different systems are listed in Table 5.5. It can be seen that for growing number of particles it is, by trend, more advantageous to use less oversampling in combination with a larger support of the window function.

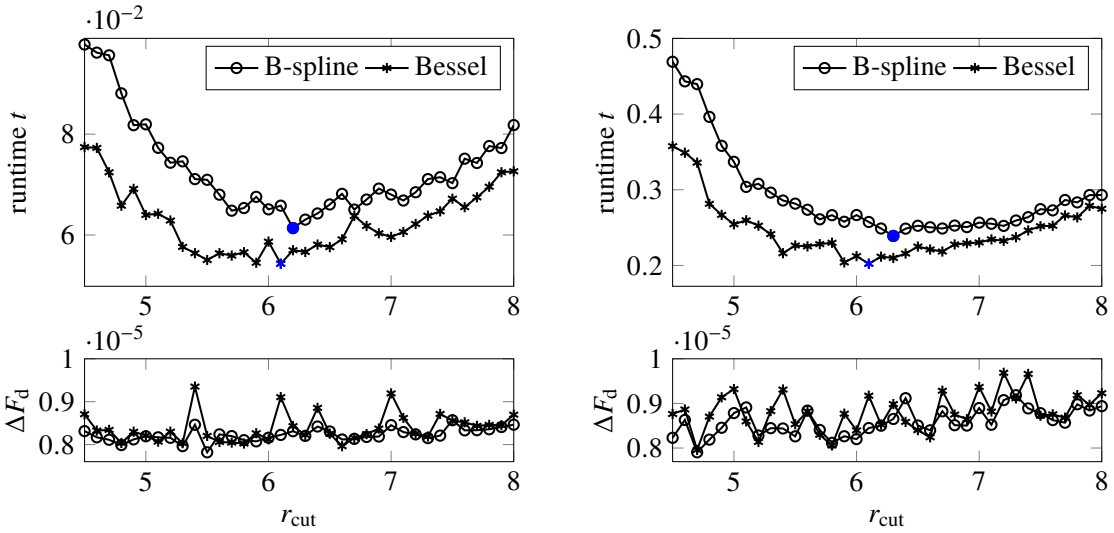
	B-spline		Bessel		
N	m	σ	m	σ	b_{opt}
600	4	1.17	3	1.17	4.84
2400	4	1.17	4	1.00	3.88
9600	4	1.17	4	1.00	3.88
38400	5	1.05	4	1.00	3.88
153600	6	1.00	4	1.00	3.88

Table 5.5: Optimal parameters regarding overall compuational cost for fixed near field cutoff $r_{cut} = 6.3$.

The measured computation times t as well as the achieved rms force errors for the dipoles ΔF_d are depicted in Figure 5.11. For all considered sets of parameters we obtained $\Delta F_c < \Delta F_d$. Thus, the measured rms force errors for the charges are omitted, for overview purposes. We also abstain from plotting the tuned optimal parameters m and σ for each of the considered particle systems.

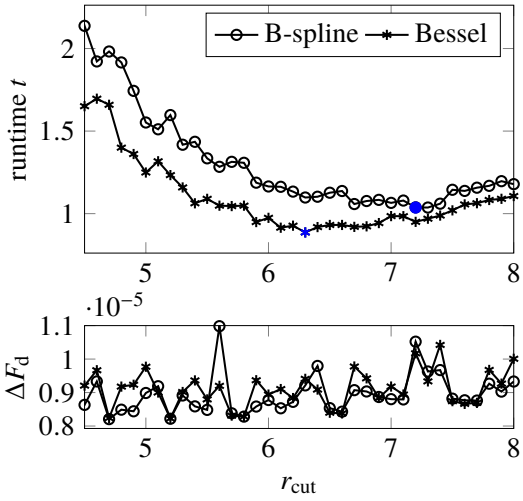
It can be seen that the required computation times are very close to the achieved optimum for different values of r_{cut}. However, a certain trend indicating that the optimal r_{cut} grows with increasing system size becomes visible. Although all possible parameter sets have been tuned for the small particle system, we achieve approximation errors of a comparable size among all systems.

In order to achieve an $O(N \log N)$ scaling we may simply apply the optimal parameters, tuned for the smallest particle system, also to the larger particle systems, i.e., the parameters r_{cut}, α, m, σ (and b_{opt}) are kept constant and the mesh size is scaled according to the box lengths L_1, L_2 and L_3 via $\boldsymbol{M} \approx \beta(L_1, L_2, L_3)$ as well as $\boldsymbol{M}_o \approx \beta\sigma(L_1, L_2, L_3)$.

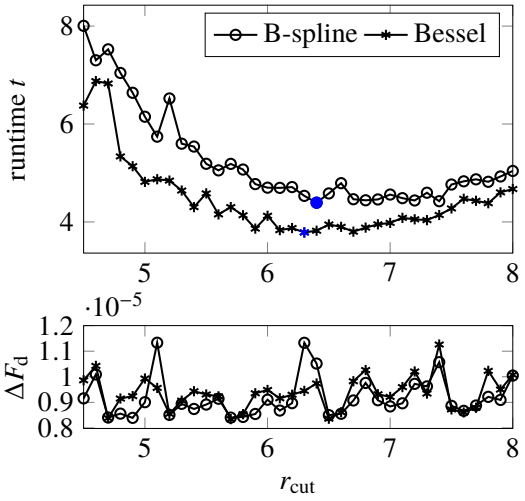


(a) $N_c = N_d = 1200$

(b) $N_c = N_d = 4800$



(c) $N_c = N_d = 19200$



(d) $N_c = N_d = 76800$

Figure 5.11: Measured computation times t (in parallel with 4 cores) and achieved rms force errors of the dipole partilces ΔF_d for the particle systems listet in Table 5.4.

In Example 5.13 the parameters

$$r_{cut} = 6.5, \alpha \approx 0.54, \beta \approx 1.18, m = 4, \sigma \approx 1.28 \tag{5.7}$$

and

$$r_{cut} = 5.5, \alpha \approx 0.65, \beta \approx 1.46, m = 4, \sigma = 1, b_{opt} \approx 3.83 \tag{5.8}$$

were found to be optimal for the B-spline and the Bessel window function, respectively. Applying these parameters to all particle systems listed in Table 5.4, we obtained the runtimes t as well as rms force errors ΔF_c and ΔF_d as depicted in Figure 5.12 on the left hand side. In contrast, on the right

hand side we plotted the results obtained by applying the optimal parameters individually found for the single systems, see Table 5.6 and blue marks in Figure 5.11. The results give numerical evidence that the method scales at least like $O(N \log N)$. □

	B-spline					Bessel					
N	r_{cut}	α	β	m	σ	r_{cut}	α	β	m	σ	b_{opt}
600	6.5	0.54	1.19	4	1.28	5.5	0.65	1.46	4	1.00	3.83
2400	6.2	0.57	1.26	4	1.20	6.1	0.58	1.29	4	1.00	3.73
9600	6.3	0.56	1.23	4	1.17	6.1	0.58	1.29	4	1.00	3.73
38400	7.3	0.47	1.03	4	1.14	6.3	0.56	1.23	4	1.00	3.88
153600	6.4	0.55	1.21	5	1.03	6.3	0.56	1.23	4	1.00	3.88

Table 5.6: Optimal parameters regarding computational cost for the considered particle systems.

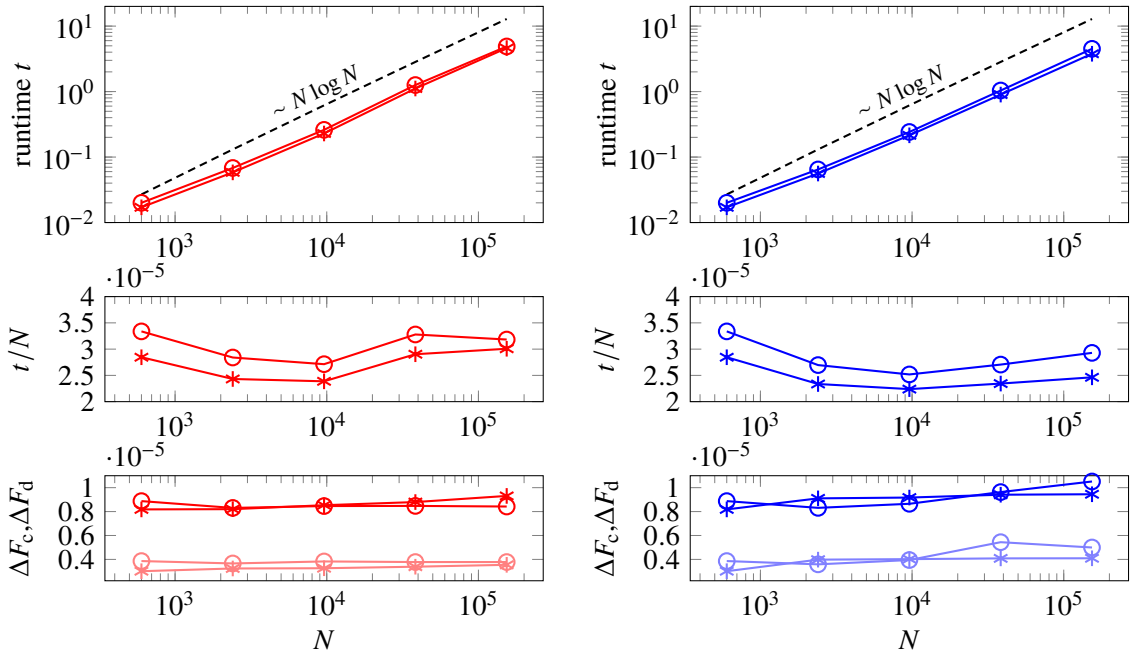


Figure 5.12: Runtime t with respect to the number of particles N, runtime per particle t/N and achieved rms force errors ΔF_c, ΔF_d. Left: We applied the parameters optimized for the small system with $N = 600$ particles, see (5.7) and (5.8). Right: We applied the optimal parameters determined individually for the different particle systems, see Table 5.6.

Example 5.16. This example will demonstrate the flexibility of our method resulting from the possibility of enabling oversampling. In this context we remark that classical particle-mesh methods, such as the P³M method [EHL80, HE88, DH98a], for instance, do not allow for the application of oversampling.

The previous examples already illustrate that it is often more advantageous in terms of computational cost to apply an oversampling factor $\sigma > 1$ in combination with a smaller support parameter m,

especially in case of the B-spline window function. We remark again that without oversampling, i.e., $\sigma = 1$, we cannot always expect to reach a specified desired accuracy, see Remark 3.6.

Here, we consider again Particle system 3 and apply the B-spline window without oversampling. In Example 5.13 the best computation times were achieved for $m \in \{4, 5\}$, combined with oversampling factors $\sigma > 1$, see Figure 5.10. Thereby, we achieved rms force errors smaller than $\varepsilon = 10^{-5}$.

In a further numerical experiment we fixed $\sigma := 1$ and increased the support parameter m step by step until (5.5) was fulfilled. As a maximum value we fixed $m_{\max} := 7$. We plot the achieved runtimes t as well as the measured rms force errors $\Delta F_c, \Delta F_d$ in Figure 5.13. It can be seen that we achieve approximation errors comparable to those obtained in Example 5.13 at the price of higher computational costs. In some cases $m = 7$ is not sufficient in order to obtain a comparably small error. □

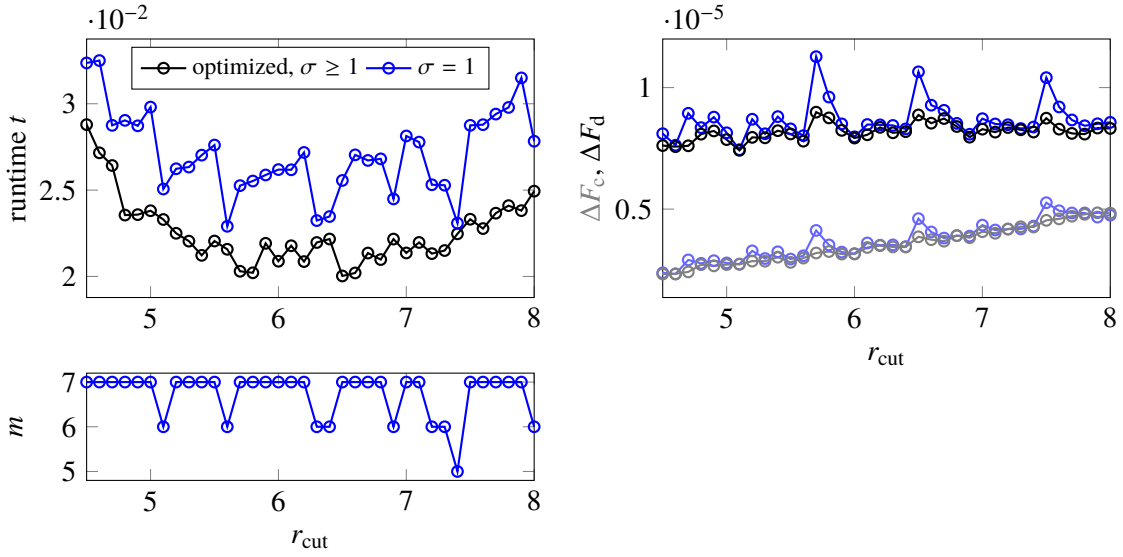


Figure 5.13: Achieved runtimes t, rms force errors $\Delta F_c, \Delta F_d$ and required support parameters $m \leq 7$. We applied the P^{2}NFFT method by using the B-spline window without oversampling (blue) to Particle system 3. For comparison, we also plot the results obtained in Example 5.13 (black).

Example 5.17. We reran Example 5.13 and applied the analytic differentiation approach, instead of i$\boldsymbol{k}$-differentiation. Again, we tuned all involved parameters for varying near field cutoff radius r_{cut} via Algorithm 5.1, where we set $\tilde{\varepsilon} := \frac{\varepsilon}{2} = 5 \cdot 10^{-6}$ in (5.6), as already applied in Example 5.12. The achieved runtimes t, the measured rms force errors $\Delta F_c, \Delta F_d$ as well as the tuned shape parameters m and oversampling factors σ are plotted in Figure 5.14. With the heuristic approach to set $\tilde{\varepsilon} := \frac{\varepsilon}{2}$ we achieve approximation errors comparable to those obtained with i$\boldsymbol{k}$-differentiation.

In Example 5.13 we achieved an optimal runtime of $t \approx 0.02$ seconds with the B-spline window. By applying the analytic differentiation approach, we could reduce the required runtime to $t \approx 0.012$ seconds. □

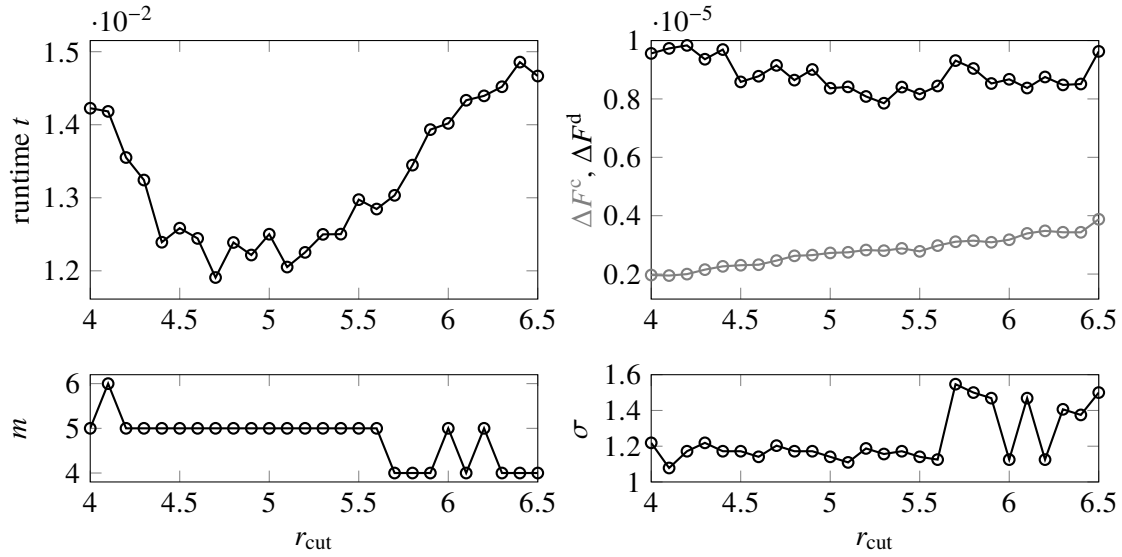


Figure 5.14: Achieved runtimes (in parallel with 4 cores) as well as the resulting rms force errors for Particle system 3, where we applied the B-spline window with analytic differentiation. We also plot the tuned parameters m and σ with respect to r_{cut}.

5.3 P²NFFT with mixed periodic and open boundary conditions

As we have seen, the 3d-periodic case is analytically well understood and, consequently, the resulting approximation errors can be predicted very precisely, see Section 5.2. In case of mixed periodic or open boundary conditions, we assume that the particles are distributed in a sphere with radius $L > 0$ with respect to the nonperiodic directions and apply the fast summation approach, cf. Section 3.4, in order to approximate the involved nonperiodic functions by trigonometric polynomials.

In the following we introduce, discuss and apply a heuristic approach to tune the parameters in case of mixed periodic or open boundary conditions, based on the error analysis for 3d-periodic constraints. This has already been introduced in [NPP15a], where charged particle systems are considered, and in [Nes17], where first tests for mixed charge-dipole systems under mixed periodic and open boundary conditions are presented.

For 2d-periodic boundary conditions, the particle system may be transformed such that the two vectors indicating the periodic directions have the form

$$\boldsymbol{\ell}_1 = \begin{pmatrix} l_{11} \\ l_{12} \\ 0 \end{pmatrix} \text{ and } \boldsymbol{\ell}_2 = \begin{pmatrix} l_{21} \\ l_{22} \\ 0 \end{pmatrix},$$

cf. Lemma 2.10 and Algorithm 2.1. In order to apply the fast summation approach in one dimension, see Section 3.4.1, we additionally assume

$$-L \leq x_{ij,3} \leq L \ \forall i, j.$$

For 1d-periodic constraints, the particle system may be transformed such that a vector of the form

$$\boldsymbol{\ell}_1 := \begin{pmatrix} L_1 \\ 0 \\ 0 \end{pmatrix}$$

indicates the periodic direction, cf. Lemma 2.14 and Algorithm 2.2. In order to apply the fast summation approach for radial kernels in $d = 2$ dimensions, cf. Section 3.4.2, we write the positions of the particles in the form $\boldsymbol{x}_j = (x_{j,1}, \tilde{\boldsymbol{x}}_j)^\top \in \mathbb{R}^3$ with $\tilde{\boldsymbol{x}}_j \in \mathbb{R}^2$ and assume

$$\|\tilde{\boldsymbol{x}}_{ij}\| \leq L \ \forall i, j.$$

In case of open boundary conditions we assume

$$\|\boldsymbol{x}_{ij}\| \leq L \ \forall i, j$$

in order to apply the fast summation approach for radial kernels in $d = 3$ dimensions.

Oftentimes, the particles are assumed to be distributed in a primary box of size $L_1 \times L_2 \times L_3$. In such a case we may simply set

$$L := \begin{cases} L_3 & : \text{2d-perioidic constraints with respect to the first two dimensions,} \\ \sqrt{L_2^2 + L_3^2} & : \text{1d-perioidic constraints with respect to the first dimension,} \\ \sqrt{L_1^2 + L_2^2 + L_3^2} & : \text{open boundary conditions.} \end{cases} \tag{5.9}$$

Finally, the periodicity applied in all nonperiodic dimensions is denoted by, cf. Section 3.4,

$$h \geq 2L. \tag{5.10}$$

For 3d-periodic boundary conditions the grid size $\boldsymbol{M} \in 2\mathbb{N}^3$ is chosen such that

$$\frac{M_j}{\|\boldsymbol{\ell}_j\|} \approx \beta \quad \forall j \in \{1, 2, 3\},$$

where we denote by $\beta > 0$ the far field cutoff parameter. In case of mixed periodic and open boundary conditions we heuristically choose the mesh size $M \in 2\mathbb{N}$ applied in all nonperiodic directions such that

$$\frac{M}{h} \approx \beta, \tag{5.11}$$

i.e., the same resolution is used for periodic and nonperiodic dimensions.

We claim that, for fixed splitting parameter α, near field cutoff radius r_{cut} and far field cutoff β, we are able to achieve approximation errors of a comparable size among all types of periodic boundary conditions, provided that the degree of smoothness p as well as the period h are chosen large enough. Furthermore, we assume that the NFFT parameters such as the support parameter m and the oversampling factor σ, for instance, are kept constant.

Recall that we may rewrite the condition $h \geq 2L$ via

$$L = h(1/2 - \varepsilon_{\text{B}}) \iff \varepsilon_{\text{B}} = \frac{h/2 - L}{h} = \frac{1}{2} - \frac{L}{h} \in [0, 1/2),$$

where we denote by $\varepsilon_B \in [0, 1/2)$ the regularization parameter, see also (4.52). For $\varepsilon_B = 0$ we have $h = 2L$, i.e., the nonperiodic function is just truncated at $x = \pm L$ and continued periodically, see Variant II in Section 3.4.1. In this case, equation (5.11) reads as

$$\frac{M_0}{2L} \approx \beta \iff M_0 \approx 2\beta L, \tag{5.12}$$

i.e., the mesh size which would be obtained without adding a regularization at the boundary of the interval is denoted by M_0.

For $\varepsilon_B > 0$ we add some additional $P \in 2\mathbb{N}$ grid points to end up with the same resolution in the nonperiodic dimensions, i.e., we set $M := M_0 + P$ to obtain

$$\frac{M}{M_0} = \frac{M_0 + P}{M_0} \approx \frac{h}{2L}. \tag{5.13}$$

In order to choose the parameters appropriately in case of mixed periodic or open boundary conditions we suggest to proceed as follows.

Algorithm 5.2 (Mesh size and regularization parameter)**.**
Input: Desired far field cutoff parameter $\beta > 0$.

i) Set $M_j \approx \beta \|\boldsymbol{\ell}_j\|$ via (5.1) for all periodic directions.

ii) Choose $M_0 \in 2\mathbb{N}$ such that $M_0 \approx 2\beta L$, i.e., round up to the next even natural number.

iii) Specify an additional number $P \in 2\mathbb{N}$ of grid points, provided for the regularization domain.

iv) The number M of mesh points used in the nonperiodic dimensions is given by $M := M_0 + P$.

v) Set

$$h := 2L \cdot \frac{M_0 + P}{M_0} \quad \text{or rather} \quad \varepsilon_B := \frac{1}{2} - \frac{M_0}{2(M_0 + P)}, \tag{5.14}$$

in order to apply the P²NFFT method, see Algorithm 4.3.

□

In the following we present a few numerical results for open as well as mixed periodic constraints. We will show that the method works for all the considered types of periodic boundary conditions. In the discussed numerical experiments, we tuned the parameters based on the 3d-periodic case heuristically, as discussed above.

Open boundary conditions

We start with the case of open boundary conditions, since reference data are easily accessible by computing the underlying finite sums, as for example (2.63), directly.

An essential ingredient for nonperiodic boundary conditions is the size of the regularization domain, which is controlled by the number of additional mesh points $P \in 2\mathbb{N}$, see (5.13), or rather the regularization parameter $\varepsilon_B \in [0, 1/2)$, see (5.14). The following example gives an idea on how the choice of the number of grid points P, spent for the regularization, influences the resulting approximation errors.

Example 5.18. We considered Particle system 3 subject to open boundary conditions and computed exact reference data via direct evaluation. Furthermore, we fixed the near field cutoff radius $r_{\text{cut}} := 5$ and the far field cutoff $\beta := 1.6$.

Note that in case of 3d-periodic boundary conditions this would result in the mesh size $\boldsymbol{M} = (32, 16, 16)^\top$. For open boundary conditions we obtain

$$L = \sqrt{20^2 + 10^2 + 10^2} = 10\sqrt{6} \approx 24.49$$

and

$$2\beta L \approx 78.38 \Rightarrow M_0 := 80.$$

In order to apply the P^{2}NFFT method subject to open boundary conditions, we added $P \in \{10, 12, 14, 16\}$ additional grid points for the regularization domain. The corresponding finally applied mesh sizes are given by $\boldsymbol{M} = (M, M, M)^\top$ with $M = M_0 + P = 80 + P$, respectively. The resulting values for h and ε_{B} are summarized in the following table.

P	M	h	ε_{B}
10	90	55.11	0.056
12	92	56.34	0.065
14	94	57.56	0.074
16	96	58.79	0.083

Furthermore, we used no oversampling, i.e., we set $\sigma := 1$, applied the support parameter $m := 5$ and the smoothness parameter $p := 12$. In Figure 5.15 we plot the achieved rms force errors with respect to the splitting parameter α. The P^{2}NFFT method has also been applied to the same particle system subject to 3d-periodic constraints, where we used the same set of parameters ($r_{\text{cut}} = 5$, $\beta = 1.6$, $m = 5$, $\sigma = 1$ and varying splitting parameter α). For comparison we also plot the achieved rms force errors under 3d-periodic boundary conditions (red lines). In case of open boundary conditions, better approximations are achieved the more grid points P for the regularization are added. For $P = 16$ the error curves almost coincide with those obtained for 3d-periodic boundary conditions.

Assuming that the errors of the fast summation approach for radial kernels in higher dimensions underlie similar principles as in the one-dimensional case, the considerations presented in Remark 3.23 give an explanation of the observations. □

Example 5.19. We consider again Particle system 3. Note that the maximum distance between two particles is given by $L = \sqrt{20^2 + 10^2 + 10^2} \approx 24.49$, i.e., the applied period h will be larger or equal to $2L \approx 48.99$, as introduced in (5.10). For 3d-periodic constraints the periods $L_1 = 20$, $L_2 = 10$ and $L_3 = 10$ regarding the single dimensions are much smaller. In other words, for a fixed far field cutoff $\beta > 0$ the computations in Fourier space are supposed to be more expensive for open boundary conditions. For this reason, the optimal parameters regarding computational costs will adopt fundamentally different values concerning the two considered types of boundary conditions.

In a numerical test we considered different near field cutoff radii r_{cut} and tuned the splitting parameter α as well as the far field cutoff β, as described in Section 2.3.4. To this end, we applied again the basic accuracy $\varepsilon := 10^{-5}$. Analogously to Example 5.13, possible combinations of the support parameter m and the oversampling factor σ were computed for each considered near field cutoff radius r_{cut}. For this purpose, we applied the error estimates concerning the B-spline window function under 3d-periodic boundary conditions. For each r_{cut} we finally kept the combination yielding the shortest computation time.

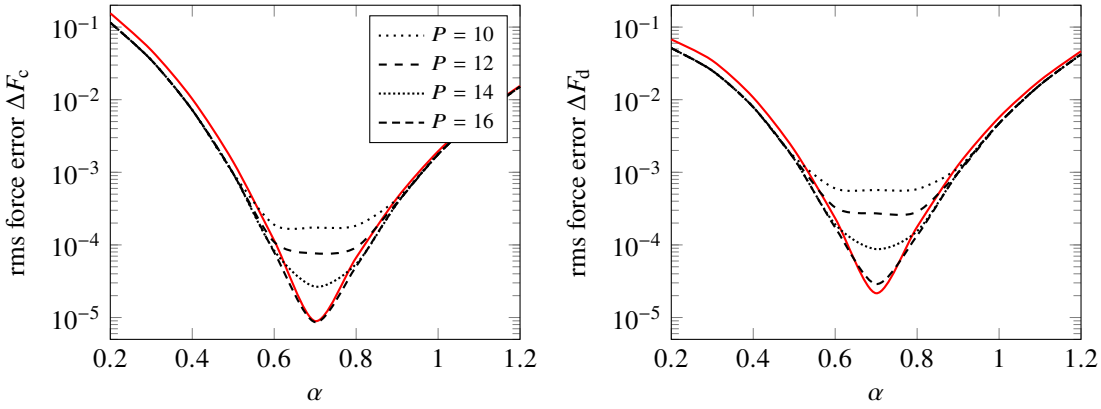


Figure 5.15: Achieved rms force errors for different values of P (see legend), where we considered Particle system 3 subject to open boundary conditions. The P²NFFT method has been applied with the B-spline window function, where we set $r_{\text{cut}} := 5$, $\beta := 1.6$, $\sigma := 1$ and $m := 5$. The corresponding errors for 3d-periodic constraints are represented by the red lines.

In order to compute the regularized kernel functions, we fixed the degree of smoothness $p := 12$. The number P of additional grid points spent for the regularization domain was increased step by step until the rms force deviation fell below ε. We obtained

$$P = \begin{cases} 20 & : 8 \leq r_{\text{cut}} \leq 12, \\ 18 & : 12 < r_{\text{cut}} \leq 16. \end{cases}$$

In Figure 5.16 we plot the achieved runtimes t, the measured rms force errors ΔF_c, ΔF_d as well as the optimal support parameters m and oversampling factors σ. By applying the proposed heuristic parameter choice we obtained somewhat larger approximation errors compared to the 3d-periodic case, cf. Figure 5.10. Nevertheless, we could always achieve rms force errors approximately of the size $\varepsilon = 10^{-5}$. We remark that the considered near field cutoff radii r_{cut} are not reasonable to be applied in case of very small particle systems, as the one considered. A direct computation of the electrostatic interactions is supposed to be much faster. However, with regard to an application of the same parameters in case of larger particle systems, such near field cutoff radii r_{cut} may be reasonable. An example is discussed in the following.

The tuned parameters have been applied to a larger particle system containing $N_c = 76800$ charges and $N_d = 76800$ dipoles, distributed in a box of the size $80 \times 80 \times 80$. As explained before, the mesh size was scaled up according to the dimensions of the given primary simulation box based on (5.12). In contrast, the number of additional grid points $P \in \{18, 20\}$ in (5.13) was kept constant. Note that keeping the number of additional grid points P constant results in a decrease of the regularization parameter ε_B for growing system size. The parameter choice for particle systems of growing size is reconsidered later on in Example 5.22. Concrete parameter sets are listed in Table 5.8, where the cutoff radius $r_{\text{cut}} := 11$ has been fixed.

The runtimes as well as rms force errors, achieved for the large particle system with $N_c = N_d = 76800$ under open boundary conditions, are summarized in Figure 5.17. We obtained somewhat larger approximation errors compared to the small particle system, cf. Figure 5.16. In this concrete example, the required computation time is minimized for $r_{\text{cut}} \approx 11$, whereas $r_{\text{cut}} \approx 6.5$ was found to be optimal in

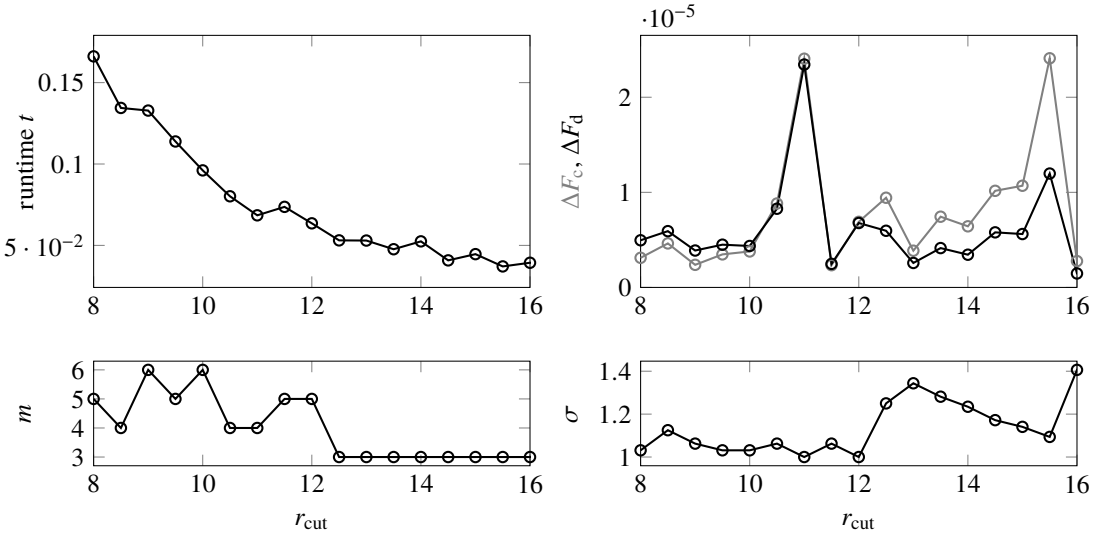


Figure 5.16: Achieved runtimes (in parallel with 4 cores), rms force errors as well as tuned optimal support parameters m and oversampling factors σ. We consider Particle system 3 subject to open boundary conditions. The P^{2}NFFT method has been applied by using the B-spline window with i$\boldsymbol{k}$-differentiation.

case of 3d-periodic boundary conditions, see Figure 5.11. We achieved an optimal runtime of $t \approx 11.5$ seconds in case of open boundary conditions, whereas $t \approx 4.5$ seconds were achieved with the B-spline window under 3d-periodic constraints. □

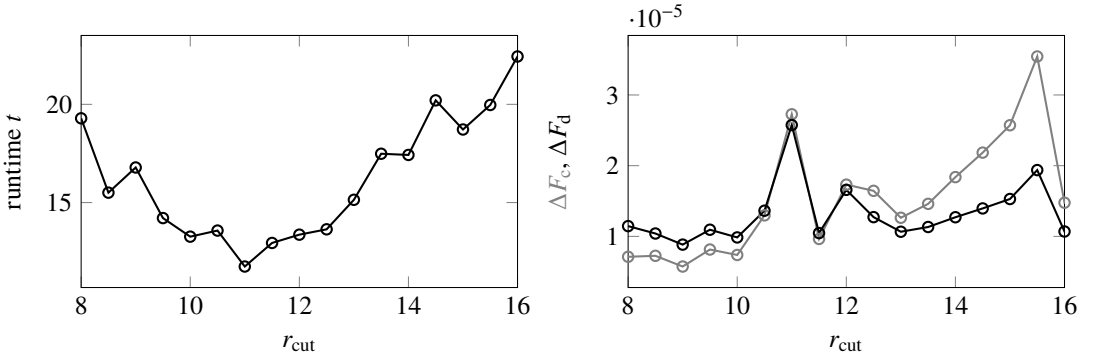


Figure 5.17: Achieved runtimes (in parallel with 4 cores) and rms force errors for the particle system containing $N_c = 76800$ charges and $N_d = 76800$ dipoles, distributed in a box of size $80 \times 80 \times 80$. Boundary conditions: open.

Mixed periodic boundary conditions

In the following example we consider Particle system 3 subject to 1d-periodic as well as 2d-periodic boundary conditions. In the discussed numerical experiments we again considered different near field cutoff radii r_{cut} and tuned the splitting parameter α, the far field cutoff β, the support parameter m

as well as the oversampling factor σ based on the error estimates for 3d-periodic constraints. Again, for each r_{cut} we measured the runtimes and kept the best combination of m and σ with respect to computational costs. We considered only the B-spline window with i$\boldsymbol{k}$-differentiation.

Example 5.20. We assume nonperiodic constraints with respect to the last two dimensions, i.e., we have $L = 10\sqrt{2} \approx 14.14$.

Analogously to Example 5.19, we applied the degree of smoothness $p := 12$ in order to compute the regularized kernel functions. The number of additional grid points P, spent for the regularization domain, was again increased step by step until the rms force deviation fell below the applied basic accuracy $\varepsilon := 10^{-5}$. For each considered near field cutoff radius, $P \in \{20, 22\}$ additional grid points for the regularization domain were required.

Note that, in contrast to the case of open boundary conditions, no exact reference data were available. In order to compute reference data for Particle system 3 under 1d-periodic boundary conditions, we used the direct summation algorithm in ScaFaCoS, which evaluates the underlying infinite sums over $\boldsymbol{n} \in \mathbb{Z} \times \{0\}^2$, cf. (2.3), directly via truncation. For this purpose, we applied a truncation scheme of the form $\boldsymbol{n} \in \{-C, C\} \times \{0\}^2$ and increased C step by step until the rms force deviation reached a value smaller than 10^{-10}. Finally, the reference data were computed with the cutoff parameter $C = 11000$. In Figure 5.18 we plot the achieved runtimes as well as rms force errors for varying near field cutoff radius r_{cut}. For overview purposes, the tuned parameters α, β, m, σ and P are not depicted. Again, the measured errors are approximately of the size $\varepsilon = 10^{-5}$.

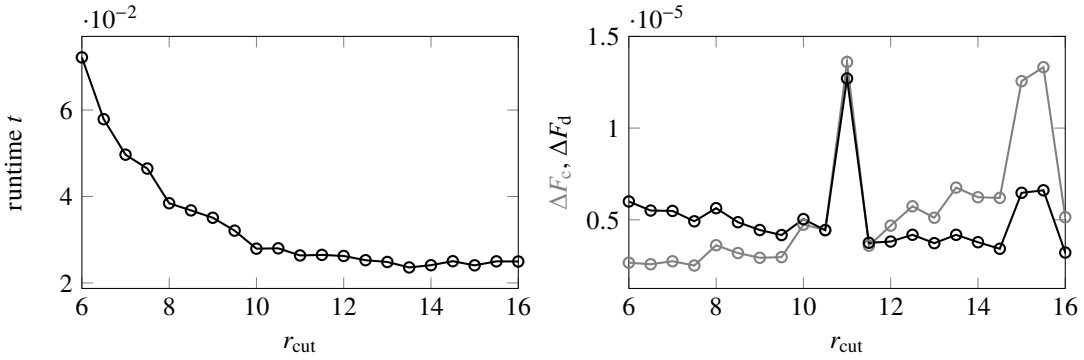


Figure 5.18: Achieved runtimes (in parallel with 4 cores) and rms force errors for Particle system 3 under 1d-periodic boundary conditions. The P²NFFT method has been applied by using the B-spline window with i$\boldsymbol{k}$-differentiation.

In order to get an impression of the optimal parameters for large systems, the tuned parameter sets have also been applied to the large particle system with $N_{\text{c}} = 76800$ charges and $N_{\text{d}} = 76800$ dipoles subject to 1d-periodic boundary conditions. The measured runtimes and rms force errors are depicted in Figure 5.19. Here, a near field cutoff radius $r_{\text{cut}} \approx 8.5$ seems to be reasonable choice in order to minimize the computational costs.

Reference data of a higher accuracy for the large particle system were obtained based on the tuned parameters for $r_{\text{cut}} := 10$, which are $\alpha \approx 0.33$, $\boldsymbol{M} = (56, 158 + 22, 158 + 22)$, $m = 4$ and $\sigma \approx 1.14$. In order to increase the accuracy we applied the following set of parameters: $r_{\text{cut}} = 12$, $\alpha \approx 0.33$, $\boldsymbol{M} = (112, 316 + 44, 316 + 44)$, $m = 6$, $\sigma \approx 1.39$. □

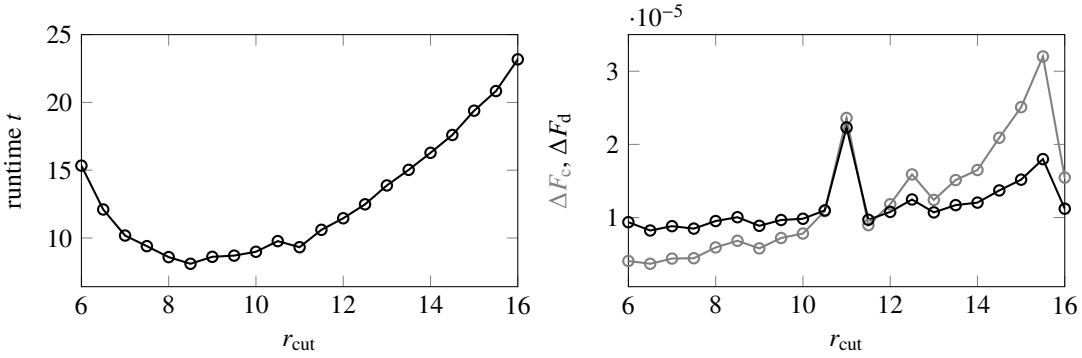


Figure 5.19: Achieved runtimes for the particle system containing $N_{\text{c}} = 76800$ charges and $N_{\text{d}} = 76800$ dipoles, distributed in a box of size $80 \times 80 \times 80$. Boundary conditions: 1d-periodic.

Example 5.21. Again, we start by considering Particle system 3 subject to 2d-periodic constraints, where we assume periodicity with respect to the first two dimensions, i.e., we have $L = L_3 = 10$.

We proceeded analogously to the 1d-periodic case and tuned the parameters based on the error estimates for 3d-periodic boundary conditions. In Figure 5.20 we plot the resulting runtimes and rms force errors for varying r_{cut}. Again, the tuned values for α, β, m, σ and P are not depicted, for overview purposes. The additional number of grid points for the regularization domain was tuned to the value $P = 16$, for each considered near field cutoff r_{cut}. With the proposed tuning we could achieve rms force errors smaller than $\varepsilon = 10^{-5}$, in each case.

Reference data of a higher accuracy have been computed as follows. We fixed the near field cutoff $r_{\text{cut}} := 10$, the degree of smoothness $p := 12$ and the basic accuracy $\varepsilon := 10^{-12}$. Then, we applied the above described parameter tuning, appropriately adapted for the purpose of applying the P^{2}NDFT Ewald method under 2d-periodic boundary conditions. Finally, the reference data were computed via the P^{2}NDFT Ewald method with parameters $r_{\text{cut}} = 10$, $\alpha \approx 0.54$, $p = 12$, $\boldsymbol{M} = (38, 20, 38 + 54)$ and $\varepsilon_{\text{B}} \approx 0.293$.

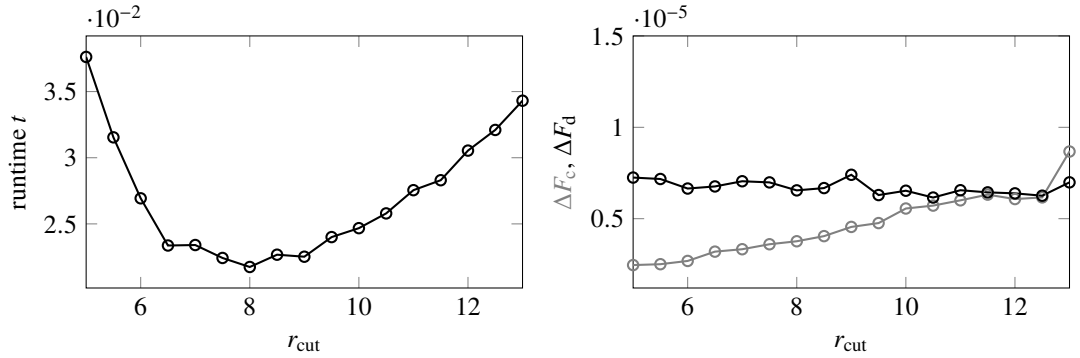


Figure 5.20: Achieved runtimes (in parallel with 4 cores) and rms force errors for Particle system 3 under 2d-periodic boundary conditions. The P^{2}NFFT method has been applied by using the B-spline window with i$\boldsymbol{k}$-differentiation.

Again, the same parameters have been applied to the large particle system with $N_{\text{c}} = 76800$ charges and $N_{\text{d}} = 76800$ dipoles subject to 2d-periodic boundary conditions. The measured runtimes and rms

force errors are depicted in Figure 5.21. An almost optimal runtime is achieved for $r_{\text{cut}} \approx 8$, for both the small and the large particle system.

Reference data of a higher accuracy for the large particle system have been obtained based on the tuned parameters for $r_{\text{cut}} := 10$, which are $\alpha \approx 0.33$, $\boldsymbol{M} = (56, 56, 112 + 16)$, $m = 4$ and $\sigma \approx 1.14$. In order to increase the accuracy we applied the following set of parameters: $r_{\text{cut}} = 12$, $\alpha \approx 0.33$, $\boldsymbol{M} = (112, 112, 224 + 32)$, $m = 6$, $\sigma \approx 1.39$. □

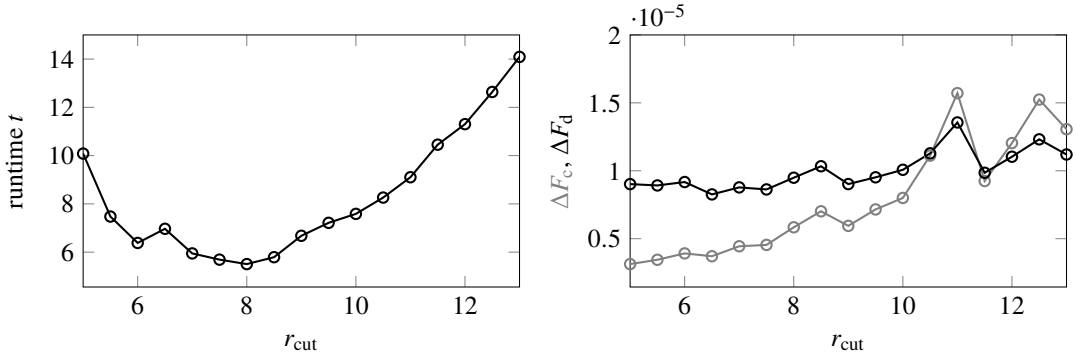


Figure 5.21: Achieved runtimes for the particle system containing $N_{\text{c}} = 76800$ charges and $N_{\text{d}} = 76800$ dipoles, distributed in a box of size $80 \times 80 \times 80$. Boundary conditions: 2d-periodic.

Example 5.22. We reconsider the first part of Example 5.15, where we applied a fixed set of parameters to particle systems of increasing size, see Table 5.4, subject to 3d-periodic boundary conditions. We applied $r_{\text{cut}} := 6.5$ in case of the B-spline window, which turned out to be a good choice for both the small particle system with $N = 600$ as well as the one with $N = 153600$ particles.

The same test was executed for open and mixed periodic boundary conditions, respectively, where we applied the near field cutoff radii r_{cut} minimizing the runtimes for the large particle system with $N = 153600$ particles, see Figures 5.17, 5.19 and 5.21. We summarize the applied parameters in case of the different periodic boundary conditions in Table 5.7.

	r_{cut}	α	β	m	σ	P
3d-periodic	6.5	0.54	1.18	4	1.28	
2d-periodic	8.0	0.43	0.92	4	1.16	16
1d-periodic	8.5	0.40	0.85	5	1.02	22
0d-periodic	11.0	0.30	0.61	4	1.00	20

Table 5.7: Applied parameters for the different types of periodic boundary conditions.

In order to apply the same parameters to the larger particle systems, the mesh size $\boldsymbol{M} \in 2\mathbb{N}^3$ was scaled according to the dimensions of the box, as explained above. The number of additional grid points P, applied in the regularization domain, was kept constant among all particle systems. Hence, the regularization parameter ε_{B} decreases with increasing system size.

In Table 5.8 we list the achieved rms force errors as well as the measured runtimes in case of open boundary conditions. We also list the values L, M_0, P, $M = M_0 + P$, h and ε_{B}, as introduced by (5.9), (5.12), (5.13) and (5.14). Finally, the mesh of precomputed Fourier coefficients is given by $\boldsymbol{M} = (M, M, M)$, since the period h is applied with respect to all three dimensions. Note that we

also list the computation times achieved by computing the underlying finite sums directly, which we denote by t_{direct}. In this specific example, we are faster by applying the P^{2}NFFT method if more than 10^4 particles are present. Furthermore, the achieved rms force errors are approximately of the size $\varepsilon = 10^{-5}$ and almost constant among the considered particle systems.

N	L	M_0	P	M	h	ε_B	ΔF_c	ΔF_d	t	t_{direct}
600	24.49	30	20	50	81.65	0.200	2.41e-05	2.34e-05	0.07	3e–3
2400	34.64	44	20	64	100.77	0.156	1.60e-05	1.48e-05	0.19	0.04
9600	108.25	74	20	94	152.43	0.106	2.28e-05	2.14e-05	0.85	0.61
38400	158.57	120	20	140	228.62	0.071	2.73e-05	2.54e-05	3.19	9.63
153600	235.15	170	20	190	309.73	0.053	2.73e-05	2.57e-05	11.49	154.43

Table 5.8: Tuned parameters, achieved runtimes t (in parallel with 4 cores) and rms force errors for all the considered particle systems. Boundary conditions: open.

The results concerning 1d-periodic and 2d-periodic boundary conditions are listed in the Tables 5.9 and 5.10, respectively. We assumed periodic boundary conditions regarding the first or the first two dimensions, respectively, for each particle system considered. Again, we also list the values for L, M_0, P, the size of the precomputed Fourier mesh $\boldsymbol{M} \in 2\mathbb{N}^3$, the period h applied in the nonperiodic dimensions and the regularization parameter ε_B. Note that the final size of the computed (inverse) FFTs is $\boldsymbol{M}_o = \sigma\boldsymbol{M}$ with the oversampling factors σ listed in Table 5.7.

N	L	M_0	P	$\boldsymbol{M}$	h	ε_B	ΔF_c	ΔF_d	t
600	14.14	26	22	(18, 48, 48)	52.22	0.229	3.19e-06	4.90e-06	0.04
2400	28.28	50	22	(18, 72, 72)	81.46	0.153	4.53e-06	7.05e-06	0.13
9600	44.72	76	22	(34, 98, 98)	115.33	0.112	5.93e-06	8.80e-06	0.53
38400	56.57	98	22	(68, 120, 120)	138.54	0.092	6.28e-06	9.13e-06	2.00
153600	113.14	194	22	(68, 216, 216)	251.93	0.051	6.83e-06	1.00e-05	8.09

Table 5.9: Tuned parameters, achieved runtimes t (in parallel with 4 cores) and rms force errors for all the considered particle systems. Boundary conditions: 1d-periodic.

N	L	M_0	P	$\boldsymbol{M}$	h	ε_B	ΔF_c	ΔF_d	t
600	10	20	16	(20, 10, 36)	36.00	0.222	3.77e-06	6.55e-06	0.02
2400	20	38	16	(20, 20, 54)	56.84	0.148	4.38e-06	7.49e-06	0.08
9600	20	38	16	(38, 38, 54)	56.84	0.148	4.86e-06	8.19e-06	0.30
38400	40	74	16	(74, 38, 90)	97.30	0.089	5.47e-06	8.98e-06	1.33
153600	80	148	16	(74, 74, 164)	177.30	0.049	5.84e-06	9.49e-06	5.52

Table 5.10: Tuned parameters, achieved runtimes t (in parallel with 4 cores) and rms force errors for all the considered particle systems. Boundary conditions: 2d-periodic.

A comparison of the achieved runtimes for all types of periodic boundary conditions is given in Figure 5.22. As expected, the required runtimes increase the more dimensions subject to nonperiodic boundary conditions are involved. This is due to the fact that the proposed parameter tuning results in larger periods and larger mesh sizes in all dimensions subject to nonperiodic constraints. However, an appropriate choice of the near field cutoff radius r_{cut} can compensate the large difference between

the required mesh sizes in Fourier space. Consequently, the value of the optimal near field cutoff radius r_{cut} increases with decreasing number of periodic dimensions involved, see Table 5.7. In case of 2d-periodic boundary conditions we only have a small increase in runtime compared to 3d-periodic constraints, for instance. The achieved increase in runtime compared to the 3d-periodic case is visualized in Figure 5.23, where we plot the runtimes for open and mixed periodic constraints, divided by the achieved runtimes for 3d-periodic boundary conditions. □

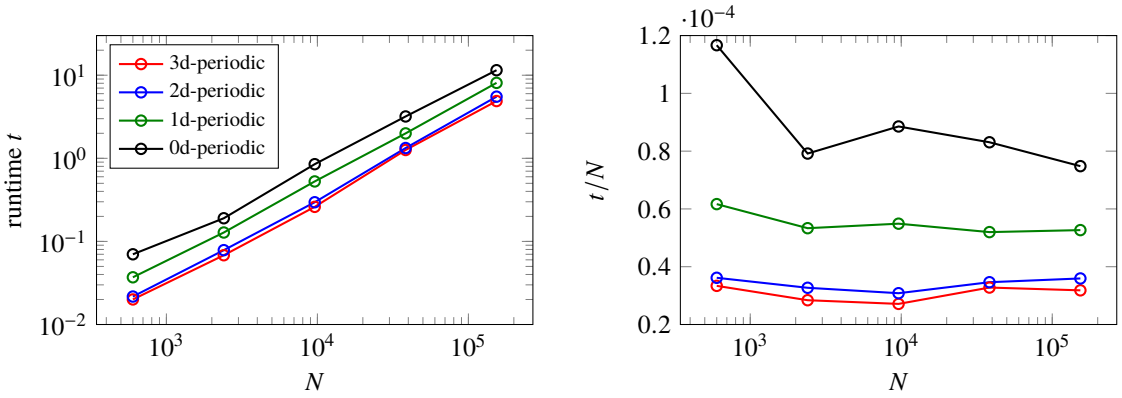


Figure 5.22: Comparison of the achieved runtimes (in parellel with 4 cores) for the different types of periodic boundary conditions. The achieved rms force accuracy is $\approx 10^{-5}$.

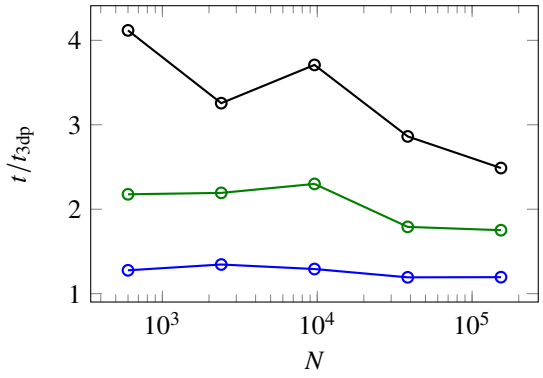


Figure 5.23: Runtime t for open and mixed periodic constraints divided by the runtime $t_{3\text{dp}}$ achieved for 3d-periodic boundary conditions. The achieved rms force accuracy is $\approx 10^{-5}$.

6 Conclusion

This work mainly dealt with the application of nonequispaced fast Fourier transforms within the scope of computing electrostatic interactions in particle systems. Main quantities of interest are the electrostatic energies, forces and torques arising in particle systems consisting of a mixture of point charges and point dipoles.

The presented method, which makes use of various NFFT modules, is named particle-particle NFFT (P^2NFFT) and is part of the publicly available software library ScaFaCoS. It is based on the widely used Ewald summation approach which splits the underlying infinite and poorly converging sums into short range and long range parts. An efficient computation of the short range parts is possible via truncation and is realized in the same way for various types of periodic boundary conditions. This is delegated to the provided ScaFaCoS near field solver. In case of the potentials we have

$$\phi^{\text{short}}(j) = \sideset{}{'}\sum_{\boldsymbol{n}\in\mathcal{S}} \sum_{i=1}^{N} \xi_i \frac{\operatorname{erfc}(\alpha\|\boldsymbol{x}_{ij}+\boldsymbol{L}\boldsymbol{n}\|)}{\|\boldsymbol{x}_{ij}+\boldsymbol{L}\boldsymbol{n}\|},$$

where the operators ξ_j have been introduced in order to handle systems containing a mixture of point charges and point dipoles, cf. (2.4). The primary simulation box has to be extended periodically according to the given boundary conditions, which is achieved by setting $\mathcal{S} := \mathbb{Z}^\rho \times \{0\}^{3-\rho}$, where $\rho \in \{0,1,2,3\}$ denotes the number of directions subject to periodic boundary conditions.

The long range parts are approximated in Fourier space. This is realized based on appropriate representations of the long range parts in Fourier space, which strongly depends on the underlying periodicity. However, we finally end up with the same modular structure for all types of periodic boundary conditions. In case of the electrostatic potentials, the Fourier space contributions are finally approximated by an expression of the form

$$\begin{aligned}\phi^{\text{F}}(j) &\approx \sum_{\boldsymbol{\kappa}\in\mathcal{I}_{\boldsymbol{M}}} c(\boldsymbol{\kappa}) \left(\sum_{i=1}^{N} \xi_i \,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{\kappa}^\top\boldsymbol{\Lambda}^{-1}\boldsymbol{x}_i}\right) \mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{\kappa}^\top\boldsymbol{\Lambda}^{-1}\boldsymbol{x}_j} \\ &= \sum_{\boldsymbol{\kappa}\in\mathcal{I}_{\boldsymbol{M}}} c(\boldsymbol{\kappa}) \left(\sum_{i=1}^{N_{\text{c}}} q_i \,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{\kappa}^\top\boldsymbol{\Lambda}^{-1}\boldsymbol{x}_i} + \sum_{i=N_{\text{c}}+1}^{N} \boldsymbol{\mu}_i^\top \nabla_{\boldsymbol{x}_i} \,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{\kappa}^\top\boldsymbol{\Lambda}^{-1}\boldsymbol{x}_i}\right) \mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{\kappa}^\top\boldsymbol{\Lambda}^{-1}\boldsymbol{x}_j}, \end{aligned} \tag{6.1}$$

with a regular matrix $\boldsymbol{\Lambda} \in \mathbb{R}^{3\times3}$, which depends on the shape of the primary simulation box and the applied periodic boundary conditions. The Fourier coefficients $c(\boldsymbol{\kappa})$ are known analytically in case of 3d-periodic constraints and computed numerically for mixed periodic and open boundary conditions, as summarized by the following table.

	$\boldsymbol{\kappa}$	computation of $c(\boldsymbol{\kappa})$
3d-periodic	$= \boldsymbol{k}$	$c(\boldsymbol{\kappa}) = (\pi V)^{-1}\hat{\psi}(\boldsymbol{k})$, see (2.30) and (2.31)
2d-periodic	$= (\boldsymbol{k}, l)$	$c(\boldsymbol{\kappa}) = \hat{\psi}^{\text{p2}}(\boldsymbol{k}, l)$, fast summation approach ($d=1$) for each $\boldsymbol{k} \in \mathcal{I}_{(M_1,M_2)}$
1d-periodic	$= (k, \boldsymbol{l})$	$c(\boldsymbol{\kappa}) = \hat{\psi}^{\text{p1}}(k, \boldsymbol{l})$, fast summation approach ($d=2$) for each $k \in \mathcal{I}_{M_1}$
0d-periodic	$= \boldsymbol{l}$	$c(\boldsymbol{\kappa}) = \hat{\psi}^{\text{p0}}(\boldsymbol{l})$, fast summation approach ($d=3$)

The indices $\boldsymbol{k}$ or k are introduced by applying the analytic Fourier transform in the periodic dimensions, and the indices $\boldsymbol{l}$ or l are introduced by applying the FFT in the nonperiodic directions, after sampling the constructed periodic functions.

In case of mixed periodic or open boundary conditions, a precomputation step is required in order to periodize the underlying nonperiodic functions and approximate them by trigonometric polynomials, following the NFFT based fast summation approach. The main portions are treated based on the so called regularization technique, extending the nonperiodic functions appropriately by polynomials in order to construct smooth periodic functions. In contrast, the well localized interaction portions are treated based on the Poisson summation formula.

It should be mentioned that the precomputations have to be conducted only once, even if the algorithm is applied in simulations involving several time steps. This is due to the fact that the precomputation step does only depend on the dimensions of the primary simulation box and not on the particle positions. The precomputations have only to be repeated if the simulation box is increased, which becomes necessary if particles are leaving the box.

Finally, the Fourier space parts of the potentials are computed based on (6.1) as follows. First, the expressions in brackets are approximated via the adjoint NFFT and the adjoint gradient NFFT. After multiplying the obtained values by the Fourier coefficients $c(\boldsymbol{\kappa})$, the outer sums are finally approximated based on the NFFT. In order to compute the acting forces, the so called gradient NFFT is applied in case of charged particles and the Hessian NFFT is used in case of dipoles. The Hessian NFFT and the adjoint gradient NFFT are only required if dipoles are present in the given system and have thus been introduced for the purpose of extending the original P^{2}NFFT for charged particle systems to mixed charge-dipole systems.

In the field of molecular dynamics simulations, the rms force error is a very common measure of accuracy. Hence, we investigated the rms force errors introduced by the P^{2}NFFT method for charge-dipole systems. Basically, there are two different types of errors to be considered, namely the truncation errors resulting from truncating the Ewald sums and, secondly, the approximation errors originating from applying the different NFFT modules. Therefor, the theory of truncation errors in the Ewald summation formulas under 3d-periodic boundary conditions has been studied and extended appropriately to mixed charge-dipole systems. The presented numerical examples confirm that a reliable prediction of the Ewald truncation errors as well as an appropriate tuning of the involved parameters with respect to accuracy is therewith possible.

In addition, the NFFT related approximation errors were taken into consideration. Corresponding error estimates have already been proposed for the classical P^{3}M method as well as for the dipolar P^{3}M algorithm. These estimates have been adapted to the P^{2}NFFT and appropriately extended to charge-dipole interactions, where we allow for the presence of an oversampling factor, which does not exist in the scope of classical particle-mesh methods. We considered both differentiation in Fourier space ($\mathrm{i}\boldsymbol{k}$) and the analytic differentiation approach. The resulting estimates concerning $\mathrm{i}\boldsymbol{k}$-differentiation are valid for arbitrary compactly supported window functions, whereas the estimates concerning analytic differentiation require differentiable window functions. The presented numerical examples confirm that high accuracies may be achieved based on both, $\mathrm{i}\boldsymbol{k}$- and analytic differentiation. If the shape parameter of the Bessel window is chosen appropriately, we may achieve smaller approximation errors than in case of the B-spline window for sufficiently large Ewald splitting parameter α.

In terms of estimating the resulting approximation errors and tuning the involved parameters, we concentrated on studying the $\mathrm{i}\boldsymbol{k}$-differentiation approach. In case of orthorhombic box shapes, the errors may be predicted very efficiently and numerically stable based on an approximation by exponential terms. For triclinic box shapes the underlying sums of length $|\mathcal{I}_{\boldsymbol{M}}|$ may be evaluated directly. Based

on the presented estimates, all involved parameters can be tuned with respect to accuracy in advance, for which only the near field cutoff radius r_{cut} has to be specified. The numerical examples show that the errors are estimated properly and that the parameter tuning provides feasible parameter sets. In order to find the optimal parameter combination with respect to runtime, a further optimization with respect to r_{cut} has to be done. This especially depends on the underlying architecture.

In terms of the NFFT window function we restricted our considerations to the widely used B-spline window and the Bessel window. In case of analytic differentiation we discussed proper estimates for the B-spline window function in connection with orthorhombic box shapes. As indicated before, the universal estimates are not readily applicable for the Bessel window since it is not differentiable at the limits of its definition area. This gives raise to continuative investigations. Furthermore, the handling of general triclinic box shapes and an investigation of the errors originating from the computation of undesired self forces may be subject of an extended analysis concerning the analytic differentiation approach.

The presented numerical results indicate that it is often advantageous in terms of computational costs to spend some oversampling while using a smaller support for the NFFT window function. The Bessel window function is oftentimes the better choice in terms of accuracy and efficiency, since a smaller oversampling factor or a smaller support parameter is required. However, which window function will be optimal regarding computational costs will always depend on the size of the given particle system as well as on the required accuracy, for instance. If all involved parameters have been tuned for a small particle system, the same set of parameters may be applied also to larger particle systems possessing the same charge and dipole densities. This rather simple approach results in a remarkable arithmetic complexity of $O(N \log N)$, which has been verified based on numerical experiments as well as theoretical considerations.

We also presented numerical results for mixed periodic and open boundary conditions, where we only applied the B-spline window function. Whereas adequate error estimates have been derived for 3d-periodic constraints, there are no equivalent estimates in the mixed periodic and open settings, since the underlying Fourier coefficients are not known analytically. Thus, a reliable prediction of the resulting approximation errors as well as an automated tuning of all involved parameters is more challenging.

In this work we presented a heuristic approach to tune the parameters for mixed periodic and open boundary conditions, which is based on the error analysis concerning the 3d-periodic setting. For a fixed set of parameters, only the applied mesh size in Fourier space has to be adapted for the different types of periodic boundary conditions. A larger mesh size is required in dimensions subject to open constraints. Consequently, the computation of the Fourier space parts is computationally more expensive in case of mixed periodic or open boundary conditions. In contrast, the computational effort required for the near field computations is almost the same as under 3d-periodic boundary conditions. Hence, the runtime is minimized for fundamentally different near field cutoff radii r_{cut}, depending on the applied periodic boundary conditions. This has been verified based on numerical examples. If the near field cutoff is chosen appropriately, the P^{2}NFFT under 2d-periodic boundary conditions only requires slightly larger computation times than under 3d-periodic constraints, for instance. Moreover, the numerical results show that the proposed heuristic parameter tuning works reliably and that the resulting approximation errors are very close to the ones obtained under 3d-periodic boundary conditions.

Bibliography

[ABD+] A. Arnold, M. Bolten, H. Dachsel, F. Fahrenberger, F. Gähler, R. Halver, F. Heber, M. Hofmann, J. Iseringhausen, I. Kabadshow, O. Lenz, and M. Pippig. ScaFaCoS - Scalable Fast Coloumb Solvers. `http://www.scafacos.de`.

[AdJH02] A. Arnold, J. de Joannis, and C. Holm. Electrostatics in periodic slab geometries. I. *J. Chem. Phys.*, 117:2496, 2002.

[AFH+13] A. Arnold, F. Fahrenberger, C. Holm, O. Lenz, M. Bolten, H. Dachsel, R. Halver, I. Kabadshow, F. Gähler, F. Heber, J. Iseringhausen, M. Hofmann, M. Pippig, D. Potts, and G. Sutmann. Comparison of scalable fast methods for long-range interactions. *Phys. Rev. E*, 88:063308, 2013.

[AH02] A. Arnold and C. Holm. MMM2D: A fast and accurate summation method for electrostatic interactions in 2D slab geometries. *Comput. Phys. Commun.*, 148:327–348, 2002.

[AH05] A. Arnold and C. Holm. MMM1D: A method for calculating electrostatic interactions in one-dimensional periodic geometries. *J. Chem. Phys.*, 123:144103, 2005.

[Arn01] A. Arnold. Berechnung der elektrostatischen Wechselwirkung in 2d+h periodischen Systemen. Diplomarbeit, Johannes Gutenberg-Universität, Mainz, 2001.

[AS72] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions*. National Bureau of Standards, Washington, DC, USA, 1972.

[AW93] R. P. Agarwal and P. J. Y. Wong. *Error inequalities in polynomial interpolation and their applications*, volume 262 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1993.

[Bal14] V. Ballenegger. Communication: On the origin of the surface term in the Ewald formula. *The Journal of Chemical Physics*, 140(16), 2014.

[BCH11] V. Ballenegger, J. Cerdà, and C. Holm. Removal of spurious self-interactions in particle-mesh methods. *Computer Physics Communications*, 182(9):1919 – 1923, 2011.

[BCH12] V. Ballenegger, J. J. Cerdà, and C. Holm. How to convert SPME to P3M: Influence functions and error estimates. *J. Chem. Theory Comput.*, 8(3):936–947, 2012.

[Bey95] G. Beylkin. On the fast Fourier transform of functions with singularities. *Appl. Comput. Harmon. Anal.*, 2:363–381, 1995.

[BH86] J. Barnes and P. Hut. A hierarchical $O(n \log n)$ force-calculation algorithm. *Nature*, 324(6096):446–449, 1986.

[Bol08] M. Bolten. *Multigrid methods for structured grids and their application in particle simulation*. PhD thesis, Bergische Universität Wuppertal, Wuppertal, 2008.

[Bró04] A. Bródka. Ewald summation method with electrostatic layer correction for interactions of point dipoles in slab geometry. *Chem. Phys. Lett.*, 400:62–67, 2004.

[BS04] A. Bródka and P. Sliwinski. Three-dimensional Ewald method with correction term for a system periodic in one direction. *J. Chem. Phys.*, 120:5518–5523, 2004.

[Cai13] W. Cai. *Computational methods for electromagnetic phenomena*. Cambridge University Press, Cambridge, 2013. Electrostatics in solvation, scattering, and electron transport.

[CBH11] J. J. Cerdà, V. Ballenegger, and C. Holm. Particle-particle particle-mesh method for dipolar interactions: On error estimates and efficiency of schemes with analytical differentiation and mesh interlacing. *J. Chem. Phys.*, 135:184110, 2011.

[CBLH08] J. J. Cerdà, V. Ballenegger, O. Lenz, and C. Holm. P3M algorithm for dipolar interactions. *J. Chem. Phys.*, 129:234104, 2008.

[CDDL09] D. S. Cerutti, R. E. Duke, T. a. Darden, and T. P. Lybrand. Staggered Mesh Ewald: An extension of the Smooth Particle-Mesh Ewald method adding great versatility. *J. Chem. Theory Comput.*, 5(9):2322, 2009.

[CGH+96] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the LambertW function. *Adv. Comput. Math.*, 5(1):329–359, 1996.

[CT65] J. W. Cooley and J. W. Tukey. An algorithm for machine calculation of complex Fourier series. *Math. Comput.*, 19:297–301, 1965.

[Dac10] H. Dachsel. An error-controlled Fast Multipole Method. *J. Chem. Phys.*, 132(11):119901, 2010.

[DAH03] M. Deserno, A. Arnold, and C. Holm. Attraction and ionic correlations between charged stiff polyelectrolytes. *Macromolecules*, 36:249 – 259, 2003.

[DH98a] M. Deserno and C. Holm. How to mesh up Ewald sums. I. A theoretical and numerical comparison of various particle mesh routines. *J. Chem. Phys.*, 109:7678–7693, 1998.

[DH98b] M. Deserno and C. Holm. How to mesh up Ewald sums. II. An accurate error estimate for the Particle-Particle-Particle-Mesh algorithm. *J. Chem. Phys.*, 109:7694–7701, 1998.

[DHR07] H. Dachsel, M. Hofmann, and G. Rünger. Library support for parallel sorting in scientific computations. In *Proc. of the 13th International Euro-Par Conference*, volume 4641 of *LNCS*, pages 695–704. Springer, 2007.

[DK00] Z.-H. Duan and R. Krasny. An Ewald summation based multipole method. *J. Chem. Phys.*, 113:3492, 2000.

[dLPS80] S. W. de Leeuw, J. W. Perram, and E. R. Smith. Simulation of electrostatic systems in periodic boundary conditions. I. Lattice sums and dielectric constants. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 373:27–56, 1980.

[DR93] A. Dutt and V. Rokhlin. Fast Fourier transforms for nonequispaced data. *SIAM J. Sci. Stat. Comput.*, 14:1368–1393, 1993.

[DS99] A. J. W. Duijndam and M. A. Schonewille. Nonuniform fast Fourier transform. *Geophysics*, 64:539–551, 1999.

[DYP93] T. Darden, D. York, and L. Pedersen. Particle mesh Ewald: An $n\log(n)$ method for Ewald sums in large systems. *J. Chem. Phys.*, 98:10089–10092, 1993.

[EHL80] J. Eastwood, R. Hockney, and D. Lawrence. P3M3DP—The three-dimensional periodic particle-particle/ particle-mesh program. *Computer Physics Communications*, 19(2):215 – 261, 1980.

[EPB+95] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. A smooth particle mesh Ewald method. *J. Chem. Phys.*, 103:8577–8593, 1995.

[Ewa21] P. P. Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.*, 369:253–287, 1921.

[Fin94] D. Fincham. Optimisation of the Ewald sum for large systems. *Mol. Simul.*, 13:1–9, 1994.

[Fou03] K. Fourmont. Non equispaced fast Fourier transforms with applications to tomography. *J. Fourier Anal. Appl.*, 9:431–450, 2003.

[FS02] D. Frenkel and B. Smit. *Understanding molecular simulation: From algorithms to applications*. Computational science series. Academic Press, San Diego, San Francisco, New York, 2002.

[FS03] J. A. Fessler and B. P. Sutton. Nonuniform fast Fourier transforms using min-max interpolation. *IEEE Trans. Signal Process.*, 51:560–574, 2003.

[FS04] M. Fenn and G. Steidl. Fast NFFT based summation of radial functions. *Sampl. Theory Signal Image Process.*, 3:1–28, 2004.

[GGB00] A. Grzybowski, E. Gwóźdź, and A. Bródka. Ewald summation of electrostatic interactions in molecular dynamics of a three-dimensional system with periodicity in two directions. *Phys. Rev. B*, 61:6706–6712, 2000.

[GL04] L. Greengard and J.-Y. Lee. Accelerating the nonuniform fast Fourier transform. *SIAM Rev.*, 46:443–454, 2004.

[GR87] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comput. Phys.*, 73:325–348, 1987.

[Hac05] W. Hackbusch. Entwicklungen nach Exponentialsummen. Technical report, Max Planck Institute for Mathematics in the Sciences, 2005.

[Har08] F. E. Harris. Incomplete Bessel, generalized incomplete gamma, or leaky aquifer functions. *J. Comput. Appl. Math.*, 215:260–269, 2008.

[HE88] R. W. Hockney and J. W. Eastwood. *Computer simulation using particles*. Taylor & Francis, Inc., Bristol, PA, USA, 1988.

[HL06] F. Hedman and A. Laaksonen. Ewald summation based on nonuniform fast Fourier transform. *Chem. Phys. Lett.*, 425:142–147, 2006.

[HNP17] M. Hofmann, F. Nestler, and M. Pippig. NFFT based Ewald summation for electrostatic systems with charges and dipoles. *Appl. Numer. Math.*, 122:39–65, 2017.

[HR09] M. Hofmann and G. Rünger. Fine-grained Data Distribution Operations for Particle Codes. In M. Ropo, J. Westerholm, and J. Dongarra, editors, *Recent Advances in Parallel Virtual Machine and Message Passing Interface, 16th European PVM/MPI Users Group Meeting*, volume 5759 of *LNCS*, pages 54–63. Springer, 2009.

[HR13] M. Hofmann and G. Rünger. Efficient Data Redistribution Methods for Coupled Parallel Particle Codes. In *Proc. of the 42nd International Conference on Parallel Processing (ICPP-2013)*, pages 40–49. IEEE, 2013.

[Jac09] M. Jacob. Optimized least-square nonuniform Fast Fourier Transform. *IEEE Trans. Signal Process.*, 57:2165–2177, 2009.

[JCW] S. Johnson, A. Carvellino, and J. Wuttke. libcerf, numeric library for complex error functions. `http://apps.jcns.fz-juelich.de/libcerf`.

[JMNM91] J. I. Jackson, C. H. Meyer, D. G. Nishimura, and A. Macovski. Selection of a convolution function for Fourier inversion using gridding. *IEEE Trans. Med. Imag.*, 10:473–478, 1991.

[Kab12] I. Kabadshow. *Periodic Boundary Conditions and the Error-Controlled Fast Multipole Method*. PhD thesis, Bergische Universität Wuppertal, Jülich, 2012.

[Kai66] J. F. Kaiser. Digital filters. In F. F. Kuo and J. F. Kaiser, editors, *System analysis by digital computer*. Wiley, New York, 1966.

[KD11] I. Kabadshow and H. Dachsel. The error-controlled Fast Multipole Method for open and periodic boundary conditions. In G. Sutmann, P. Gibbon, and T. Lippert, editors, *Fast Methods for Long-Range Interactions in Complex Systems*, IAS-Series, pages 85–113, Jülich, 2011. Forschungszentrum Jülich.

[KKP09] J. Keiner, S. Kunis, and D. Potts. Using NFFT3 - a software library for various nonequispaced fast Fourier transforms. *ACM Trans. Math. Software*, 36:Article 19, 1–30, 2009.

[KP92] J. Kolafa and J. W. Perram. Cutoff errors in the Ewald summation formulae for point charge systems. *Mol. Simul.*, 9(5):351–368, 1992.

[KS04] K. N. Kudin and G. E. Scuseria. Revisiting infinite lattice sums with the periodic Fast Multipole Method. *J. Chem. Phys.*, 121(7):2886–2890, 2004.

[KS05] N. A. Kumar and C. Seidel. Polyelectrolyte brushes with added salt. *Macromolecules*, 38:9341 – 9350, 2005.

[KW86] M. H. Kalos and P. A. Whitlock. *Monte Carlo methods*, volume 1 of *A Wiley-Interscience Publication*. John Wiley & Sons, Inc., New York, 1986.

[LAM+06] H. J. Limbach, A. Arnold, B. A. Mann, C. Holm, and G. Berne. ESPResSo — an extensible simulation package for research on soft matter systems. *Comp. Phys. Comm.*, 174:704–727, 2006.

[LT11] D. Lindbo and A.-K. Tornberg. Spectral accuracy in fast Ewald-based methods for particle simulations. *J. Comput. Phys.*, 230:8744–8761, 2011.

[LT12] D. Lindbo and A.-K. Tornberg. Fast and spectrally accurate Ewald summation for 2-periodic electrostatic systems. *J. Chem. Phys.*, 136:164111, 2012.

[MMY+04] P. Minary, J. A. Morrone, D. A. Yarne, M. E. Tuckerman, and G. J. Martyna. Long range interactions on wires: a reciprocal space based formalism. *J. Chem. Phys.*, 121:11949–11956, 2004.

[Moo38] C. Moore. *Summable series and convergence factors*. Dover books on intermediate and advanced mathematics. Dover Publications, 1938.

[MT99] P. Minary and M. E. Tuckerman. A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *J. Chem. Phys.*, 110:2810–2821, 1999.

[MTPM02] P. Minary, M. E. Tuckerman, K. A. Pihakari, and G. J. Martyna. A new reciprocal space based treatment of long range interactions on surfaces. *J. Chem. Phys.*, 116:5351–5362, 2002.

[Nes12] F. Nestler. Approximationsverfahren zur schnellen Energieberechnung in Partikelsystemen. Diplomarbeit, Fakultät für Mathematik, Technische Universität Chemnitz, 2012.

[Nes16a] F. Nestler. Automated parameter tuning based on RMS errors for nonequispaced FFTs. *Adv. Comput. Math.*, 42(4):889–919, 2016.

[Nes16b] F. Nestler. An NFFT based approach to the efficient computation of dipole-dipole interactions under various periodic boundary conditions. *Appl. Numer. Math.*, 105:25 – 46, 2016.

[Nes16c] F. Nestler. Parameter tuning for the NFFT based fast Ewald summation. *Front. Phys.*, 4(28), 2016.

[Nes17] F. Nestler. Fast Ewald summation for electrostatic systems with charges and dipoles for various types of periodic boundary conditions. *Proceedings of the 12th International Conference on Sampling Theory and Applications*, pages 465–469, 2017.

[NH10] A. Neelov and C. Holm. Interlaced P3M algorithm with analytical and ik-differentiation. *J. Chem. Phys.*, 132(23):234103, 2010.

[NP13] F. Nestler and D. Potts. Fast Ewald summation under 2d- and 1d-periodic boundary conditions based on NFFTs. *Proceedings of the 10th International Conference on Sampling Theory and Applications*, pages 568–571, 2013.

[NPP15a] F. Nestler, M. Pippig, and D. Potts. Fast Ewald summation based on NFFT with mixed periodicity. *J. Comput. Phys.*, 285:280–315, 2015.

[NPP15b] F. Nestler, M. Pippig, and D. Potts. NFFT based fast Ewald summation for various types of periodic boundary conditions. In G. Sutmann, J. Grotendorst, G. Gompper, and D. Marx, editors, *Computational Trends in Solvation and Transport in Liquids*, volume 28 of *IAS-Series*, pages 575–598, Jülich, 2015. Forschungszentrum Jülich GmbH.

[Pipa] M. Pippig. PFFT - Parallel FFT software library. `https://github.com/mpip/pfft`.

[Pipb] M. Pippig. PNFFT - Parallel Nonequispaced FFT software library. `https://github.com/mpip/pnfft`.

[Pip13] M. Pippig. PFFT - An extension of FFTW to massively parallel architectures. *SIAM J. Sci. Comput.*, 35:C213–C236, 2013.

[Pip15] M. Pippig. *Massively Parallel, Fast Fourier Transforms and Particle-Mesh Methods*. Dissertation. Universitätsverlag Chemnitz, 2015.

[Por00] M. Porto. Ewald summation of electrostatic interactions of systems with finite extent in two of three dimensions. *J. Phys. A*, 33:6211–6218, 2000.

[PP11] M. Pippig and D. Potts. Particle simulation based on nonequispaced fast Fourier transforms. In G. Sutmann, P. Gibbon, and T. Lippert, editors, *Fast Methods for Long-Range Interactions in Complex Systems*, volume 6 of *IAS-Series*, pages 131–158, Jülich, 2011. Forschungszentrum Jülich GmbH.

[PP13] M. Pippig and D. Potts. Parallel three-dimensional nonequispaced fast Fourier transforms and their application to particle simulation. *SIAM J. Sci. Comput.*, 35:C411–C437, 2013.

[PPD88] J. W. Perram, H. G. Petersen, and S. W. De Leeuw. An algorithm for the simulation of condensed matter which grows as the 3/2 power of the number of particles. *Molecular Phys.*, 65:875–893, 1988.

[PPvR06] G. Pöplau, D. Potts, and U. van Rienen. Calculation of 3D space-charge fields of bunches of charged particles by fast summation. In A. Anile, G. Alì, and G. Mascaly, editors, *Scientific Computing in Electrical Engineering*, pages 241–246, Berlin Heidelberg, Germany, 2006. Springer.

[PS03] D. Potts and G. Steidl. Fast summation at nonequispaced knots by NFFTs. *SIAM J. Sci. Comput.*, 24:2013–2037, 2003.

[PSN04] D. Potts, G. Steidl, and A. Nieslony. Fast convolution with radial kernels at nonequispaced knots. *Numer. Math.*, 98:329–351, 2004.

[PST01] D. Potts, G. Steidl, and M. Tasche. Fast Fourier transforms for nonequispaced data: A tutorial. In J. J. Benedetto and P. J. S. G. Ferreira, editors, *Modern Sampling Theory: Mathematics and Applications*, pages 247–270, Boston, MA, USA, 2001. Birkhäuser.

[PT13] D. Potts and M. Tasche. Parameter estimation for nonincreasing exponential sums by Prony-like methods. *Linear Algebra Appl.*, 439:1024–1039, 2013.

[RK89] R. Roy and T. Kailath. ESPRIT—estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoustic speech and Signal Process.*, 37:984–994, 1989.

[SKE+05] Y. Shan, J. L. Klepeis, M. P. Eastwood, R. O. Dror, and D. E. Shaw. Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. *The Journal of chemical physics*, 122(5):54101, 2005.

[SS10] R. M. Slevinsky and H. Safouhi. A recursive algorithm for the G transformation and accurate computation of incomplete Bessel functions. *Appl. Numer. Math.*, 60:1411–1417, 2010.

[ST17] D. S. Shamshirgar and A.-K. Tornberg. The spectral Ewald method for singly periodic domains. *Journal of Computational Physics*, 347(Supplement C):341 – 366, 2017.

[Ste98] G. Steidl. A note on fast Fourier transforms for nonequispaced grids. *Adv. Comput. Math.*, 9:337–353, 1998.

[SW15] R. M. Slevinsky and E. J. Weniger. A fast algorithm for incomplete Bessel functions. Personal communication, 2015.

[TMPM02] M. E. Tuckerman, P. Minary, K. Pihakari, and G. J. Martyna. A new reciprocal space based method for treating long range interactions in ab initio and force-field based calculations for surfaces, wires, and clusters. In *Computational methods for macromolecules: challenges and applications (New York, 2000)*, volume 24 of *Lect. Notes Comput. Sci. Eng.*, pages 381–410. Springer, Berlin, 2002.

[TOS00] U. Trottenberg, C. W. Oosterlee, and A. Schuller. *Multigrid*. Academic Press, Inc., Orlando, FL, USA, 2000.

[WA97] A. H. Widmann and D. B. Adolf. A comparison of Ewald summation techniques for planar surfaces. *Comput. Phys. Commun.*, 107:167–186, 1997.

[War98] A. F. Ware. Fast approximate Fourier transforms for irregularly spaced data. *SIAM Rev.*, 40:838–856, 1998.

[WH01] Z. W. Wang and C. Holm. Estimate of the cutoff errors in the Ewald summation for dipolar systems. *J. Chem. Phys.*, 115:6277–6798, 2001.

[WHP+14] Y.-L. Wang, F. Hedman, M. Porcu, F. Mocci, and A. Laaksonen. Non-uniform FFT and its applications in particle simulations. *Appl. Math.*, 5:520–541, 2014.

[WNW+18] R. Weeber, F. Nestler, F. Weik, M. Pippig, D. Potts, and C. Holm. Accelerating the calculation of dipolar interactions in particle based simulations with open boundary conditions by means of the P2NFFT method. Preprint, 2018.

[YB99] I.-C. Yeh and M. L. Berkowitz. Ewald summation for systems with slab geometry. *J. Chem. Phys.*, 111(7):3155–3162, 1999.

[Zyg93] A. Zygmund. *Trigonometric Series*. Cambridge University press, 1993.

Theses

(1) The well known Ewald summation approach has become a helpful tool in order to efficiently compute electrostatic interactions in particle systems subject to periodic boundary conditions. Based on this technique, the interactions are split into short range and long range portions. The former converge rapidly in spatial domain, whereas the latter converge fast in Fourier space. For 3d-periodic boundary conditions the Fourier coefficients are known analytically and the resulting Fourier space representation allows the separation of the particle distances regarding the single dimensions, which serves as the basis for efficient algorithms.

In contrast, for mixed periodic or open constraints the computation of the analytic Fourier transform is only possible with respect to the periodic dimensions or not readily possible, respectively. Hence, the development of efficient algorithms is more challenging.

(2) Often one considers charged particle systems containing a set of charges $q_j \in \mathbb{R}$ at positions $\boldsymbol{x}_j \in \mathbb{R}^3$. In case of dipolar systems, containing a set of dipoles with dipole moments $\boldsymbol{\mu}_j \in \mathbb{R}^3$, one may formally replace the charges q_j by the operators $\boldsymbol{\mu}_j^\top \nabla_{\boldsymbol{x}_j}$, as is known. In order to study particle systems containing a mixture of charges and dipoles one may introduce the operators

$$\xi_j := \begin{cases} q_j & : \text{particle } j \text{ is a charge,} \\ \boldsymbol{\mu}_j^\top \nabla_{\boldsymbol{x}_j} & : \text{particle } j \text{ is a dipole,} \end{cases}$$

to end up with a convenient way to express the Ewald summation formulas.

(3) In the field of molecular dynamics simulations the root mean square error (rms) in the forces is commonly considered as an adequate measure of accuracy. Estimates concerning the truncation of the Ewald sums have already been derived in case of charge-charge and dipole-dipole systems. A handling of mixed charge-dipole systems is possible by considering the different types of interactions separately (charge-charge, charge-dipole, dipole-charge and dipole-dipole). The derivation of estimates concerning interactions between charges and dipoles can be realized based on similar techniques.

(4) The particle-particle NFFT (P^2NFFT) method for charged particle systems enables an efficient computation of the Fourier space parts of the electrostatic potentials by applying the NFFT and its adjoint. In order to compute the forces or rather the therefor required partial derivatives, the so called gradient NFFT may be applied. The resulting arithmetic complexity is $O(N \log N)$, where N denotes the number of present particles.

In order to treat mixed periodic and open boundary conditions one may combine the Ewald summation formulas and the NFFT based fast summation approach to end up with an efficient algorithm of the same structure as well as the same complexity. For this purpose, a precomputation step is required in order to compute the approximating Fourier coefficients numerically based on the FFT. However, in view of particle simulations involving several time steps, this precomputation step has to be conducted only once, since it does not depend on the positions of the particles. The precomputation step has to be repeated as soon as particles leave the primary simulation box.

(5) In order to extend the P^{2}NFFT method to mixed charge-dipole systems, two new NFFT variants are required, namely the adjoint gradient NFFT and the Hessian NFFT. The adjoint gradient NFFT enables an efficient computation of the so called structure factors of the dipole particles. The Hessian NFFT is needed in order to approximate the electric field gradients, which are required in order to compute the forces acting on the dipole particles. The gradient NFFT, the Hessian NFFT and the adjoint gradient NFFT may be applied with differentiation in Fourier space (i$\boldsymbol{k}$) or analytic differentiation.

The nature of the underlying Fourier coefficients is not relevant for the presented extensions for mixed charge-dipole systems. Thus, both features may be combined. Finally, the first $O(N \log N)$ algorithm for charge-dipole systems is presented, that in addition allows the application of various types of periodic boundary conditions.

(6) The current implementation of the P^{2}NFFT method allows the treatment of mixed charge-dipole systems under various types of periodic boundary conditions. It is part of the publicly available C software library ScaFaCoS. Concerning the shape of the primary simulation box the present implementation poses some restrictions in case of mixed periodic and open boundary conditions. However, the case of a general triclinic box can be traced back to the available special cases.

(7) The application of the different NFFT modules introduces further approximation errors, in addition to the Ewald truncation errors. Appropriate estimates have already been derived for the particle-particle particle-mesh (P^{3}M) method under 3d-periodic boundary conditions, where pure charge as well as pure dipole systems have been studied so far. These error estimates are valid for the P^{2}NFFT method in a very similar form, where additionally the application of an oversampling factor has to be taken into consideration. An extension to charge-dipole systems is realized based on similar techniques. The obtained estimates are valid for arbitrary compactly supported window functions in case of i$\boldsymbol{k}$-differentiation. For analytic differentiation we additionally require differentiability. In contrast to the i$\boldsymbol{k}$-differentiation approach, systematic errors are introduced in case of analytic differentiation, in addition.

Based on these estimates it is possible to automatically determine appropriate sets of parameters in order to achieve a certain rms force accuracy, provided that a near field cutoff radius is specified. For relatively small particle systems the optimal near field cutoff radius may be tuned based on measuring the resulting runtimes. The rather simple approach to apply the optimal parameters, determined for a small particle systems, to larger particle systems with the same particle density, already results in an arithmetic complexity of $O(N \log N)$.

(8) In case of mixed periodic and open boundary conditions, the analysis underlying the P^{2}NFFT method is much more complicated as in the 3d-periodic setting. Thus, reliable error estimates have not been derived yet. Nevertheless, a heuristic approach in order to predict the resulting approximation errors or rather to tune the involved parameters, that is based on the error analysis for 3d-periodic boundary conditions, may be applied.

(9) The NFFT allows the application of various types of window functions. Within the scope of P^{2}NFFT, the best results are obtained by the widely used B-spline window and the so called Bessel window. The optimal shape parameter for the Bessel function very much influences the resulting accuracy and depends on the decay of the underlying Fourier coefficients. If the shape parameter is tuned appropriately, the Bessel window may outperform the B-spline window in terms of accuracy. Relating to the application within the P^{2}NFFT method, the optimal window function will finally depend on the size of the considered particle system and the required accuracy, for instance.

(10) The computation of oversampled FFTs is possible within the P^2NFFT framework. Such an oversampling factor is not present in classical particle-mesh methods. In terms of computational efficiency it is oftentimes advantageous to spend some oversampling while using a window function with smaller support. The optimality of certain parameter constellations may strongly depend on the used software and hardware.