Fast computation of attenuated Radon transform

av
Sven F. Berglund

2004 - No 19
Fast computation of attenuated Radon transform

Sven F. Berglund

Examensarbete i matematik 20 poäng
Handledare: Jan-Olov Strömberg
2004
Abstract

The subject of this thesis is an algorithm for fast computation of a large set of line integrals, primarily in 2-dimensional space. The algorithm in question is previously known and is here documented, analyzed and extended.

The Radon transform of a function with domain \( \mathbb{R}^2 \) is given by the function integrated on straight one-dimensional lines of all locations and orientations. Given discrete data sampled from a function with compact support this algorithm can approximate the Radon transform with an, under reasonable conditions, arbitrary accuracy.

The algorithm has been extended to computation of attenuated line integrals, i.e. with the integrand weighted by a function belonging to a certain class. The attenuated Radon transform arises as a model in tomography where fast computation has applications in numerical inversion methods.

Furthermore, the algorithm has been generalized to arbitrary dimensions. In more than two dimensions it can be regarded as a discrete approximation of the X-ray transform (since the Radon transform in \( n \)-dimensional space integrates a function on \((n - 1)\)-dimensional hyperplanes whereas this algorithm, like the X-ray transform, sticks with integration along one-dimensional lines).

Coordinates and parameterizations are given so that the algorithm can be defined as a composition of discrete operators. In this context an alteration of the algorithm enables computation of the adjoint operator as well.
The thesis was supervised by Prof. Jan-Olov Strömberg at the Royal Institute of Technology (KTH) in Stockholm. The subject of the thesis was outlined by Ozan Öktem at Sidec Technologies, who also gave complementary supervision. Support was also given by Jan Boman at Stockholm University.
Appendices

**Appendix A - Sizes**

A.I Size of the set $\Sigma_k$ ........................................ i
A.II Size of the set $\hat{\Sigma}_k$ .................................. iii
A.III Size of the set $\mathcal{L}_k$ .................................. iii

**Appendix B - Accuracy analysis**

(2-dimensional algorithm)

B.I Error analysis at a single upcount of level ........................................ i
B.II Angular distribution of total error at level $k$ ............................... iv
B.III $h$-dependence of $k$ .................................................. viii
B.IV Shrinkage of error ..................................................... ix
B.V Angle of greatest error ............................................... ix
B.VI Correlation of interpolation errors ......................................... x
B.VII Conclusions ............................................................. xi

**Appendix C - Accuracy analysis**

($n$-dimensional algorithm with attenuation)

C.I Interpolation errors .................................................... i
C.II Propagation of the error of $\hat{E}^k$ to $\hat{E}^{k+1}$ ......................... v
C.III Propagation of the error of $\hat{I}_\mu^k$ to $\hat{I}_\mu^{k+1}$ .................. vi
C.IV Improved accuracy by the second approach in computation of $\mathcal{F}$ ... vii
C.V Estimating total error at level $k$ ........................................ ix
C.VI Conclusions ............................................................. xi
## Symbols and notations

Symbols listed with page of first appearance/definition.

### Coordinates, variables

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathcal{L} )</td>
<td>3</td>
</tr>
<tr>
<td>((x, \omega, L))</td>
<td>3</td>
</tr>
<tr>
<td>((x, \xi))</td>
<td>3</td>
</tr>
</tbody>
</table>

### Discretization of variables

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>([a, b])</td>
<td>5</td>
</tr>
<tr>
<td>(D)</td>
<td>10</td>
</tr>
<tr>
<td>(h_i)</td>
<td>5</td>
</tr>
<tr>
<td>(h_{\text{max}}, \text{mesh size})</td>
<td>5</td>
</tr>
<tr>
<td>(J_k,l, \Sigma_k, x_l)</td>
<td>5</td>
</tr>
<tr>
<td>(\hat{J}_k,j, \hat{\Sigma}_k, \xi_j)</td>
<td>5</td>
</tr>
<tr>
<td>(L_k, \ell)</td>
<td>7</td>
</tr>
</tbody>
</table>

### Related to level \(k\)

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{cells of level } k)</td>
<td>6</td>
</tr>
<tr>
<td>(2^k)-multiple</td>
<td>5</td>
</tr>
<tr>
<td>(2^k)-component</td>
<td>5</td>
</tr>
<tr>
<td>(k)</td>
<td>6</td>
</tr>
<tr>
<td>(k_L(h_{\text{min}}))</td>
<td>16</td>
</tr>
</tbody>
</table>

### Related to \(f\) and \(\mu\)

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f_\Omega, f, \hat{f})</td>
<td>10</td>
</tr>
<tr>
<td>(\mu_\Omega, \mu, \tilde{\mu})</td>
<td>26/App. C:iii</td>
</tr>
<tr>
<td>(D)</td>
<td>26/App. C:iii</td>
</tr>
</tbody>
</table>

### Parameterizations

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>({\alpha_i}_j^m)</td>
<td>21</td>
</tr>
<tr>
<td>(B(i), \hat{\iota}, \xi_j(s))</td>
<td>24</td>
</tr>
<tr>
<td>(\xi_j(s))</td>
<td>13/22</td>
</tr>
</tbody>
</table>

### Operators

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mathcal{P}(f))</td>
<td>1, 3</td>
</tr>
<tr>
<td>(\mathcal{P}_\mu(f))</td>
<td>1, 17</td>
</tr>
<tr>
<td>(\mathcal{I}(f))</td>
<td>4</td>
</tr>
<tr>
<td>(\mathcal{I}_\mu(f))</td>
<td>18</td>
</tr>
<tr>
<td>(\mathcal{E}(\mu))</td>
<td>18</td>
</tr>
</tbody>
</table>

### Discretization of operator

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mathcal{I}^\mathcal{L}<em>k(f</em>{\Omega}))</td>
<td>35</td>
</tr>
</tbody>
</table>

### Discrete operators and sets to compute

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{J}^k)</td>
<td>9/10, 12</td>
</tr>
<tr>
<td>(\hat{I}^k_L)</td>
<td>19/20, 21</td>
</tr>
<tr>
<td>(\hat{J}^k_{lj})</td>
<td>9, 19</td>
</tr>
<tr>
<td>(\hat{E}^k, {\hat{E}^k_{lj}})</td>
<td>19/20, 21</td>
</tr>
</tbody>
</table>

### In second approach:

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{M}^k, {\hat{M}^k_{lj}})</td>
<td>25</td>
</tr>
</tbody>
</table>

### Approximations

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mathcal{J}^k) approximates (\mathcal{I}(f))</td>
<td>12/13</td>
</tr>
<tr>
<td>(\mathcal{J}^k_\mu) approximates (\mathcal{I}_\mu(f))</td>
<td>21/22, 24</td>
</tr>
<tr>
<td>(\mathcal{F}^k) approximates (\mathcal{E}(\mu))</td>
<td>21/22, 24</td>
</tr>
</tbody>
</table>

### In second approach:

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mathcal{F}^k) approximates (\mathcal{E}(\mu))</td>
<td>/25</td>
</tr>
<tr>
<td>(\mathcal{N}^k) approximates (\mathcal{M}(\mu))</td>
<td>25</td>
</tr>
</tbody>
</table>

### Interpolation errors

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mathcal{R}^I)</td>
<td>14/App. B:iii</td>
</tr>
</tbody>
</table>

### Level-dependent:

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(k^R\mathcal{I}^I)</td>
<td>26/App. C:i</td>
</tr>
<tr>
<td>(k^R\mathcal{E}^I)</td>
<td>26/App. C:i</td>
</tr>
</tbody>
</table>

### Total errors

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mathcal{R}^k_\mathcal{L})</td>
<td>16, 28/App. B:viii</td>
</tr>
<tr>
<td>(\mathcal{R}^k_\mu)</td>
<td>29</td>
</tr>
</tbody>
</table>

### Error that is propagated to adjoint

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mathcal{R}^k_\mathcal{L}(f_{\Omega}))</td>
<td>36</td>
</tr>
</tbody>
</table>
1 Introduction

Attenuated X-ray transform and Radon transform

A line in $\mathbb{R}^n$ is the set of all linear combinations $\lambda_1 x + \lambda_2 y$ of two fixed points $x$ and $y$ in $\mathbb{R}^n$ such that $\lambda_1 + \lambda_2 = 1$. A line segment is given by the subset of a line where $\lambda_1$ and $\lambda_2$ are positive. When we distinguish between the line or line segment given by $(x, y)$ and the one given by $(y, x)$ we speak of oriented lines or oriented line segments.

For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ with compact support the X-ray transform $P(f)$ is given by the collection of all integrals on lines that intersect the support of $f$. For any line $\ell$ in $\mathbb{R}^n$, $P(f)(\ell)$ is defined as

$$P(f)(\ell) = \int_{\ell} f(x) \, dm(x)$$

with $m$ being the one dimensional Lebesgue-measure. When $f$ is a function on $\mathbb{R}^2$ the X-ray transform coincides with the Radon transform, which is the function integrated on $n - 1$ dimensional hyperplanes (whereas the X-ray transform sticks with "one dimensional hyperplanes" regardless of $n$).

When $\mu$ is another function on $\mathbb{R}^n$, that is positive, the attenuated X-ray transform is defined as

$$P_\mu(f)(\ell) = \int_{\ell} f(x) \rho_\mu(x, \ell) \, dm(x)$$

where $\ell$ is an oriented line and $\rho_\mu$ is a weight function. The attenuated X-ray transform arises as a model in single photon emission computed tomography (SPECT) in which case $\rho_\mu$ is the function

$$\rho_\mu(x, \ell) = e^{-\int_{\ell(x)} \mu(y) \, dm(y)}$$

where $\ell(x)$ is one of the components of $\ell \setminus \{x\}$. The attenuated X-ray transform reduces to the X-ray transform if $\mu$ is identically zero.

An algorithm

Inverse problems in these transforms arise in computed transmission tomography (CT) and SPECT. These are often handled with iterative numerical methods. In such case, when given a set of samples of $f$, being able to fast compute $P(f)$ or $P_\mu(f)$ numerically and with high accuracy is crucial. This thesis is dedicated to documentation, analysis and extension of a particular algorithm for computation of $P(f)$ or $P_\mu(f)$, so called forward projection. The algorithm, which is based on interpolation, has been described in [1] for the computation of line integrals in the plane. In this thesis it shall be extended to the computation of attenuated line integrals in $\mathbb{R}^n$ where $n \geq 2$. An interesting feature of the algorithm, that shall also be investigated, is that it can be altered to also compute the adjoint operator of $P_\mu$ so called back-projection. Fast computation of the
adjoint operator is useful e.g. in iterative solution methods to inverse problems. The algorithm is fast and the number of operations in two dimensions is $O(M^2 \log_2 M)$ where $M^2$ is the number of pixels. In the $n$-dimensional algorithm with $n > 2$ the operation count is $O(M^{2n-2})$ where $M^n$ is the number of discrete data points.

Emphasis has been put on mathematical aspects of the algorithm. From this point of view the algorithm can be defined as a composition of discrete operators. We shall not claim to give an exhaustive instruction for implementation. Several issues concerning implementation remain to be taken interest in. Rather, we shall concentrate in giving explicit coordinate settings and parameterizations that will in turn enable well defined extensions as well as a thorough accuracy analysis. We will look upon properties of the different operators that describes the forward projection algorithm and its adjoint version. Not least, shall the combinatorial properties of the different discrete sets involved be dealt with in the process of investigating their sizes.

**Structure of the text**

To enable a better overview of the text, some calculations that are quite extensive and not necessary for understanding the algorithm, have been assigned to appendices. In such cases the results of the calculations have been summarized in the main text. This concerns the sizes of the discrete sets involved (which are calculated in Appendix A and summarized in section 2.1.4) and the accuracy analysis (with calculations included in Appendix B and Appendix C and summaries in sections 3.3 and 5.4).

The material of the main text starts with setting up the coordinates and defining the sets needed in section 2. Since the coordinates and sets are best motivated by the algorithm itself, it may be most convenient to read this section just briefly and rather look back on the definitions when needed.

In section 3 the algorithm is described for computing the forward projection of functions on $\mathbb{R}^2$.

The concept of attenuation is discussed in section 4 and in 5 the algorithm is extended to computation of attenuated forward projection for functions on $\mathbb{R}^n$ with $n$ being arbitrary.

Section 6 concerns computation of the adjoint operator or back-projection and deals with issues such as propagation of accuracy-estimates from the forward projection to the adjoint.

Finally in section 7 we pay attention to some issues concerning implementation of the algorithm.

**Notations**

We shall in this thesis consider functions defined on the manifold $\mathbb{R}^n$. $\mathbb{R}^+$ shall denote the set of all positive reals and $S^{n-1}$ is the unit sphere in $\mathbb{R}^n$, i.e. $S^{n-1} = \{ x \in \mathbb{R}^n : |x| = 1 \}$. Lower-case $n$ shall consistently be used for the
number of dimensions of \( \mathbb{R}^n \).

In the task of enumerating discrete sets we shall utilize multi-indexes \( l \) and \( j \), that are \( n \)-tuples of integers, i.e. \( l = (l_1, l_2, \ldots, l_n) \) and \( j = (j_1, j_2, \ldots, j_n) \) where \( l_i \) and \( j_i \) are integers for all \( i \).

When performing component-wise operations on \( n \)-tuples of reals we shall use the circled operator \( \odot \) as

\[
\mathbf{x} \odot \mathbf{y} = (x_1 * y_1, x_2 * y_2, \ldots, x_n * y_n)
\]

where \( * \) is an operator for reals.

### 2 Coordinates on the set of lines in \( \mathbb{R}^n \)

We will be working with functions defined on the manifold \( \mathcal{L} \) of oriented line segments of finite length in \( \mathbb{R}^n \). In order to work with such functions we need to introduce coordinates on that manifold.

Let \( \ell \in \mathcal{L} \), i.e. \( \ell \) is an oriented line segment in \( \mathbb{R}^n \) of finite length. Any such \( \ell \) can be uniquely described with a triple \(( \mathbf{x}, \mathbf{\omega}, L) \in \mathbb{R}^n \times S^{n-1} \times \mathbb{R}^+\) where

\[
\ell = \{ \mathbf{x} + s \mathbf{\omega} \in \mathbb{R}^n : 0 \leq s \leq L \}
\]

i.e. \( \mathbf{x} \) is an endpoint of the line segment \( \ell \), \( \mathbf{\omega} \) is its direction, and \( L \) is its length.

In this way, the manifold \( \mathcal{L} \) is naturally identified with \( \mathbb{R}^n \times S^{n-1} \times \mathbb{R}^+ \). An alternative description is given by identifying \( \ell \) with a pair \(( \mathbf{x}, \mathbf{\xi}) \in \mathbb{R}^n \times \mathbb{R}^n \) where \( \mathbf{x} \) as before is an endpoint of the line segment \( \ell \) and \( \mathbf{\xi} \) contains both the length of the line segment (given by \( |\mathbf{\xi}| \)) and its direction (given by \( \mathbf{\xi}/|\mathbf{\xi}| \)). In these coordinates, \( \mathbf{x} \) is called the spatial variable and \( \mathbf{\xi} \) is called the angular variable. The relation between these two coordinates on \( \mathcal{L} \) is given as \( \ell = (\mathbf{x}, \mathbf{\xi}) = (\mathbf{y}, \mathbf{\omega}, L) \) whenever \( \mathbf{x} = \mathbf{y} \) and \( \mathbf{\xi} := L \mathbf{\omega} \).

**Definition 2.0.1.** Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a fixed integrable function. Then we define \( P(f) : \mathcal{L} \to \mathbb{R} \) as the line integral of \( f \) over an oriented line segment of finite length, i.e.

\[
P(f)(\ell) := \int_{\ell} f(\mathbf{x}) \, dm(\mathbf{x}) \quad \text{for } \ell \in \mathcal{L},
\]

where \( dm \) is the one-dimensional Lebesgue measure.

**Remark 2.0.1.** Assume that \( f \) has compact support and consider a fixed line \( \ell^* \in \mathbb{R}^n \) that intersects the support of \( f \). Note that \( \ell^* \) is a line, i.e. it is not a member of \( \mathcal{L} \) since it has infinite length. Also, let \( L_{\max} \in \mathbb{R}^+ \) denote the length of a line segment that contains the intersection of the line \( \ell^* \) with the support of \( f \). Now fix the variables \( \mathbf{x} \) and \( \mathbf{\omega} \), defining endpoint and direction of that line segment. Then, \( P(f)(\ell') = P(f)(\ell'') \) for all line segments \( \ell', \ell'' \in \mathcal{L} \) that are subsets of the fixed line \( \ell^* \) and which have length at least equal to \( L_{\max} \). This is simply because the parts of the line segments outside the support of \( f \) do not
Contribute to the line integral. In fact, for such line segments $P(f)$ equals the X-ray transform of $f$, i.e.

$$P(f)(t') = \int_{t'} f(y) \, dm(y)$$

holds for all line segments $t' \in L$ that has length at least $L_{\text{max}}$.

**Remark 2.0.2.** Expressed in the $\mathbb{R}^n \times S^{n-1} \times \mathbb{R}^+$-coordinates we get that

$$P(f)(x, \omega, L) = \int_0^L f(x + s\omega) \, ds \quad \text{for} \ (x, \omega, L) \in \mathbb{R}^n \times S^{n-1} \times \mathbb{R}^+,$$

and expressed in the $\mathbb{R}^n \times \mathbb{R}^n$-coordinates we get

$$P(f)(x, \xi) = \int_0^{|\xi|} f \left( x + s\frac{\xi}{|\xi|} \right) \, ds \quad \text{for} \ (x, \xi) \in \mathbb{R}^n \times \mathbb{R}^n.$$

We also introduce the weighted projection $I(f)$, which in the $\mathbb{R}^n \times S^{n-1} \times \mathbb{R}^+$-coordinates is given as

$$I(f)(x, \omega, L) := \frac{1}{L} P(f)(x, \omega, L)$$

and in the $\mathbb{R}^n \times \mathbb{R}^n$-coordinates, it is given as

$$I(f)(x, \xi) = \frac{1}{|\xi|} P(f)(x, \xi) = \int_0^1 f(x + s\xi) \, ds. \quad (1)$$

The task of the algorithm that we will describe is to efficiently compute the values of $P(f)$ on a specific finite subset of $L$. The algorithm will be described as consisting of levels, starting from zero, where in each level $P(f)$ is computed for line segments of certain lengths. We will however find it convenient to work with the weighted projection $I(f)$ in $\mathbb{R}^n \times \mathbb{R}^n$-coordinates given in (1). Computing the zeroth level of the algorithm can be done by any method where the error is under control and it is the only time when the sample values of $f$ needs to be accessed. Every following level of the algorithm then uses the integrals obtained in the previous level to calculate integrals of paths twice as long.

### 2.1 Discretization of $L$

The discretization of $L$ is based on a discretization of the domain of $f$. However, this discretization does not have to coincide with the set of points where $f$ is sampled, see e.g. the discussion in section 3.1.

As previously mentioned, the algorithm consists of levels $k$. In each level one computes the line integrals of $f$ over a specific finite set $L_k$ of line segments. The set $L_k$ is in turn defined by specific discretization $\Sigma_k$ and $\hat{\Sigma}_k$ of the spatial and angular variables, respectively.
2.1.1 Discretization of the spatial and angular variables

Assume that the support of $f$ is contained in an $n$-dimensional interval

$$[a, b] := [a_1, b_1] \times \ldots \times [a_n, b_n].$$

Now, let the vector $h \in \mathbb{R}^n$ be chosen such that for all $j = 1, \ldots, n$ there exists an integer $d_j$ where $(b_j - a_j)/h_j = 2^{d_j}$ holds. Given such $h$ we define the vectors $N, d$ of integers as

$$N := (b - a) \odot h = \left(\frac{b_1 - a_1}{h_1}, \ldots, \frac{b_n - a_n}{h_n}\right)$$

$$d := (d_1, \ldots, d_n) \quad \text{where} \quad 2^{d_j} = N_j.$$

Our next step is to introduce the following sets of multi-indexes that will be used in the discretization of the spatial and angular variables.

**Definition 2.1.1.** Assume that $N_i$ are defined as above given an $n$-dimensional interval $[a, b]$ and vector $h$. Then we define

$$J_0 := \left\{ l \in \mathbb{Z}^n : 0 \leq l_i \leq N_i \text{ for } i = 1, \ldots, n \right\}$$

and for $k \geq 0$ we define

$$J_k := \left\{ l \in J_0 : \text{There exists } i = 1, \ldots, n \text{ such that } 2^k | l_i \right\}$$

$$\hat{J}_k := \left\{ j \in \mathbb{Z}^n : |j_i| \leq 2^k \text{ for all } i = 1, \ldots, n \text{ and there exists } i \right. \left. \text{such that } |j_i| = 2^k \right\}.$$ 

We also define the finite set $\Sigma_k \subset [a, b]$ in the spatial variable as

$$\Sigma_k := \{ a + h \odot l \in \mathbb{R}^n : l \in J_k \}$$

and the finite set $\hat{\Sigma}_k \subset \mathbb{R}^n$ in the angular variable as

$$\hat{\Sigma}_k := \{ h \odot j \in \mathbb{R}^n : j \in \hat{J}_k \}$$

where $h \odot l = (h_1 l_1, h_2 l_2, \ldots, h_n l_n)$ and $h \odot j$ has an analogous meaning.

Note that from the definition we get that

$$\hat{J}_k = \left\{ j \in \mathbb{Z}^n : k = \max_{i=1, \ldots, n} \{ \log_2 |j_i| \} \right\}. \quad (2)$$

Moreover, $\Sigma_{k+1} \subset \Sigma_k$.

**Definition 2.1.2.** Any $x_l := a + h \odot l \in \Sigma_k$ has at least one component $x_i := a_i + l_i h_i$, which we refer to as a $2^k$-multiple of $x_l$, for which $l_i$ is a multiple of $2^k$. In a similar fashion, we see that $\xi_j := h \odot j \in \Sigma_k$ has at least one component of length $|\xi_i| = 2^k h_i$, which we refer to as a $2^k$-component of $\xi_j$.

In figure 1 the sets $\Sigma_k$ are depicted for some levels in a simple 2-dimensional case. From the figure it is also made clear why we shall refer to the maximal component $h_{\text{max}}$ of $h$ as the mesh size.
Cells

It shall be useful, particularly to motivate the coming discretization of $\mathcal{L}$, to think of the space $[a, b]$ as decomposed into cells by the set $\Sigma_k$.

**Definition 2.1.3.** The $k$’th discretization of the spatial variable divides the $n$-dimensional interval $[a, b]$ into a number of intervals with length $2^k h_i$ in the $i$’th dimension (figure 1). These $n$-dimensional intervals has endpoints with all components being $2^k$-multiples. For instance there is the first such interval $[a, a + 2^k h]$. These shall be referred to as the cells of level $k$. Note that $\Sigma_k \cap [a, a + 2^k h] = \emptyset$. In other words the points of $\Sigma_k$ are situated on the boundaries of the cells of level $k$ and not in their interiors.

### 2.1.2 Level index $k$

As already mentioned, the level index $k$ determines a certain finite subset $\mathcal{L}_k$ of $\mathcal{L}$. The values of $k$ are

$$k = 0, 1, \ldots, k_{\text{max}} \quad \text{where} \quad k_{\text{max}} = \log_2 \left\{ \max_{i=1, \ldots, n} \{ N_i \} \right\} \quad (3)$$

This choice of $k_{\text{max}}$ will guarantee that the line segments in $\mathcal{L}_{k_{\text{max}}}$ are of sufficient length for the X-ray transform of $f$ to be computed according to remark 2.0.1.
2.1.3 Defining the set $L_k$ of line segments

We now define $L_k$ which is the $k$'th set of line segments in $L$. An oriented line segment is determined by its spatial and angular variables $(x, \xi)$ through

$$\ell = \{x + s \xi \in \mathbb{R}^n : 0 \leq s \leq 1\}.$$

Given level $k$ we now consider spatial variables in $\Sigma_k$ and angular variables in $\hat{\Sigma}_k$. The set $L_k$ of line segments is now defined such that it only contains line segments with a spatial variable in $\Sigma_k$ and an angular variable in $\hat{\Sigma}_k$. Those line segments has at least one endpoint in $\Sigma_k$ whereas the other endpoint is allowed to be either in $\Sigma_k$ or outside $[a, b]$. Formally, this translates into requiring that a line segment with coordinates $(x_l, \xi_j) \in \Sigma_k \times \hat{\Sigma}_k$ is in $L_k$ whenever $\xi_j$ has at least one $2^k$-component on the same entry as a $2^k$-multiple of $x_l$. This will assure that the endpoint $x_l + \xi_j$ of the line segment $(x_l, \xi_j)$, if it remains inside $[a, b]$, also has at least one $2^k$-multiple and therefore it is also in $\Sigma_k$, see figure 1. Now, whenever $x_l := a + h \odot l$ and $\xi_j := h \odot j$, then

$$x_l + \xi_j = a + h \odot (l + j) =: x_{l+j}.$$

Hence, because $2^k \mid (l_i + j_i)$ for some $i$, we can characterize $L_k$ as follows:

$$2^k \mid l_i \text{ and } |j_i| = 2^k \text{ for some } i = 1, \ldots, n \implies (x_l, \xi_j) \in L_k,$$

so for $k > 0$, $L_k$ is a proper subset of $\Sigma_k \times \hat{\Sigma}_k$:

$$L_k = \left( \Sigma_k \times \hat{\Sigma}_k \right) \setminus \left\{ (x_l, \xi_j) \in \mathbb{R}^n \times \mathbb{R}^n : 2^k \mid l_i \implies |j_i| < 2^k \text{ for } i = 1, \ldots, n \right\}.$$

**Remark 2.1.1.** The set $L_k$ can be understood by looking at one of the $n$-dimensional intervals referred to as *cells* by definition 2.1.3, say the cell given by $[a, a + 2^k h]$. This particular cell, for the 3-dimensional case, is depicted in figure 2. The cell is bounded by hyperplanes given by setting one component of the spatial variable to a $2^k$-multiple (either $a_i$ or $a_i + 2^k h_i$ in this case). A point of $\Sigma_k$ can be a on more than one such hyperplane, since it can have more than one $2^k$-multiple. We shall use the parameter $\nu$ for the number of $2^k$-multiples (definition 2.1.2) in a particular point $x_l$. The line segments of $L_k$ that intersect any such cell can be regarded as parts of lines (infinite length). Then these lines strike such a point $x_l$ in an angle of incidence that is bounded by the relation $|\xi_i/\xi_j| \leq 2^k$ for $i = 1, \ldots, n$, where $x_j$ is a $2^k$-multiple.

2.1.4 The sizes of various sets

In Appendix A we express the number of elements in the various sets of section 2. Closed form expressions are derived for the case $N_{\text{max}} = N_{\text{min}} := N$ (see section 2.1.1 for the definition of $N$). We state here as a theorem some of the results.
Figure 2: The cell $[a, a + 2^k h]$. The angle of incidence $\theta$ is bounded by the relation $|\xi_i/\xi_j| \leq 1$ for $i = 1, \ldots, n$ where the $j$'th component of the spatial variable determines the plane in which the line intersect. A certain point of incidence may be on several ($\nu$) such planes and the limitation for angle of incidence is indicated in the figure for 3 different cases.
Theorem 2.1. In n dimensions the k’th discretization of the angular variable has the size 
\[ |\Sigma_k| = (2^{k+1} + 1)^n - (2^{k+1} - 1)^n \]
and in the special case \( N_{\text{max}} = N_{\text{min}} := N \) the k’th discretization \( \Sigma_k \) of the spatial variable and the set \( \mathcal{L}_k \) has the sizes:
\[ |\Sigma_k| = \left(1 + N\right)^n - \left(N - \frac{N}{2^k}\right)^n \]
\[ |\mathcal{L}_k| = \left[(2^{k+1} + 1)(N + 1)\right]^n - \left[(2^{k+1} + 1)(N + 1) - 2(2^{-k}N + 1)\right]^n \]

The proof of theorem 2.1 is supplied in Appendix A, where particularly the expression for \( |\mathcal{L}_k| \) is somewhat complicated to derive due to the fact that for \( k > 0 \), \( \mathcal{L}_k \) is a proper subset of \( \Sigma_k \times \hat\Sigma_k \). Being able to express \( |\mathcal{L}_k| \) may have some interest if one is to apply the results of section 6.2 where the sizes of the sets involved has importance for the result of a certain accuracy analysis.

3 The algorithm in two dimensions

An algorithm for computing approximations of \( I(f) \) (equation (1)) on the sets \( \mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_{k_{\text{max}}} \) (section 2.1) will be given. The algorithm needs the zeroth level (i.e. the approximations on the set \( \mathcal{L}_0 \)) as input while the levels for \( k > 0 \) are computed recursively.

In 3.1 a method of computing level zero is suggested. For \( k > 0 \) computation on the line segments from \( \mathcal{L}_k \) is described in 3.2. A thorough error analysis has been made and formulae bounding the errors are stated without proofs in section 3.3. References will be made into Appendix B where the results concerning the errors are proved. We begin with some basic definitions.

3.0.5 The set \( \{\hat{\mathcal{I}}_{ij}\}^k \)

The approximation of \( I(f)(x_l, \xi_j) \) on a line segment from the set \( \mathcal{L}_k \) will be denoted \( \hat{I}^k(x_l, \xi_j) \) or \( \hat{\mathcal{I}}_{ij}^k \). The entire set of such approximations will be denoted \( \{\hat{\mathcal{I}}_{ij}\}^k \) (and together with the set \( \mathcal{L}_k \) they will be referred to simply as level k). Note that the pair \( (l, j) \) of indexes determines a member of \( \{\hat{\mathcal{I}}_{ij}\}^k \) only when the line segment \( (x_l, \xi_j) \) is contained in \( \mathcal{L}_k \).

3.0.6 Near horizontal paths

As has been shown a line segment in \( \mathcal{L}_k \) is determined by a spatial variable \( x_l \) and an angular variable \( \xi_j \). In two dimensions the angular variable has the form
\[ \xi_j = (j_1h_1, j_2h_2) \]
where \( j_1, j_2 = 0, 1, \ldots, 2^k \) in level \( k \). We will distinguish between those members of \( \mathcal{L}_k \) for which \( |j_1/j_2| \leq 1 \) and \( |j_1/j_2| \geq 1 \). The former will be called
near-vertical and the latter near-horizontal. One important feature of the 2-dimensional algorithm is that integrals of near-vertical and near-horizontal paths are treated independently. We will therefore describe the algorithm for near-horizontal paths whilst the near-vertical ones can be treated analogously simply by swapping place of notations.

Remark 3.0.2. The 'diagonal' members of $\mathcal{L}_k$, i.e. the members for whom $|j_1| = |j_2|$, are both near-vertical and near-horizontal. These are exactly the members for whom $\xi_j$ has two $2^k$-components. The near-horizontal members of $\mathcal{L}_k$ are the members for which $|j_1| = 2^k$. In figure 1 near-horizontal and near-vertical paths of $\mathcal{L}_1$ are depicted separately.

3.0.7 Functions related to $f$

Some functions that are related to $f$ shall also be defined. $f_\mathcal{D}$ is a discretization of $f$ and $f_\mathcal{\downarrow}$ is a step-function based on $f_\mathcal{D}$. $\tilde{f}$ is a general expression for a function that approximates $f$ and that can be uniquely determined from a set of samples of $f$, more precisely from $f_\mathcal{D}$. These definitions shall be referred back to when needed.

Definition 3.0.4. Suppose that $\mathcal{D} \subset \mathbb{R}^n$ is a discrete set intersecting the support of $f$. The function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is then discretized to the function $f_\mathcal{D} : \mathcal{D} \mapsto \mathbb{R}$ such that

$$f_\mathcal{D}(x) := f(x) \quad x \in \mathcal{D}$$

Definition 3.0.5. In the special case that $\mathcal{D} = \Sigma_0$, we define the step-function $f_\downarrow$ that approximates $f$ as constant in semi-open $n$-dimensional intervals surrounding the points of $\mathcal{D}$:

$$f_\downarrow(y) := \begin{cases} f_\mathcal{D}(x) = f(x) & \text{if } y \in [x - \frac{h}{2}, x + \frac{h}{2}] \\ 0 & \text{otherwise.} \end{cases}$$

Definition 3.0.6. The notation $\tilde{f}$ shall sometimes be used for an approximation of $f$ defined on $\mathbb{R}^n$ and that can be uniquely determined from $f_\mathcal{D}$. For instance $\tilde{f}$ can be the step-function $f_\downarrow$. It can also be a smooth function that is adapted to $f_\mathcal{D}$.

3.1 Level zero

The algorithm takes the set $\{\tilde{I}_{lj}\}_0$ as input data. This set can be obtained however, although usually one assumes that it is computed from samples of $f$ on a discrete set $\mathcal{D} \subset \mathbb{R}^2$, which intersects the support of $f$. If $\tilde{f}$ is an approximation of $f$ that can be determined from samples of $f$ on the set $\mathcal{D}$ (definition 3.0.6), then one way of defining the level zero is by letting $\mathcal{D} := \Sigma_0$.

The set $\{\tilde{I}_{lj}\}_0$ can then be computed by letting $\tilde{f} = f_\downarrow$ (definition 3.0.5) and

$$\tilde{I}_{lj} := \frac{f(x_l) + f(x_l + \xi_j)}{2} = \frac{f(x_l) + f(x_{l+j})}{2} \quad \text{for } (l, j) \in J_0 \times \hat{J}_0.$$
There are of course other ways to obtain \( \{ \tilde{I}_{lj} \}^0 \). Some, such as computing \( \tilde{I}^0 \) by spline interpolation, are discussed in [1]. In such cases the sampling set \( \mathcal{D} \) could also be made denser than \( \Sigma_0 \).

Regardless of the computational method it shall be useful, particularly as a preparation for section 6, to define level zero as:

\[
\tilde{I}_{lj} := \mathcal{I} \left( \tilde{f} \right) (x_l, \xi_j) \quad \text{for} \ (l, j) \in J_0 \times \hat{J}_0.
\]

(Note that this definition implies the previous when \( \tilde{f} = f \).) This definition is useful because it enables a well defined discrete level-zero operator \( \tilde{\mathcal{I}}^0 \) mapping functions defined on \( \mathcal{D} \) on functions defined on \( \mathcal{L}_0 \). It can be expressed in terms of this operator:

\[
\tilde{\mathcal{I}}^0 (f_D) (x_l, \xi_j) := \mathcal{I} \left( \tilde{f} \right) (x_l, \xi_j) \quad \text{for} \ (l, j) \in J_0 \times \hat{J}_0.
\]

### 3.2 Recursive computation of level \( k > 0 \)

The following concerns the method of obtaining level \( k+1 \) from level \( k \). Figure 3 depicts a few paths of integrals computed in zeroth, first and second level. The integral paths of level 1 will be twice as long as those in level 0, and the angular resolution of the integral paths will be doubled. There are by necessity two ways of computing integrals in the \( k+1 \)th level of the algorithm. If there are two level \( k \) paths coinciding with a new level \( k+1 \) path the integral on the latter is simply computed by taking the average of the two level \( k \) integrals. Otherwise, there are four level \( k \) paths surrounding the new path, see magnification figure.
3 THE ALGORITHM IN TWO DIMENSIONS

3. Averaging these four integrals is the same as linearly interpolating the surrounding integrals two by two in order to achieve two integrals coinciding in path with the one that is to be computed. Higher order interpolation schemes could of course be used, utilizing more surrounding paths. However, we shall show that given fine enough initial resolution, linear interpolation will suffice to yield arbitrarily good accuracy. To examine, and later generalize, this algorithm we make the following observations. Computing members of level \( k + 1 \) is based on the simple relation

\[
\mathcal{I}(f)(x, \xi) = \frac{\mathcal{I}(f)(x, \frac{\xi}{2}) + \mathcal{I}(f)(x + \frac{\xi}{2}, \xi)}{2}
\]

(5)

that holds for the operator \( \mathcal{I} \) defined in (1). If there are two level \( k \) integrals coinciding in path with the new level \( k + 1 \) integral, then the latter can be computed by the formula

\[
\tilde{\mathcal{I}}^{k+1}(x_l, \xi_j) = \frac{\tilde{\mathcal{I}}^k(x_l, \frac{\xi_j}{2}) + \tilde{\mathcal{I}}^k(x_l + \frac{\xi_j}{2}, \frac{\xi_j}{2})}{2}
\]

where the right side must be well defined in terms of level \( k \) members. This equation can then be expressed by relation (4) in terms of indexes \( l, j \) as:

\[
\tilde{J}_{lj} := \frac{\tilde{\mathcal{I}}^k(x_l, \frac{\xi_j}{2}) + \tilde{\mathcal{I}}^k(x_l + \frac{\xi_j}{2}, \frac{\xi_j}{2})}{2}
\]

(6)

where:

\[
(l, j) \in L_{k+1} \text{ and } \left(l, \frac{j}{2}\right), \left(l + \frac{j}{2}, \frac{j}{2}\right) \in L_k
\]

The conditions for \( l \) and \( j \) of (6) is fulfilled if no component in \( j \) is an odd number (see section 2.1.1). By remark 3.0.2, component \( j_1 \) of \( (j_1, j_2) = j \) is always \( \pm 2^{k+1} \) thus the only concern is whether or not \( j_2 \) is odd. For the general case, with \( j_2 \) being even or odd, a few more definitions will be made in order to replace (6) with a general expression.

3.2.1 The approximation \( J^k \)

Since (6) is not well defined for all members \( (l, j) \) of \( L_{k+1} \), the following general definition will implement the relation (5) in the algorithm:

\[
\tilde{\mathcal{I}}^{k+1}(x_l, \xi_j) := \frac{J^k(x_l, \frac{\xi_j}{2}) + J^k(x_l + \frac{\xi_j}{2}, \frac{\xi_j}{2})}{2}
\]

(7)

Where \( J^k \) is some approximation, yet to be defined, calculated by using elements of \( \{\tilde{J}_{lj}\}^k \). The introduction of \( J^k \), as an intermediate step between \( \mathcal{I}^k \) and \( \mathcal{I}^{k+1} \), will be of particularly great benefit in section 5 where attenuation is accounted for and (7) shall be replaced with a more complicated relation. A path of \( J^k \) is illustrated in figure 4. This far the definitions given are applicable\(^1\) linear in the second component \( \xi_2 \) of \( \xi \)
when using any method of calculating \( J^k \) (e.g. higher order interpolations). Although we shall concentrate on linear interpolation. In order to define \( J \) by linear interpolation an explicit parameterization of \( \xi_j \) shall be given.

### 3.2.2 Parameterization of \( \xi_j \)

For any \( \xi_j = j \odot h \) index \( j \) has either zero or one component that is odd. If there is one such component \( j_2 \) then the even and odd parts \( j_e \) and \( j_o \) of \( j \) will be defined such that the odd part is also a function of one real variable \( s \in [0, 2] \) and that \( j = (j_1, j_2) = j_e + j_o(1) \) The even and odd parts are defined:

\[
\begin{align*}
  j_e &= (j_1, 0) \\
  j_o(s) &= \left(0, j_2 \left(1 + \frac{s-1}{12j_2} \right) \right)
\end{align*}
\]

Note that \( j_o(0) \) and \( j_o(2) \) are even in all components. If, on the other hand, no odd component in \( j \) is odd then we simply define the parts as:

\[
\begin{align*}
  j_e &= j \\
  j_o &= (0, 0)
\end{align*}
\]

Now also \( \xi_j \) can be replaced by a function of \( s \) such that:

\[
\xi_j(s) = (j_e + j_o(s)) \odot h.
\] (8)

Note that \( \xi_j \), written without argument, always refers to as defined in section 2.1.1 while this is equal to \( \xi_j(1) \) when considered as a function of \( s \). This shall cause no confusion since the function will always be written with argument.

**Definition 3.2.1.** We are now ready to give a precise definition of the approximation \( J^k \) of (7) based on linear interpolation:

\[
\begin{align*}
  J^k(x_l, \xi_j) &:= \frac{1}{2} \tilde{I}^k \left(x_l, \frac{\xi_j(0)}{2} \right) + \frac{1}{2} \tilde{I}^k \left(x_l, \frac{\xi_j(2)}{2} \right) \\
  J^k(x_l + \frac{\xi_1}{2}, \frac{\xi_2}{2}) &:= \frac{1}{2} \tilde{I}^k \left(x_l + \frac{\xi_1(0)}{2}, \frac{\xi_2(0)}{2} \right) + \frac{1}{2} \tilde{I}^k \left(x_l + \frac{\xi_1(2)}{2}, \frac{\xi_2(2)}{2} \right)
\end{align*}
\] (9)

Note that \( (l, j) \in \mathcal{L}_{k+1} \). By the following remark, 3.2.1, the right side terms of (9) are all either well defined members of \( \{\tilde{I}_{l,j}\}^k \) or zero.
Remark 3.2.1. This remark concerns the function \( \xi_j(s) \) when used in the definition of \( \mathcal{J}^k \). The following observations are to assure \( \mathcal{J}^k \) being well defined by equation (9): If \( (x_l, \xi_j) \in \mathcal{L}_{k+1} \) by \( l \) and \( j \) fulfilling the condition in section 2.1.3, then there will be \( 2^k \)-components (definition 2.1.2) in \( \frac{\xi_j(s)}{2} \) positioned as the \( 2^{k+1} \)-components of \( \xi_j \). Therefore
\[
\left( x, \frac{\xi_j(s_1)}{2} \right) \in \mathcal{L}_k
\]
for \( s_1 \) being 0 or 2. While
\[
\left( x + \frac{\xi_j(s_2)}{2}, \frac{\xi_j(s_3)}{2} \right),
\]
for \( s_2, s_3 \) taking any combination of the values 0 and 2, is either in \( \mathcal{L}^k \) or outside\(^2\) the support of \( f \) (since \( x + \frac{\xi_j(s_2)}{2} \) is either a point in \( \Sigma_k \) or outside \( [a, b] \)).

Any member of level \( k+1 \) can consequently be computed from level-\( k \) members by equations (7) and (9). The description that has now been given hence provides a method of how to in two dimensions approximate \( \mathcal{I}(f) \) on the set \( \mathcal{L}_k \) by means of the composition
\[
\tilde{I}(f_D) := \tilde{I}^k \circ \tilde{I}^{k-1} \circ \ldots \circ \tilde{I}^0(f_D)
\]
where \( \tilde{I} \) is an operator that maps a function defined on the discrete set \( D \) (section 3.1) on a function on the set \( \mathcal{L}_k \). In Appendix B the error \( |\mathcal{I}(f) - \tilde{I}(f_D)| \) is analyzed. The results are stated in the following section, with references to proofs in the appendix.

### 3.3 Errors

Analyzing the accuracy of the algorithm has been done in several steps and the calculations are quite lengthy. We shall therefore in this section give a summary of the results concerning the errors, that are all proved in Appendix B.

**Interpolation error**

In section B.I of the appendix, we show that the error caused by a single interpolation has a bound that is independent of level. The interpolation error is defined as a number \( R^I \) such that:
\[
\left| \mathcal{I}(f) \left( x_l, \frac{\xi_j}{2} \right) - \mathcal{I}(f) \left( x_l, \frac{\xi_j(0)}{2} \right) + \mathcal{I}(f) \left( x_l, \frac{\xi_j(2)}{2} \right) \right| \leq R^I
\]
\(^2\)The fact that level \( k+1 \) is defined partially with line-segments that are not in \( \mathcal{L}_k \) is commented further in section 7.2.
Given positive numbers \( |\frac{\partial^2 f}{\partial x^2}|_{\text{max}} \) and \( h_{\text{max}} \), bounding the second partial derivatives of \( f \) and the components of \( h \), then the following bound for the interpolation error can be used:

\[
R^I := \frac{1}{24} |\frac{\partial^2 f}{\partial x^2}|_{\text{max}} h_{\text{max}}^2
\]  

(10)

However, when estimating the error at an arbitrary level \( k \) all the errors from interpolations in previous levels has to be propagated, and this is the next task in the error analysis.

**Mean and upper bound of angular error-distribution**

We will assume that there is a global bound of the error at level zero as well as a global bound of the partial second derivatives of \( f \). Under these conditions a function bounding the error at any level \( k \) can be derived. This function will remain independent of the spatial variable but depend on the angular variable. More precisely, we define the error at level zero as a number \( R^0 \), such that

\[
|\tilde{I}(f)(x_l, \xi_j) - \mathcal{I}(f)(x_l, \xi_j)| \leq R^0
\]

for all \((x_l, \xi_j) \in \mathcal{L}_0\). The error at any level \( k > 0 \) on the other hand is an angle-dependent number \( R^k_I(\xi_j) \) such that

\[
|\tilde{I}(f)(x_l, \xi_j) - \mathcal{I}(f)(x_l, \xi_j)| \leq R^k_I(\xi_j)
\]

for all \((x_l, \xi_j) \in \mathcal{L}_k\). The buildup of interpolation errors \( R^I \) varies depending on angle and the error-distribution over an angular interval can be computed recursively by level. The angular error-distribution is computed by formulae (54) and (55) in section B.I, Appendix B, and shown in figure 5.

In appendix B.II, formula (58), we show that the mean \( \bar{R}^k_I \), taken over the angular variable, of the distribution on figure 5 is a linear function of \( k \):

\[
\bar{R}^k_I = \frac{k}{2} R^I + R^0
\]

For the maximum of the same distribution a recursive formula (60) is derived

\[
a_k = \frac{a_{k-1} + a_{k-2}}{2} + 1, \quad a_0 = 0, \quad a_1 = 1
\]

such that the maximal error in level \( k \) is given by

\[
\max_{j \in J_k} \{ R^k_I(\xi_j) \} = a_k R^I + R^0.
\]
This recursively defined maximum is shown to be tightly bounded by another function linear with \(k\), denoted plainly \(\mathcal{R}_k\):

\[
\mathcal{R}_k = \frac{2k+1}{\mathcal{I}} + \mathcal{R}^0
\]

where substituting the interpolation error with (10) yields:

\[
\mathcal{R}_k = \frac{2k+1}{\mathcal{I}} h_{\text{max}}^2 \left| \frac{\partial^2 \mathcal{I}}{\partial x^2} \right|_{\text{max}} + \mathcal{R}^0
\]

**h-dependence of \(k\)**

When considering the shrink-rate of the error with decreasing mesh size \(h_{\text{max}}\) one has to take into account that the mesh size will not be shrunken without affecting the length of the integrals computed. In appendix B.III a function \(k_L(h_{\text{min}})\) is defined (equation (62)) as the lowest level \(k\) where all integrals computed has pathlength equal to or exceeding a certain number \(L\). The function is bounded according to

\[
k_L(h_{\text{min}}) < \log_2 \left( \frac{L}{h_{\text{min}}} \right) + 1
\]

**A level-independent formula bounding the error**

If one is to shrink the error in the final level \(k_{\text{max}}\) by making the mesh-size smaller, perhaps rather than holding \(k_{\text{max}}\) fixed one wants to assure that integrals of paths exceeding a certain fixed length \(L\) is computed in the final level.
4 ATTENUATION

Obviously for sufficient shrinkage of \( h_{\text{max}} \) we have \( h_{\text{max}} = h_{\text{min}} := h \). Combining the interpolation error with equation (11) and the function \( k_L(h) \) of (12) enables a \( k \)-independent formula bounding the error when integrals are to be computed on paths exceeding length \( L \). This formula is:

\[
\mathcal{R}_{\frac{L}{l}}^{kL(h)} = \frac{(2 \log_2 L + 3 + 2 \log_2 \frac{1}{h})}{72} \left| \frac{\partial^2 f}{\partial x^2_{\text{max}}} \right| h^2 + \mathcal{R}^0
\]  

(13)

If \( \mathcal{R}^0 \) is neglected we see that the error \( \mathcal{R}_{\frac{L}{l}}^{kL(h)} \) decreases with \( h \) like \( \log \left( \frac{1}{h} \right) h^2 \).

Angles corresponding to maximal error

In appendix B.V it is shown that the values of the angular variable corresponding to the maximal error are in each level the ones closest approximating

\[
\frac{|j_2|}{j_1} = \frac{1}{3} \text{ or } \frac{2}{3}
\]

Positively correlated interpolation errors

In appendix B.VI the subject of correlation of interpolation errors is discussed. The question of interest is whether different errors affecting the computation tend to strike in the same direction and in the worst case cause a resulting error to grow close to an upper bound (in this case the angular error-distribution) and the aim is to derive what is the condition for positive correlation of the interpolation errors. It is concluded that interpolation errors, of integration paths that are sufficiently close to each other and with similar orientation, can be assumed to be positively correlated provided that

\[
f \in C^2
\]

particularly for sufficiently short paths.

4 Attenuation

In section 5 the algorithm of section 3 shall be generalized in two ways. Attenuated line integrals in \( n \) dimensions shall be computed where \( n \) is not necessarily 2. The coordinates and discretizations given in 2.1 are well prepared for generalization of the algorithm from \( \mathbb{R}^2 \) to \( \mathbb{R}^n \) but the generalization to attenuation requires a few preparations.

Definition 4.0.1. Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a fixed integrable function and \( \mu : \mathbb{R}^n \to \mathbb{R}^+ \) another fixed integrable function that is positive. We shall call \( \mu \) the attenuation and define \( \mathcal{P}_\mu(f) : \mathcal{L} \to \mathbb{R} \) as the attenuated line integral of \( f \) over an oriented line segment of finite length, i.e.

\[
\mathcal{P}_\mu(f)(\ell) := \int_{\ell} f(x) e^{-\int_{\ell(x)} \mu(y) dm(y)} dm(x)
\]
where $dm$ is the one-dimensional Lebesgue measure and $\ell'$ is a set that intersects $\ell$ depending on $x$. The relationship between $\ell$ and $\ell'$ is given by the following parameterization. A point $x = a + s\omega$ on $\ell$ determines $\ell'$ as:

$$
\ell = \{a + s\omega : s \in [0, L]\} \quad \ell' = \{a + (s + t)\omega : t \in [0, \infty]\}
$$

**Remark 4.0.1.** If $\mu$ has compact support we can integrate the attenuation over a line segment of finite length instead of over $\ell'$. Particularly for sufficiently large $L$ we can replace $\ell'$ with

$$
\ell'' = \{a + (s + t)\omega : t \in [0, L - s]\}
$$

in definition 4.0.1. This can be realized by expressing the attenuated projection in the $R^n \times S^{n-1} \times R^+$-coordinates of section 2

$$
P_\mu(f)(x, \omega, L) = \int_0^L f(x + \gamma \omega) e^{-\int_0^\gamma \mu(x + t\omega)dt} d\gamma
$$

$$
= \frac{e^{f_L^\infty \mu(x+t\omega)dt}}{L} \int_0^L f(x + \gamma \omega) e^{-\int_0^\gamma \mu(x + t\omega)dt} e^{-\int_0^\gamma \mu(x + t\omega)dt} d\gamma
$$

where $e^{-\int_0^\infty \mu(x+t\omega)dt} = 1$ for sufficiently large $L$ if $\mu$ has compact support.

**Attenuated weighted projection**

The weighted projection that was defined in section 2 shall now be generalized to the **attenuated weighted projection**:

$$
I_\mu(f)(x, \omega, L) = \frac{e^{f_L^\infty \mu(x+t\omega)dt}}{L} P_\mu(f)(x, \omega, L)
$$

$$
= \frac{1}{L} \int_0^L f(x + \gamma \omega) e^{-\int_0^\gamma \mu(x + t\omega)dt} d\gamma
$$

In the $R^n \times R^n$-coordinates, where $(x, \xi) = (x, \omega, L)$ if $|\xi| = L$ and $\omega = \frac{\xi}{|\xi|}$, the attenuated weighted projection is expressed:

$$
I_\mu(f)(x, \xi) = \int_0^1 f(x + \gamma \xi) e^{-|\xi| \int_0^\gamma \mu(x + t\xi)dt} d\gamma
$$

(14)

By remark 4.0.1, for sufficiently large $|\xi|$ and if $\mu$ has compact support: $I_\mu(f)(x, \xi) = \frac{1}{|\xi|} P_\mu(f)(x, \xi)$.

**Integrating the attenuation**

In the algorithm to be described in section 5 we shall compute approximations of $I_\mu(f)(x, \xi)$ over sets of line segments. In doing this we shall soon find it necessary to compute the **exponential function of the integral of the attenuation** over the same sets of line segments. Therefore the following operator is defined:

$$
E(\mu)(x, \xi) = e^{-|\xi| \int_0^1 \mu(x + t\xi)dt}
$$

(15)
5 The algorithm for $n \geq 2$ with attenuation

The algorithm of section 3 shall now be generalized in two ways. To include attenuation and to $n \geq 2$. In other words, an algorithm shall be given to compute approximations of the function $I_\mu(f)(x, \xi)$ defined by equation (14) over the oriented line segments of the sets $L_0, L_1, \ldots, L_{k_{max}}$ (defined in section 2.1). In the first approach to generalize the algorithm it will be necessary to approximate the operator $\mathcal{E}(\mu)(x, \xi)$ of equation (15) on the same sets. We shall thereafter present a second approach which. The second approach renders some more computational effort with some improvement in accuracy. Basically it replaces $\mathcal{E}(\mu)(x, \xi)$ with another operator $\mathcal{M}(\mu)(x, \xi)$ which calls for a slight change in the computational scheme. Frequent references shall be made back to sections 2, 3 and 4. The discretizations of variables made in section 2.1 allows the number of dimensions $n$ to be arbitrary.

In section 5.1 methods of obtaining level zero are discussed. In 5.2 the recursive computation of levels $k > 0$ is described. The second approach mentioned, with improved accuracy, is given in section 5.3. An error analysis has been conducted and the resulting formulae are stated without proofs in section 5.4. Proofs of all the results concerning the errors are contained in Appendix C. Finally, in section 5.5 the order of the computation-rate is investigated.

First a few additional definitions are needed.

5.0.1 The sets $\{\tilde{I}_{lj}\}^k$ and $\{\tilde{E}_{lj}\}^k$

Analogously to section 3.0.5, approximations of $I_\mu(f)$ and $\mathcal{E}(\mu)$ on the line segment $(x_l, \xi_j)$ from the set $L_k$ shall be denoted $\tilde{I}_\mu^k(x_l, \xi_j)$ and $\tilde{E}^k(x_l, \xi_j)$ or $\tilde{I}_{lj}^k$ and $\tilde{E}_{lj}^k$, with the level-index $k$ sometimes omitted. The entire sets of these approximations shall be denoted

$$\{\tilde{I}_{lj}\}^k$$

and together with $L_k$ they shall be referred to as level $k$.

Functions related to $\mu$

Analogously to the definitions 3.0.4, 3.0.5 and 3.0.6 some functions related to $\mu$ shall be needed:

**Definition 5.0.2.** Corresponding to the discrete set $D \subset \mathbb{R}^n$ of definition 3.0.4, the function $\mu : \mathbb{R}^n \rightarrow \mathbb{R}^+$ is discretized to the function $\mu_D : D \rightarrow \mathbb{R}^+$ such that

$$\mu_D(x) := \mu(x) \quad x \in D$$

**Definition 5.0.3.** In the special case that $D = \Sigma_0$ we define the step-function $\mu_\downarrow$ which approximates $\mu$ as constant in semi-open $n$-dimensional intervals sur-
ranging the points of $D$:
\[
\mu_\ast(y) := \begin{cases} 
\mu_\ast(x) = \mu(x) & \text{if } y \in [x - \frac{h}{2}, x + \frac{h}{2}] \text{ for some } x \in D, \\
0 & \text{otherwise.}
\end{cases}
\]

**Definition 5.0.4.** The notation $\tilde{\mu}$ shall be used for an approximation of $\mu$ defined on $\mathbb{R}^n$ and that can be uniquely determined from $\mu_\ast D$. $\tilde{\mu}$ can be the step-function $\mu_\ast$ or a smooth function that is adapted to $\mu_\ast D$.

### 5.1 Level zero

This algorithm takes the sets $\{\tilde{T}_{ij}\}_0^0$ and $\{\tilde{E}_{ij}\}_0^0$ as input data. Defining level zero serves an important purpose apart from suggesting computational method. As a preparation for section 6 we shall need to define discrete operators mapping functions defined on $D$ on functions defined on $L_0$. Therefore it shall be useful to define level zero, regardless of how it is to be computed, by the functions of definitions 3.0.6 and 5.0.4.

\[
\tilde{E}_{ij} := E(\tilde{\mu})(x_l, \xi_j) \quad \text{for all } l \in J_0, j \in \hat{J}_0
\]

\[
\tilde{T}_{ij} := T(\tilde{\mu})(x_l, \xi_j) \quad \text{for all } l \in J_0, j \in \hat{J}_0
\]

Expressed in the discrete level-zero operators the zeroth level is defined:

\[
\tilde{E}^0_\ast(x_l, \xi_j) := E(\tilde{\mu})(x_l, \xi_j) \quad \text{for all } x_l \in \Sigma_0, \xi_j \in \hat{\Sigma}_0
\]

\[
\tilde{T}^0_\ast(f)(x_l, \xi_j) := T(\tilde{\mu})(f)(x_l, \xi_j) \quad \text{for all } x_l \in \Sigma_0, \xi_j \in \hat{\Sigma}_0
\]

One particular way of doing the computations is to generalize the method of section 3.1, letting $D = \Sigma_0$, $\tilde{f} = f_\ast$, and $\tilde{\mu} = \mu_\ast$, and compute level zero with few operations straight from discrete data sampled on $D$.

### 5.2 Computation of level $k > 0$

The levels $k > 0$ are computed recursively by the same idea as described in section 3.2. Although the attenuation makes things a little more complicated. The simple relation of equation (5) is not valid for the attenuated weighted projection but can be replaced with the two relations

\[
\mathcal{E}(\mu)(x, \xi) = \mathcal{E}(\mu)(x, \xi \frac{h_2}{2}) \mathcal{E}(\mu)(x + \xi, \xi \frac{h_2}{2}) \quad (16)
\]

\[
\mathcal{I}_\mu(f)(x, \xi) = \frac{\mathcal{I}_\mu(f)(x, \xi \frac{h_2}{2}) \mathcal{E}(\mu)(x + \xi, \xi \frac{h_2}{2}) + \mathcal{I}_\mu(f)(x + \xi, \xi \frac{h_2}{2})}{2} \quad (17)
\]

that can be derived from equations (14) and (15). Equation (17) is the reason why we need $\mathcal{E}(\mu)$ approximated on the same set of lines as $\mathcal{I}_\mu(f)$ (with the exception of $L_{k_{\max}}$). Suppose now that the left sides of (16) and (17) shall
be approximated on some line segment \((x_l, \xi_j)\) in level \(k + 1\). Just like shown in the two-dimensional case in equation (6) the right sides of the equation will then be available as level \(k\) members if and only if all the components of \(j\) are even. Therefore (16) and (17) shall be approximated with relations that are well defined in the general case.

5.2.1 The approximations \(F^k\) and \(J^k_\mu\)

The definition made in section 3.2.1 shall be generalized. Members of the \(k + 1\)’th level are defined

\[
\tilde{E}^{k+1}(x, \xi) := F^k(x, \xi/2) F^k(x + \xi/2, \xi/2)
\]

\[
\tilde{I}^{k+1}(x, \xi) := J^k_\mu(x, \xi/2) F^k(x + \xi/2, \xi/2) + J^k_\mu(x + \xi/2, \xi/2)
\]

where \(F^k\) and \(J^k_\mu\) are approximations somehow calculated by using members of the sets \(\tilde{E}_{ij}\) and \(\tilde{I}_{ij}\). Two alternative ways of defining \(F^k\) and \(J^k_\mu\) by linear interpolation shall be given by two alternative parameterizations of \(\xi_j\). In section 5.3 we shall discuss yet another way of computing \(F^k\) without involving the set \(\tilde{E}_{ij}\).

Remark 5.2.1. From the relations (18) and (19) it follows that the sets \(\tilde{E}_{ij}\) (actually the set of approximations \(F^k\)) can be computed independently of the sets \(\tilde{I}_{ij}\). This means the sets \(\tilde{E}_{ij}\) can be computed once, for use with projections of several different functions \(f\).

Definition 5.2.1. For any multi index \(j\) that has \(m\) out of totally \(n\) components that are odd the set

\[\{\alpha_i\}_j\]

is the index set for those \(m\) components. In other words \(j_{\alpha_i}\) is odd for all \(i \in \{1, 2, \ldots, m\}\).

5.2.2 Univariate parameterization of \(\xi_j\)

For any \(\xi_j = j \odot h\) out of any set \(\tilde{\Sigma}_k\) the even and odd parts of \(j\) shall be defined

\[
\tilde{j}_e(s) = (j_{e1(s), j_{e2(s)}, \ldots, j_{en(s)})}
\]

\[
\tilde{j}_o(s) = (j_{o1(s), j_{o2(s)}, \ldots, j_{on(s)})}
\]

such that the odd part is also function of one real variable \(s \in [0, 2]\) and:

\[
j = \tilde{j}_e + \tilde{j}_o(1)
\]

The definition is made such that both \(\tilde{j}_o(0)\) and \(\tilde{j}_o(2)\) are even in all components:

- if \(i \in \{\alpha_i\}_j\) then: \(j_{o1}(s) = j_i \left(1 + \frac{s-1}{|j_i|}\right)\), \(j_{e1} = 0\)
- if \(i \notin \{\alpha_i\}_j\) then: \(j_{o1}(s) = 0\), \(j_{e1} = j_i\)
Now $\xi_j$ can also be expressed as a function of $s$ such that:

$$\xi_j(s) = (j_e + j_o(s)) \odot h$$

$\xi_j$ written without argument always refers to the original definition from section 2.1.1 which corresponds to $\xi_j(1)$ when regarded as a function of $s$.

**Definitions of $F^k$ and $J^k_{\mu}$**

We can now give precise definitions of the approximations $F^k$ and $J^k_{\mu}$ of (18) and (19):

$$F^k(x_l, \frac{\xi_j}{2}) = \frac{1}{2} \tilde{c}^k \left( x_l, \frac{\xi_j^{(0)}}{2} \right) + \frac{1}{2} \tilde{c}^k \left( x_l, \frac{\xi_j^{(2)}}{2} \right)$$

$$F^k(x_l + \xi_l, \frac{\xi_j}{2}) = \frac{1}{2} \tilde{c}^k \left( x_l + \frac{\xi_j^{(0)}}{2}, \frac{\xi_j^{(2)}}{2} \right) + \frac{1}{2} \tilde{c}^k \left( x_l + \frac{\xi_j^{(2)}}{2}, \frac{\xi_j^{(0)}}{2} \right)$$

$$J^k_{\mu}(x_l, \frac{\xi_j}{2}) = \frac{1}{2} \tilde{r}^k_{\mu} \left( x_l, \frac{\xi_j^{(0)}}{2} \right) + \frac{1}{2} \tilde{r}^k_{\mu} \left( x_l, \frac{\xi_j^{(2)}}{2} \right)$$

$$J^k_{\mu}(x_l + \xi_l, \frac{\xi_j}{2}) = \frac{1}{2} \tilde{r}^k_{\mu} \left( x_l + \frac{\xi_j^{(0)}}{2}, \frac{\xi_j^{(2)}}{2} \right) + \frac{1}{2} \tilde{r}^k_{\mu} \left( x_l + \frac{\xi_j^{(2)}}{2}, \frac{\xi_j^{(0)}}{2} \right)$$

Remark 3.2.1 of section 3.2 holds for the $n$-variable case as well which is why the right side terms of (20) and (21) are either well defined members of the sets $\{\tilde{c}^k\}$ and $\{\tilde{r}^k_{\mu}\}$ respectively or zero.

By the method of section 5.2.2 each approximation $F^k$ or $J^k_{\mu}$ is computed as the mean of not more than two members of level $k$, regardless of the number $m$ of components in the angular index $j$ that are odd. Alternatively they can be computed as a mean of $2^m$ level-$k$ members. This will be described next. The difference in these interpolation methods is illustrated in figure 6.

**5.2.3 Multivariate parameterization of $\xi_j$**

For any $\xi_j = j \odot h$ out of any set $\hat{\Sigma}_k$ the even and odd parts of $j$ shall now be defined

$$j_e = (j_{e1}, j_{e2}, \ldots, j_{en})$$

$$j_o(s) = (j_{o1}(s), j_{o2}(s), \ldots, j_{on}(s))$$

such that the odd part is also function of an $m$-dimensional variable $s = (s_1, s_2, \ldots, s_m)$ where $m$ is the number of odd components in $j$ (i.e. $m = |\{\alpha_{i,j}\}|$)
5 THE ALGORITHM FOR $N \geq 2$ WITH ATTENUATION

Figure 6: Illustration of the difference in interpolation methods in $\mathbb{R}^3$. Two paths of integrals are depicted, both with one $2^k$-component in $\xi_j$ ($\nu = 1$). The upper of the paths has two components in $\mathbf{j}$ that are odd ($m = 2$) and the lower has one ($m = 1$).
by definition 5.2.1). The components of \( s \) are real numbers all in the interval \([0, 2]\) and

\[
j = j_e + j_o(1)
\]

The definition of \( j_o(s) \) shall be made such that \( j_o(0) \), \( j_o(2) \) as well as \( j_o(s) \), when \( s \) is any \( m \)-tuple of zeroes and twos, are even in all components.

If \( i = \alpha_i \in \{\alpha_i\}_j \) then: \( j_o(s) = j_i \left(1 + \frac{s_i - 1}{\|\alpha_i\|}\right) j_{ei} = 0 \)

If \( i \notin \{\alpha_i\}_j \) then: \( j_o(s) = 0 j_{ei} = j_i \)

Now \( \xi_j \) can also be expressed as a function of \( s \) such that:

\[
\xi_j(s) = (j_e + j_o(s)) \circ h
\]

\( \xi_j \) written without argument always refers to the original definition from section 2.1.1 which corresponds to \( \xi_j(1) \) when regarded as a function of \( s \).

**Definitions of \( F^k \) and \( J^k_\mu \)**

**Definition 5.2.2.** Let \( B(i) \) be an bijective mapping of the integers \( i = 0, 1, \ldots, 2^m - 1 \) on all \( m \)-tuples of ones and zeros. This will be accomplished by representing \( i \) binary in \( m \) bits. If these \( m \)-tuples are regarded as representing the corners of an \( m \)-dimensional cube, then the opposite corner of \( B(i) \) is given by \( B(2^m - 1 - i) \), i.e. these two \( m \)-tuples has no common value in any component.

For the number \( 2^m - 1 - i \) we shall use the notation \( i \), so that:

\[
B(i) + B(i) = (1, 1, \ldots, 1_m)
\]

for \( 0 \leq i \leq 2^m - 1 \)

By the mapping of definition 5.2.2 explicit definitions of \( F^k \) and \( J^k_\mu \) by the multivariate parameterization \( \xi_j(s) \) can now be given. The right side terms of (18) and (19) can be defined:

\[
F^k(x_l, \frac{\xi_j}{2}) = \frac{1}{2^m} \sum_{i=0}^{2^m-1} \hat{\xi}^k \left( x_l + \frac{\xi_j(2B(i))}{2}, \frac{\xi_j(2B(i))}{2} \right)
\]

\[
F^k(x_l + \frac{\xi_j}{2}, \frac{\xi_j}{2}) = \frac{1}{2^m} \sum_{i=0}^{2^m-1} \hat{\xi}^k \left( x_l + \frac{\xi_j(2B(i))}{2}, \frac{\xi_j(2B(i))}{2} \right)
\]

\[
J^k_\mu(x_l, \frac{\xi_j}{2}) = \frac{1}{2^m} \sum_{i=0}^{2^m-1} \hat{J}^k_\mu \left( x_l + \frac{\xi_j(2B(i))}{2}, \frac{\xi_j(2B(i))}{2} \right)
\]

\[
J^k_\mu(x_l + \frac{\xi_j}{2}, \frac{\xi_j}{2}) = \frac{1}{2^m} \sum_{i=0}^{2^m-1} \hat{J}^k_\mu \left( x_l + \frac{\xi_j(2B(i))}{2}, \frac{\xi_j(2B(i))}{2} \right)
\]

By remark 3.2.1 of section 3.2 the right side terms of (22) and (23) are either well defined members of the sets \( \{\xi_j\}^k \) and \( \{\hat{J}_{lj}\}^k \) respectively or zero.

Any member of level \( k + 1 \) can consequently be computed from level-\( k \) members by equations (18) and (19). Either univariate parameterization by (20) and (21) or multivariate parameterization by (22) and (23) can then be used to define
and compute $\mathcal{F}$ and $\mathcal{J}_\mu$ (an alteration of the method will be outlined in section 5.3). The description that has now now been given hence provides a method of how to in any dimensions approximate $\mathcal{I}_\mu(f)$ on the set $\mathcal{L}_k$ by means of the compositions

$$\tilde{\mathcal{I}}_\mu(f_D) := \tilde{\mathcal{I}}_\mu^k \circ \tilde{\mathcal{I}}_\mu^{k-1} \circ \ldots \circ \tilde{\mathcal{I}}_\mu^0(f_D)$$

where $\tilde{\mathcal{I}}_\mu$ is an operator that maps a function defined on the discrete set $\mathcal{D}$ (section 3.1) on a function on the set $\mathcal{L}_k$ and computation of $\tilde{\mathcal{I}}_\mu^k$ calls for computation of $\tilde{\mathcal{I}}_\mu^{k-1}$.

5.3 improved accuracy in computation of $\mathcal{F}$

An alternative method for computing $\mathcal{F}^k$, to those of equations (20) and (22), shall be outlined. Let

$$\mathcal{M}(\mu)(x, \xi) := \int_0^1 \mu(x + t\xi)dt$$

and suppose that there is a set $\{\mathcal{M}_{lj}\}_k$ of approximations

$$\tilde{\mathcal{M}}(x_l, \xi_j) \approx \mathcal{M}(\mu)(x_l, \xi_j)$$

of $\mathcal{M}(\mu)$ on the set $\mathcal{L}_k$.

**Remark 5.3.1.** Note that we have the algorithm for computing $\tilde{\mathcal{M}}$ as the weighted projection in $\mathbb{R}^n$ with zero attenuation. This is done by extending the 2-dimensional algorithm to $\mathbb{R}^n$ as described in the previous section, but without bothering about the attenuation.

This means that instead of the relations (16) and (17) we use the following relations:

$$\mathcal{M}(\mu)(x, \xi) = \frac{\mathcal{M}(\mu)\left(x, \frac{\xi}{2}\right) + \mathcal{M}(\mu)\left(x + \frac{\xi}{2}, \frac{\xi}{2}\right)}{2}$$

$$\mathcal{I}_\mu(f)(x, \xi) = \frac{\mathcal{I}_\mu(f)\left(x, \frac{\xi}{2}\right) e^{-\frac{\xi^2}{2} \mathcal{M}(\mu)(x + \frac{\xi}{2})} + \mathcal{I}_\mu(f)\left(x + \frac{\xi}{2}, \frac{\xi}{2}\right)}{2}$$

The approximation $\mathcal{F}^k\left(x_l, \frac{\xi_j}{2}\right)$, equation (18), can then be computed as

$$\mathcal{F}\left(x_l, \frac{\xi_j}{2}\right) = \exp\left\{-\left|\frac{\xi_j}{2}\right| N\left(x_l, \frac{\xi_j}{2}\right)\right\}$$

where $N$ is an approximation computed by either univariate parameterization as

$$N\left(x_l, \frac{\xi_j}{2}\right) := \frac{\tilde{\mathcal{M}}\left(x_l, \frac{\xi_j(0)}{2}\right) + \tilde{\mathcal{M}}\left(x_l, \frac{\xi_j(2)}{2}\right)}{2}$$

or by the corresponding multivariate parameterization.
Remark 5.3.2. By using this method we do not compute the sets \(\{\tilde{E}_{ij}\}\) \(k\) \((\text{section } 5.0.1)\) but compute the approximation \(\mathcal{F}^k\), for use in relation (19), straight from the set \(\{\tilde{M}_{ij}\}\) \(k\).

The approximations \(J^k\) are computed as suggested in the previous method. In section C.IV of Appendix C a comparison between the error-terms of the two methods is conducted and it is established that this second approach is an improvement, see section 5.4. The improvement can be regarded as a consequence of the term \(\left|\frac{\xi_j^2}{2}\right|\) in equation (25) where the ability to supply an exact value for this term eliminates one source of error compared to the first method. This improvement is particular for short integral paths.

In the following section formulae for estimation of the error \(|I_\mu(f) - \tilde{I}_\mu(f_D)|\) shall be presented.

5.4 Errors

Appendix C consists of an error analysis with attempts to generalize the calculations of Appendix B to the biggest extent possible. The attenuation makes things a great deal more complicated and some results shall now be given as recursive formulae. However, the second approach, described in section 5.3, offers better conditions for accuracy analysis than the first approach which is why in this case more useful formulae can be presented. A difference compared to the calculations made in Appendix B is that some approximative error propagation has been made in Appendix C by a commonly used formula (73), which basically approximates a function with its tangent hyperplanes.

The important results of the error analysis shall now be summed-up with references to proofs in the appendix. Most important are formulae (26), (27), (28) and (29) below.

Interpolation Errors

In Appendix C.I the interpolation errors for \(\mathcal{E}(\mu)\), and \(I_\mu(f)\) are investigated. The definitions of these interpolation errors (definition C.I.1) are made analogously to the interpolation error defined in section 3.3 for the zero-attenuation case in \(\mathbb{R}^2\) (see also definition B.I.1). In other words, they are the errors caused by a single interpolation. In this case though the errors are a lot more complicated than in the zero-attenuation case. Particularly, they are level-dependent. Functions bounding these errors are denoted \(R^k_\mu\) and \(R^k_{\mu f}\) respectively (and derived by applying theorem B.1, Appendix B, on the univariate parameterizations given by equations (20) and (21)). There are several terms of these errors which are all given in formulae (65) and (68) in the appendix.

Simplified expressions for the errors can be made by defining a positive constant \(D\) as follows: In the case that the functions \(\mu\) and \(f\) are bounded, as well as their partial first and second derivatives, \(D\) (definition C.I.2, Appendix C) denotes an upper bound of the absolute values of all the zeroth, first and second
THE ALGORITHM FOR N ≥ 2 WITH ATTENUATION

5

derivatives of \( f \) and \( \mu \).

The interpolation errors expressed in terms of \( D \) are given in Appendix C (formulae (66), (67), (69) and (70)). Particularly the function \( k R_{I\mu}^L \), bounding the interpolation error of \( I_{\mu} \) at level \( k \) can be defined as:

\[
k R_{I\mu}^L := \frac{m^2 D^2}{2} \left[ \frac{n}{64} 2^k h_{\max}^4 + \frac{\sqrt{n}}{4} 2^k h_{\max}^3 + \frac{5}{6} h_{\max}^2 + \frac{h_{\max}}{2^k} \right]
\]

(26)

where \( m \) is the number of dimensions over which to interpolate i.e. \( 0 < m < n - 1 \), see definition 5.2.1. This formula can be regarded as the generalization of equation (10). For integrals of pathlength \( L \) (see remark C.I.1, Appendix C) and with \( h_{\max} = h_{\min} := h \), the \( k \)-dependence can be substituted with the upper bound for the function \( k_L(h) \) of equation (12). All the terms of the error are then proportional to \( h^2 \) and we get the \( k \)-independent form:

\[
k_{L}(h) R_{I\mu}^L := \frac{m^2 D^2 h^2}{4} \left[ \frac{n}{8} L^2 + \sqrt{n} L + \frac{5}{3} + \frac{1}{L} \right]
\]

Formula (26) has been derived for univariate parameterization of \( \xi_j \) (section 5.2.2). Multivariate parameterization (section 5.2.3) should yield same or better error-estimates.

Recursive formulae for angular error distribution

In appendices C.II and C.III it is investigated how error estimates can be propagated from level \( k \) to level \( k+1 \). Formulae (71), (72), (75) and (76) describes this process depending on the angular variable and recursively by level and hence they generalize the angular error distribution of appendix B.II.

Improvement provided by the second approach

In appendix C.IV error analysis is conducted for the alternative computation of \( F \) described as the second approach (by the operator \( M(\mu) \), section 5.3). It is noted that the second approach avoids one interpolation for the pathlength \( \xi_j \) that has to be made under the first approach (by the operator \( E(\mu) \)).

A comparison can be made by studying an important part of the error of \( F \), for both methods. Specifically, we shall compare the error of \( F \) generated by one single interpolation.

In both methods \( F \) is an approximation of \( E(\mu) \), equation (15), and before proceeding we write \( E(\mu) \) in the form:

\[
E(\mu) \left( x_t, \frac{\xi_j}{2} \right) = e^{-\mathcal{X}(1)M(1)}
\]

where the right side is parametrized with \( s \) as

\[
\mathcal{X}(s) := \left| \frac{\xi_j(s)}{2} \right| \quad \text{and} \quad M(s) := M(\mu) \left( x_t, \frac{\xi_j(s)}{2} \right).
\]
For the first approach all the terms of the interpolation error of $E^k$ is specified in equation (65) and the error propagates times one to $F$:

Interpolation error of $F^k = kR_M^l$

$$= \exp\left(-\frac{XM}{2}\right) \left| (\lambda'M + \lambda'\lambda')^2 - (\lambda''M + 2\lambda'M' + M''\lambda') \right|_{\max}$$

In the second approach, on the other hand, $F^k$ is computed from $N^k$, equation (25), which has the error $kR_M^l$ (interpolation error of $M^k$). This error can be calculated according to theorem B.1 and propagated first order to $F$ by the error propagation formula (73), which yields:

Interpolation error of $F^k \approx \frac{\partial F^k}{\partial N^k} \cdot kR_M^l$

$$= \exp\left(-\frac{XM}{2}\right) |X'M'_{\max}|$$

This means that, up to first order approximation, all the terms except $\exp\left(-\frac{XM}{2}\right) |X'M'_{\max}|$, in the interpolation error of $F$ will be eliminated using the second approach. In the appendix we also show that this remaining term (79) is proportional to pathlength $L$

(wheras the rest of the terms are proportional to either one of $L^2$, $L$, $L^0$ or $L^{-1}$) and hence the second approach is particularly an improvement for short paths. There is by remark C.IV.1 a predicament concerning how big an improvement of the algorithm that this method provides. In short, the problem is that all the terms that can be eliminated from the interpolation error of $E$ will still be present in the interpolation error of $I_{\mu}$.

**Error of $\hat{M}(\mu)$ at level $k$ and zero-attenuation case**

In appendix C.V the error at level $k$ is estimated. First we note that an upper bound $R_E^k$ for the error of $\hat{E}$ at level $k$ has to grow with a rate at least exponential by level.

Secondly, from now on we concentrate on the second approach, we have already remarked that approximating the operator $M(\mu)$ is just a case of generalizing the zero-attenuation algorithm of section 3. Formula (11) bounding the error of the weighted projection with no attenuation is generalized to $\mathbb{R}^n$ (80) by multiplication with $(n-1)^2$ which yields

$$R_E^k = (n-1)^2 \cdot \frac{2k+1}{h^2} \left| \frac{\partial^2 f}{\partial^2 x} \right|_{\max} + R^0$$

and in the case $\hat{I}(f) = \hat{M}(\mu)$ we have:

$$R_M^k = (n-1)^2 \cdot \frac{2k+1}{h^2} \left| \frac{\partial^2 \mu}{\partial^2 x} \right|_{\max} + R^0_M$$
where $R_{\tilde{M}}^k$ is a number such that
\[ |\bar{M}(\mu)(x_1, \xi_j) - M(\mu)(x_1, \xi_j)| \leq R_{\tilde{M}}^k \]
for all $(x_1, \xi_j) \in \mathcal{L}^k$
and $R_{\tilde{M}}^0$ is the error at level zero. The formula has by remark C.V.1 only been proved for $n = 2$. Formula (27) has two uses. First giving an error-estimate for the zero-attenuation case in $n$ dimensions. Second, propagating the error to $\mathcal{F}^k$ and finally to $\tilde{I}_{\mu}^k(f)$ we shall investigate the error in the ‘attenuated’ algorithm.

Error of $\mathcal{F}^k$

The error of $\tilde{M}$ given by (27), when propagated by first order approximation (formula (73)) gives the following error estimate for $\mathcal{F}^k$. If $R_{\mathcal{F}}^k$ is a number such that
\[ |\mathcal{F}(x_1, \xi_j) - \mathcal{E}(\mu)(x_1, \xi_j)| \leq R_{\mathcal{F}}^k \]
for all $(x_1, \xi_j) \in \mathcal{L}^{k+1}$
it can be estimated:
\[ R_{\mathcal{F}}^k \approx \sqrt{n} 2^k R_{\tilde{M}}^k \]

(28)

where $R_{\tilde{M}}^k$ is as given in (27). The growth-rate with $k$ is $k2^k$. Also a closed form $k$-independent formula can be given by substituting the expression (12) for the function $k_L(h)$ into (28). The error of $\mathcal{F}$ computed on paths equal to or exceeding a certain length $L$ can be estimated as bounded by:
\[ R_{\mathcal{F}}^{k_L(h)} \approx \sqrt{n} 2^k R_{\tilde{M}}^{k_L(h)} \]

where we can see that if the level-zero error is neglected, the remaining terms of the formula are
\[ O\left(h \log \frac{1}{h}\right) \]

Error of $\tilde{I}_{\mu}^k(f)$

The error $R_{\mathcal{F}}^k$ given by (28) shall finally be propagated to $\tilde{I}_{\mu}^k(f)$. We shall do so using a recursive formula (77) that is derived in appendix C.III. First, $R_{\tilde{I}_{\mu}}^k$ is defined as a number such that
\[ |\tilde{I}_{\mu}(f)(x_1, \xi_j) - I_{\mu}(f)(x_1, \xi_j)| \leq R_{\tilde{I}_{\mu}}^k \]
for all $(x_1, \xi_j) \in \mathcal{L}^k$.

The error $R_{\tilde{I}_{\mu}}^{k+1}$ has three terms of which the second is recursively defined:
\[ R_{\tilde{I}_{\mu}}^{k+1} = f_{\text{max}} R_{\mathcal{F}}^k + C_k R_{\tilde{I}_{\mu}}^k + k R_{\tilde{I}_{\mu}}^k \]

(29)

Where $C_k \in [\frac{1}{2}, 1]$ for all $k$.

$R_{\mathcal{F}}^k$ is given by (28).

\[ \text{The term } C_k \text{ is derived by substituting equation (75) into (77)} \]
5. THE ALGORITHM FOR $N \geq 2$ WITH ATTENUATION

$kR^I_{I\mu}$ is the interpolation error of $I_{\mu}$ given by (26) and it is also shown that the interpolation error is, for fixed pathlength, $O(h^2)$.

Consequently, in the zero attenuation case the formula (11) is generalized in a very simple way to (27), whereas for the 'attenuated' algorithm, we do not present the error at level $k$ in such a "ready-to-use" form, but as the partially recursive formula (29). A closed form expression can possibly be estimated from the latter, though this subject has not been treated.

5.5 Rate of computation

The number of operations is proportional to the number of line segments through the entire algorithm.

Let $M = N_{\text{max}} + 1$ i.e. $M^n \geq |\Sigma_0|$. In order to estimate the growth-rate with $M$ of the total number of operations we shall look at the cells of definition 2.1.3.

If $c_k$ is the number of cells in level $k$ then

$$c_k = \frac{c_{k-1}}{2^n}$$

due to the fact that the size of the cells is doubled in every dimension for every upcount of level (see figure 1).

Now, let $l_k$ be the number of line segments in level $k$ that intersects a certain cell and that has starting point on the boundary of that same cell (see figure 2). In every upcount of level the number of such possible starting points is increased with with factor $2^{n-1}$ (since the number of integers in the interval $[0, 2^k]$ is doubled). The number of line segments intersecting the cell from every such point will also increase with factor $2^{n-1}$ since the angular resolution is doubled. Therefore

$$l_k = l_{k-1} \cdot 2^{2n-2}$$

and

$$c_k l_k = c_0 l_0 2^{k(n-2)}$$

The total count of operations is limited proportional to:

$$\sum_{k=0}^{k_{\text{max}}} c_k l_k = c_0 l_0 \sum_{k=0}^{k_{\text{max}}} 2^{k(n-2)} \approx |\Sigma_0| |\hat{\Sigma}_0| \sum_{k=0}^{\log_2 M} 2^k(n-2)$$

$$\leq M^n |\Sigma_0| \sum_{k=0}^{\log_2 M} 2^k(n-2) = \begin{cases} M^n |\Sigma_0| \frac{M^{n-2} - 1}{2^{n-1} - 1} & \text{if } n > 2 \\ M^2 |\Sigma_0| \log_2 M + 1 & \text{if } n = 2 \end{cases}$$

By theorem 2.1 $|\hat{\Sigma}_0| = (3^n - 1)$ and we find that the total number of operations is $O(M^2 \log_2 M)$ if $n = 2$ and $O(M^{2n-2})$ if $n > 2$, where $M^n$ is the size of the initial discrete point set.

For $n > 2$ the number of operations is also proportional to $\frac{3^n - 1}{2^{n-1} - 1}$. 
6 The adjoint operator

In sections 3 and 5 a method has been given of how to approximate $I_\mu(f)$ on the set $L_k$ by means of the composition

$$\tilde{I}_\mu = \tilde{I}_k^h \circ \tilde{I}_k^{h-1} \circ \ldots \circ \tilde{I}_0^\mu$$

where $\tilde{I}_\mu$ is an operator that maps a function defined on the discrete set $D$ on a function on the set $L_k$. We will now provide a method of calculating the adjoint of $\tilde{I}_\mu$ w.r.t. a discrete inner product. Still we will work under the condition that the level-zero operator $\tilde{I}_0^\mu$ can be defined in several different ways (see remark 6.0.1 however concerning linearity of the level-zero operator). Below in 6.0.1 the necessary definitions are given. In section 6.1 the algorithm for computing the adjoint is presented. It is practically a reversal of the forward projection method developed in sections 3 and 5. That is, for certain values of parameters $h$, $k_{\text{max}}$, etc. (that are given under 2.1) we will be able to present not only an algorithm for computing $\tilde{I}_\mu$ but also a closely corresponding one for the adjoint $\tilde{I}_\mu^*$. In section 6.2 we check how the error estimates developed for $\tilde{I}_\mu$ can provide information on the accuracy of $\tilde{I}_\mu^*$.

6.0.1 Definitions

First we involve any way of obtaining level zero, see section 5.1, from any discretization of the function $f$. That is, let $D \subset \mathbb{R}^n$ be any discrete set intersecting the support of $f$ and upon which $f$ is discretized to $f_D$:

$$f_D : D \mapsto \mathbb{R}$$

$$f_D(x) = f(x)$$

The inner products for functions $f_1, f_2$ on $D$ and $g_1, g_2$ on $L_{k_{\text{max}}}$ is defined:

$$\langle f_1, f_2 \rangle = \sum_{x \in D} f_1(x)f_2(x)$$

$$\langle g_1, g_2 \rangle = \sum_{\ell \in L_{k_{\text{max}}}} g_1(\ell)g_2(\ell)$$

For any operator $F$ that maps a function $f : D \mapsto \mathbb{R}$ on a function $F(f) : L_{k_{\text{max}}} \mapsto \mathbb{R}$ the adjoint operator $F^*$ is the operator that fulfills

$$\langle F(f), g \rangle = \langle f, F^*(g) \rangle$$

for all functions $g : L_{k_{\text{max}}} \mapsto \mathbb{R}$. Now let $\tilde{I}_0^\mu$ be any linear level-zero operator (section 5.1 and remark 6.0.1), i.e. some approximation of $I_\mu$ on $L_0$ (section 2.1) obtained from samples on $D$.

$$\tilde{I}_0^\mu(f_D) : L_0 \mapsto \mathbb{R}$$

$$\tilde{I}_0^\mu(f_D)(x_l, \xi_j) \approx I_\mu(f)(x_l, \xi_j)$$
The operator $\hat{I}_\mu(f_D)$ can now be defined as the composition of the operators

$$\hat{I}_\mu = \hat{I}_{\mu,\text{max}}^{k} \circ \hat{I}_{\mu,\text{max}}^{k-1} \circ \ldots \circ \hat{I}_{\mu,\text{max}}^{0}$$

and has the adjoint (w.r.t the inner product defined):

$$\hat{I}_\mu^* = \hat{I}_{\mu,\text{max}}^{0\ast} \circ \hat{I}_{\mu,\text{max}}^{1\ast} \circ \ldots \circ \hat{I}_{\mu,\text{max}}^{k\ast}$$  \hspace{1cm} (30)

This is realized by:

$$\langle \hat{I}_\mu(f), g \rangle = \langle \hat{I}_{\mu,\text{max}}^{k\ast} \circ \hat{I}_{\mu,\text{max}}^{k-1} \circ \ldots \circ \hat{I}_{\mu,\text{max}}^{0\ast}(f), g \rangle$$

Remark 6.0.1. The adjoint operator is well defined for linear operators only. We must therefore assume that the level-zero operator $\hat{I}_\mu^{0\ast}$ is linear, e.g. by using the stepfunctions of definitions 3.0.5 and 5.0.3 in the definition of level zero (section 5.1). See also remark 6.2.2.

6.1 An algorithm for the adjoint operator

From (30) it follows that given the adjoint operator of $\hat{I}_\mu^{k}$, the composition for all $k \in \{0, 1, \ldots, k_{\text{max}}\}$ gives the adjoint of $\hat{I}_\mu$. We shall give the adjoint $\hat{I}_\mu^{k\ast}$ of $\hat{I}_\mu^{k}$ for any $k$ except the level zero operator.

$\hat{I}_\mu^{k\ast}$ is an operator that maps a function $G : \mathcal{L}_k \mapsto \mathbb{R}$ on a function $\hat{I}_\mu^{k\ast}(G) : \mathcal{L}_{k-1} \mapsto \mathbb{R}$.

We shall assume that $\hat{I}_\mu^{k}$ is computed by the multivariate parameterization defined by equations (19) and (23). The description can easily be changed to the univariate parameterization simply by replacing $s$ with $s$ and $2^m$ with $2$ in a few formulae. The function $B(i)$ of definition 5.2.2, that maps all integers $i$ from $\{0, \ldots, 2^m - 1\}$ on all $m$-tuples of ones and zeros, shall be used. Since we are computing the attenuated weighted projection and its adjoint we shall also assume that we have access to all the approximations $F^{k-1}$ (equation (18)) that are obtained from either the set $\{\tilde{E}_{ij}\}^{k-1}$ or the sets $\{\tilde{M}_{ij}\}^{k-1}$.

The description that follows will be somewhat implicit in the sense that we do not give a closed form expression for $\hat{I}_\mu^{k\ast}(G)(x_1, \xi_j)$. When computing the forward projection $\hat{I}_\mu^{k}(g)$ for a function $g : \mathcal{L}_{k-1} \mapsto \mathcal{L}_k$ we walk through all the members of $\mathcal{L}_k$. For every such line segment we collect data from the projections related to surrounding members of $\mathcal{L}_{k-1}$. When computing the adjoint we shall also walk through the elements of $\mathcal{L}_k$ but instead distribute data to the projections on surrounding paths of $\mathcal{L}_{k-1}$. Initially we set $\hat{I}_\mu^{k\ast}(G)$ to zero for all $(x_1, \xi_j) \in \mathcal{L}_{k-1}$.
Definition 6.1.1. When increasing a number $b$ by adding a number $a$ the notation

$$b + = a$$

shall be used.

Definition 6.1.2. Surrounding line segments.

For any member $(x_l, \xi_j)$ of $\mathcal{L}_k$, suppose that $j$ has $m$ components that are odd numbers. Then there are $2^{m+1}$ members of $\mathcal{L}_{k-1}$ related to $(x_l, \xi_j)$ by the parameterization of $\xi_j$ as

$$\left(x_l + \frac{\xi_j(2B(i))}{2}, \frac{\xi_j(2B(i))}{2}\right) \in \mathcal{L}_{k-1}$$

for $i = 0, 1, \ldots, 2^m - 1$ by (definition 5.2.2). For a specific line segment of level $k$, this set will be referred to as the surrounding level $k-1$ paths.

Note now that $\mathcal{F}^{k-1}$ is defined on the line segment $\left(x_l, \frac{\xi_j}{2}\right)$ (equation (18) and figure 4). Data is distributed from $G(x_l, \xi_j)$ to the projections on surrounding line segments according to:

$$\tilde{T}_{\mu}^{k-1}(G) \left(x_l, \frac{\xi_j(2B(i))}{2}\right) = \frac{\mathcal{F}(\mu)(x_l, \frac{\xi_j}{2})}{2^{m+1}} G(x_l, \xi_j)$$

$$\tilde{T}_{\mu}^{k-1}(G) \left(x_l + \frac{\xi_j(2B(i))}{2}, \frac{\xi_j(2B(i))}{2}\right) = \frac{1}{2^{m+1}} G(x_l, \xi_j)$$

for $i = 0, 1, \ldots, 2^{m-1}$.

$\tilde{T}_{\mu}^{k-1}(G)$ on every member of $\mathcal{L}_{k-1}$ this way gets contributions from $G$ on several members of $\mathcal{L}_k$.

Proof of the adjoint operator

It shall now be proved that $\tilde{T}_{\mu}^{k-1}$ as described above truly is the adjoint of $\tilde{T}_{\mu}^k$ by considering the product:

$$\langle \tilde{T}_{\mu}^k(g), G \rangle$$

of functions defined on $\mathcal{L}_k$. The product can be expressed

$$\sum_{(l,j) \in \mathcal{L}_k} G(x_l, \xi_j) \left[ \frac{\mathcal{J}_\mu(g) \left(x_l, \frac{\xi_j}{2}\right) \mathcal{F} \left(x_l + \frac{\xi_j}{2}, \frac{\xi_j}{2}\right) + \mathcal{J}_\mu(g) \left(x_l + \frac{\xi_j}{2}, \frac{\xi_j}{2}\right)}{2} \right]$$

(by equation (19)). Note that $\mathcal{J}_\mu(g) \left(x_l^{k-1}, \frac{\xi_j}{2}\right)$ is computed as a mean of $2^{k}$ elements $g(x_l^{k-1}, \xi_j^{k-1})$ of level $k-1$ that are situated on surrounding paths.
The same holds for $\mathcal{J}_\mu(g) \left( x^k_l + \xi^k_{l,j} \right)$ thus the product can be expressed:

$$\langle \tilde{I}_{\mu}^k(g), G \rangle = \sum_{(l,j) \in \mathcal{L}_k} \left[ G(x^k_l, \xi^k_j) \cdot \sum_{(l,j) \in \mathcal{L}_{k-1}} w[(l,j) \cdot (l,j)^{-1}] g(x^{k-1}_l, \xi^{k-1}_j) \right]$$

where $w[(l,j) \cdot (l,j)^{-1}]$ is a weight that is determined by the relationship between two line segments of level $k$ and level $k - 1$ respectively. The weight is only non-zero if the shorter line segment belongs to the set surrounding the long one. Specifically:

$$w[(l,j) \cdot (l,j)^{-1}] = \begin{cases} \frac{1}{2^{m+1}} & \text{if } (x^{k-1}_l, \xi^{k-1}_j) = \left( x^k_l, \frac{\xi^k_{l,j}(2B(i))}{2} \right) \\
\text{for some } i \in \{0, \ldots, 2^m - 1\} \\
\frac{1}{2^{m+1}} & \text{if } (x^{k-1}_l, \xi^{k-1}_j) = \left( x^k_l + \frac{\xi^k_{l,j}(2B(i))}{2}, \frac{\xi^k_{l,j}(2B(i))}{2} \right) \\
\text{for some } i \in \{0, \ldots, 2^m - 1\} \\
0 & \text{otherwise} \end{cases}$$

Therefore, by change in order of summation:

$$\langle \tilde{I}_{\mu}^k(g), G \rangle = \sum_{(l,j) \in \mathcal{L}_{k-1}} \left[ g(x^{k-1}_l, \xi^{k-1}_j) \cdot \sum_{(l,j) \in \mathcal{L}_k} w[(l,j) \cdot (l,j)^{-1}] G(x^k_l, \xi^k_j) \right]$$

$$= \langle g, \tilde{I}_{\mu}^k(G) \rangle$$

6.2 Accuracy in computation of the adjoint operator

We will in this section investigate how the accuracy propagates from $\tilde{I}_\mu$ to $\tilde{I}_{\mu}^*$ (see section 6.0.1). In other words the following question shall be dealt with: The operator $\tilde{I}_\mu$ approximates $I_\mu$ with an accuracy that can be estimated. Now, how well does $\tilde{I}_{\mu}^*$ approximate $I_{\mu}^*$? Note that the operators are defined on different function spaces ($\tilde{I}_\mu$ is a discrete operator while $I_\mu$ is the continuous operator defined by equation (14)). We shall therefore in section 6.2.1 define a discretization of the operator $I_\mu(f)$ which will be denoted $I_{\mu}^*(f_D)$. 
In section 6.2.2 we show how an error estimate depending on the underlying discrete function $f_D$, in the forward algorithm, can provide information of the accuracy in computation of the adjoint

In section 6.2.3 we show that when the error is defined relative to the discrete operator (of section 6.2.1) the error in terms of operator norm is actually bounded independently of $f_D$ and propagates straight to the adjoint in a simple way.

First a few definitions shall be given.

In the previous sections we have defined the discrete operator

$$\tilde{I}_\mu(f_D) = \tilde{I}_\mu^k \circ \tilde{I}_\mu^{k-1} \circ \ldots \circ \tilde{I}_\mu^0(f_D)$$

that has the adjoint

$$\tilde{I}_\mu^*(g) = \tilde{I}_\mu^0 \circ \tilde{I}_\mu^1 \circ \ldots \circ \tilde{I}_\mu^k(g)$$

where $g$ is a real function defined on the set $L_k$ and the inner products are defined

$$\langle f_1, f_2 \rangle = \sum_{x \in D} f_1(x)f_2(x) \quad (g_1, g_2) = \sum_{\ell \in L_k} g_1(\ell)g_2(\ell). \quad (31)$$

By the inner products of formula (31) the norm of a function $f$, with domain any of the spaces $D$ or $L_k$, is defined as a number

$$\|f\| = \sqrt{\langle f, f \rangle} \quad (32)$$

and for a discrete operator $F$ the operator norm is defined as the number

$$\|F\| = \sup_{\|f\|=1} \{\|F(f)\|\} \quad (33)$$

if such a number exists. $F$ is then called a bounded operator.

**6.2.1 Discretization of $I_\mu(f)$ to $I_{\mu}^{L_k}(f_D)$**

For a function $g : L_k \to \mathbb{R}$, in order to define the error of $\tilde{I}_\mu^*(g)$ we shall define a discrete operator $I_{\mu}^{L_k}(f_D)$ such that the adjoint $I_{\mu}^{L_k,*}(g)$ represents the ‘true’ function. This operator can be regarded as the $k$'th discretization of $I_\mu(f)$. It takes a function defined on $D$ and maps it on a function defined on $L_k$.

**Definition 6.2.1.** Let $\hat{f}$ be an integrable approximation of the function $f$ that can be uniquely reconstructed from $f_D$. For instance $\hat{f}$ can be the step function $f_\mu$ (definitions 3.0.5 and 3.0.6). It can also be a smooth function that is a higher order adaption to $f_D$. The discrete operator $I_{\mu}^{L_k}(f_D)$ is defined (analogously to the level-zero operator) such that:

$$I_{\mu}^{L_k}(f_D)(x_1, \xi_j) = I_\mu\left(\hat{f}\right)(x_1, \xi_j) \quad \text{for all } (x_1, \xi_j) \in L_k$$
Remark 6.2.1. The function \( \tilde{f} \) is involved in definition 6.2.1 in order for \( I_{\mu}^{Dk} \) to be well defined. The reason is that \( f \) itself can not be uniquely determined by \( f_D \) unless \( f \) is restricted to some special class of step functions, splines, et c. The use of \( \tilde{f} \) can therefore be regarded as the restriction of \( f \) to such a class of functions that can be uniquely determined from its discretization \( f_D \). Note that the set \( D \) can be an arbitrarily large discrete set (section 6.0.1).

Remark 6.2.2. In order for the adjoint of \( I_{\mu}^{Dk} \) to be well defined we shall however assume that \( I_{\mu}^{Dk} \) is a linear operator which leads to the following requirement on the adapted function \( \tilde{f} \). For any two sampled functions \( f_1 \) and \( f_2 \) and any two reals \( \lambda_1 \) and \( \lambda_2 \) the relation

\[
(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 \tilde{f}_1 + \lambda_2 \tilde{f}_2
\]

must hold in definition 6.2.1 (which is the case e.g. if \( \tilde{f} = f_D \)).

6.2.2 An estimate in terms of function norm

The error of \( \tilde{I}_\mu \) can be expressed in two parts as

\[
\left| \tilde{I}_\mu(f_D) - I_\mu(f) \right| \leq \left| \tilde{I}_\mu(f_D) - I_{\mu}^{Dk}(f_D) \right| + \left| I_{\mu}^{Dk}(f_D) - I_\mu(f) \right|. \tag{34}
\]

In sections 3.3 and 5.4 we have estimated \( R_{\tilde{I}_\mu}^k \) which is a number such that

\[
\left| \tilde{I}_\mu(f_D) - I_\mu(f) \right| \leq R_{\tilde{I}_\mu}^k \text{ for all } (x_1, \xi_j) \in L_k.
\]

Note that the operators in this definition are defined on different function spaces. Also that our estimates of \( R_{\tilde{I}_\mu}^k \) are depending on \( f \). We shall now work with the error of \( \tilde{I}_\mu \) relative to the discrete operator \( I_{\mu}^{Dk} \) (i.e. the first term on the right side of (34)) and allow the error to depend on the function \( f_D \).

Definition 6.2.2. By the error of \( \tilde{I}_\mu(f_D) \) we shall refer to a number \( \tilde{R}_{\tilde{I}_\mu}^k(f_D) \) such that

\[
\left| \tilde{I}_\mu(f_D) - I_{\mu}^{Dk}(f_D) \right| \leq \tilde{R}_{\tilde{I}_\mu}^k(f_D) \text{ for all } (x_1, \xi_j) \in L_k. \tag{35}
\]

This is not the same as the error \( R_{\tilde{I}_\mu}^k \) that we have previously estimated (but given sufficiently fine mesh-size it may be justified to approximate the errors as same \( ^4 \)).

\(^4\)By remark 6.2.1 there are restrictions that can be put on \( f \) so that \( \tilde{R}_{\tilde{I}_\mu}^k \approx R_{\tilde{I}_\mu}^k \) or \( \tilde{R}_{\tilde{I}_\mu}^k = R_{\tilde{I}_\mu}^k \). See also B.II.1, Appendix B, concerning the definition of error \( \tilde{R}_{\tilde{I}_\mu}^k \).
For any two functions $f_\mathcal{D} : \mathcal{D} \mapsto \mathbb{R}$ and $g : \mathcal{L}_k \mapsto \mathbb{R}$

$$\left| \langle f_\mathcal{D}, \tilde{T}_\mu^*(g) \rangle - \langle f_\mathcal{D}, I_\mu L_k^*(g) \rangle \right| = \left| \langle g, \tilde{T}_\mu(f_\mathcal{D}) \rangle - \langle g, I_\mu L_k^*(f_\mathcal{D}) \rangle \right| \leq g_{\text{max}} |\mathcal{L}_k| \tilde{R}_k^{L_k}(f_\mathcal{D}).$$  \hspace{1cm} (36)

We now define the operator

$$E := \tilde{T}_\mu - I_\mu L_k.$$  

Now since $E^* = \tilde{T}_\mu^* - I_\mu L_k^*$ we put $f_\mathcal{D} := E^*(g)$ in the left side of (36) whereupon it takes the form

$$\left| \langle f_\mathcal{D}, \tilde{T}_\mu^*(g) \rangle - \langle f_\mathcal{D}, I_\mu L_k^*(g) \rangle \right| = \| \tilde{T}_\mu^*(g) - I_\mu L_k^*(g) \|^2$$

thus

$$\| \tilde{T}_\mu^*(g) - I_\mu L_k^*(g) \|^2 \leq g_{\text{max}} |\mathcal{L}_k| \tilde{R}_k^{L_k} \left( \tilde{T}_\mu^*(g) - I_\mu L_k^*(g) \right).$$  \hspace{1cm} (37)

Hence, for a function $g : \mathcal{L}_k \mapsto \mathbb{R}$, formula (37) gives an estimate for the error of the adjoint in terms of

$$\| \tilde{T}_\mu^*(g) - I_\mu L_k^*(g) \|^2 = \sum_{\mathcal{D}} \left( \tilde{T}_\mu^*(g) - I_\mu L_k^*(g) \right)^2$$

whenever $\tilde{R}_k^{L_k}$ (definition 6.2.2) can be estimated for the function $E^*(g) : \mathcal{D} \mapsto \mathbb{R}$.

### 6.2.3 An estimate in terms of operator norm

We shall again look at the accuracy of $\tilde{T}_\mu$ relative to $I_\mu L_k$ (i.e. the first term on the right side of the inequality (34)). The accuracy can be expressed in terms of the norm (formula (33)) of the operator $E = \tilde{T}_\mu - I_\mu L_k$. Conversely, the norm of $E^*$ provides information of the accuracy of $\tilde{T}_\mu^*$. In fact, provided that the operators are linear (see remarks 6.0.1 and 6.2.2) $E$ is a bounded operator $^5$ and $\|E^*\| = \|E\|$. This is a result of the following theorem.

**Theorem 6.1.** If $L$ is a linear operator and the operator norm is as defined by equation (33), then

$$\|L^*\| = \|L\|$$

To prove theorem 6.1 we shall use the following common relations for function norm and operator norm

$$\|\langle f_1, f_2 \rangle\| \leq \|f_1\| \|f_2\|$$

$$\|L(f_1)\| \leq \|L\| \|f_1\|$$

$^5$Linear operators are bounded. See i.e. [4] for a proof.
7 SOME ISSUES REGARDING IMPLEMENTATION

where $f_1$ and $f_2$ are functions in the domain of $L$ which is a bounded operator (for proof see i.e. [4]). Now, if $g$ is any function in the domain of $L^*$
\[
\langle L^*(g), L^*(g) \rangle = |\langle L(L^*(g)), g \rangle| \leq \|L(L^*(g))\| \|g\| \leq \|L\| \|L^*\| \|g\|^2
\]
therefore
\[
\|L^*\| = \sup_{\|g\|=1} \left\{ \sqrt{\langle L^*(g), L^*(g) \rangle} \right\} \leq \sqrt{\|L\| \|L^*\|}
\]
thus $\|L^*\| \leq \|L\|$. Since $L^{**} = L$ we have analogously that $\|L\| \leq \|L^*\|$ thus consequently $\|L^*\| = \|L\|$. □

It is now clear that the error in terms of operator norm propagates from $\tilde{I}_\mu$ to $\tilde{I}_\mu^*$ such that:
\[
\|\tilde{I}_\mu^* - I_{\mu k}^*\| = \|\tilde{I}_\mu - I_{\mu k}\|
\]
In other words, in terms of discrete operator norm, the back-projection algorithm has the same accuracy as the forward projection algorithm.

7 Some issues regarding implementation

In this section we shall mention some important issues to handle if the algorithm is to be implemented.

7.1 Enumeration of the sets

A big concern in an implementation is how to enumerate and store the sets $\mathcal{L}_k$. We have seen in section 2 that for $k > 0$ the set $\mathcal{L}_k$ is a proper subset of $\Sigma_k \times \hat{\Sigma}_k$ (see also Appendix A regarding the properties of the sets). The most handy approach may be to allocate in each iteration an entire $\Sigma_k \times \hat{\Sigma}_k$-matrix. In fact it may be even larger, as we shall see below in section 7.2.

Regarding $\mathbb{R}^n$ where $n > 2$ the algorithm can first be implemented in $\mathbb{R}^2$. The 2-dimensional algorithm can then be executed on 2-dimensional planes in $\mathbb{R}^3$ with different orientations. Note that in the univariate parameterization of section 5.2 the algorithm is not more complex on a plane where the angular variables may have two $2^k$-components than on planes where it has at most one. In four dimensions such a 3-dimensions algorithm can possibly be implemented on 3-dimensional hyperplanes and so on.

7.2 Line segments not in the sets

By remark 3.2.1 we have defined level $k$ by operations partially on line segments that are not in $\mathcal{L}_{k-1}$. This is shown in figure 7 and realized by the fact that for a given cell (definition 2.1.3) of level $k$, e.g. the cell $[a, a + 2^k h]$, we shall compute integrals on line segments that lies partially outside the cell. This calls for presence, in level $k - 1$, of integrals on line segments $(x, \xi)$ that are only intersecting the cell in the point $x$. These in turn, are defined by integrals on line
7 SOME ISSUES REGARDING IMPLEMENTATION

Figure 7: Line segments that are partially or totally outside the sets.

Figure 8: Parallel line segments intersecting a cell with convex subset enclosed

segments that does not intersect the cell in any point. In an implementation this can be taken care of by making some abundance of space in the data-matrices allocated. On the other hand there are line segments by definition in \( L_k \) that has no interest at all, such as the line segment given by \((a, \xi_j)\) when components of \( \xi_j \) are negative. In this sense the sets \( L_k \) are not optimal.

7.3 Restrictions for angle of incidence

By remark 2.1.1, section 2, the angle of incidence has a limitation for the line segments intersecting a cell. This is also shown in figure 2 in the same section. In this sense we may not be able to claim that level \( k \) provides a good discretization of the X-ray transform on the cell. Particularly if the orientation of the line segments has importance (i.e. attenuation). However we may be interested in a convex subset enclosed in the cell, see figure 8. Such a set has a higher limitation for angle of incidence.

It is also possible to alter the algorithm in order to compute two integrals (once
in each direction) on every line segment, or some. Such an alteration shall also increase the angle of incidence for some line segments.

References


Appendix A Sizes

In this appendix we will investigate the sizes, i.e. number of elements, of the sets defined in section 2.1. The contents of this appendix is in no way crucial for the description of the algorithm. However, some formulae derived here are referred to from previous sections. Most important are (49) and (50), expressing the size of \( \mathcal{L}_k \), that are useful in section 6.2. Before we continue, let us introduce two additional sets of indexes. For \( 1 \leq \nu \leq n \) we define

\[
J'_k := \{ l \in J_k : \text{There are precisely } \nu \text{ components of } l \text{ that divides } 2^k \}
\]

\[
\hat{J}'_k := \{ j \in \hat{J}_k : \text{There are precisely } \nu \text{ components of } j \text{ that equals } 2^k \}.
\]

For \( 0 \leq \nu \leq n \), we also introduce two subsets \( \Sigma'_k \subset \Sigma_k \) and \( \hat{\Sigma}'_k \subset \hat{\Sigma}_k \) which are defined as

\[
\Sigma'_k := \{ a \oplus h \odot l : l \in J'_k \}
\]

\[
\hat{\Sigma}'_k := \{ h \odot j : j \in \hat{J}'_k \}.
\]

These subsets will in turn provide a means to investigate the set \( \mathcal{L}_k \). When deriving the formulae of this appendix the following four terms will be used:

When \( N \) is chosen as in section 2.1.1 there is in the interval \([0, N]\) the number

\[
1 + \frac{N}{2^k}
\]

of integers that are multiples of \( 2^k \). And since the total number of integers in the interval is \( 1 + N \), there are

\[
N - \frac{N}{2^k}
\]

integers that are not multiples of \( 2^k \). There are in the interval \([-2^k, 2^k]\) totally

\[
2^{k+1} + 1
\]

integers and, excluding the endpoints, there are in \([-2^k, 2^k]\] totally

\[
2^{k+1} - 1
\]

integers.

A.I Size of the set \( \Sigma_k \)

The set \( \Sigma_0 \), defined in section 2.1.1, by (39) contains \( \prod_i^n (1 + N_i) \leq (1 + N_{\text{max}})^n \) points. Particularly if \( N_{\text{min}} = N_{\text{max}} := N \) then

\[
|\Sigma_0| = (1 + N)^n
\]

In every level \( k \) the set is reduced to \( \Sigma_k \) containing only those elements \( x_l = a + l \odot h \) where \( 2^k \| l_i \) for some \( i = 1, \ldots, n \). By (40) there are in the \( i \)'th
dimension \( N_i - \frac{N}{2^k} \) points that are not multiples of \( 2^k \). In the set \( \Sigma_k \) all \( n \)-tuples of such points are removed thus the number of points remaining is limited by:

\[
(1 + N_{\min})^n - \left( N_{\max} - \frac{N_{\max}}{2^k} \right)^n \leq |\Sigma_k| \leq (1 + N_{\max})^n - \left( N_{\min} - \frac{N_{\min}}{2^k} \right)^n
\]

In the case \( N_{\min} = N_{\max} := N \) there is the size of the set:

\[
|\Sigma_k| = (1 + N)^n - \left( N - \frac{N}{2^k} \right)^n
\]  (44)

**Size of the set \( \Sigma_k^\nu \)**

We will also derive a formula for the number of points in \( \Sigma_k \) that has \( \nu \) exactly \( 2^k \)-multiples, i.e. all elements \( x_l = a + l \odot h \) that has \( \nu \) indexes \( i \) s.t. \( 2^k | i \). This subset will be denoted \( \Sigma_k^\nu \). For instance in figure 1, section 2, we can count to 15 members of \( \Sigma_{2^1} \subset \Sigma_1 \) and 6 members of \( \Sigma_{2^2} \subset \Sigma_2 \), while the remaining points of the sets form \( \Sigma_{1^1} \) and \( \Sigma_{1^2} \) respectively. For deriving a general formula, first \( \nu \) positions in the vector \( x \) can be selected in \( \binom{n}{\nu} \) ways. By (39) these \( \nu \) components can be initiated in \( \prod_\nu (\frac{N_{\min}}{2^k} + 1) \) different ways (where the product is over \( \nu \) terms). The remaining \( n - \nu \) components can be initiated in \( \prod_{\nu=0}^{n-\nu} (N_{\max} - \frac{N_{\min}}{2^k}) \) ways. This gives the following limitations for the size of the set \( \Sigma_k^\nu \):

\[
\binom{n}{\nu} \left( \frac{N_{\min}}{2^k} + 1 \right) \nu \left( N_{\min} - \frac{N_{\min}}{2^k} \right)^{n-\nu} \leq |\Sigma_k^\nu| \leq \binom{n}{\nu} \left( \frac{N_{\max}}{2^k} + 1 \right) \nu \left( N_{\max} - \frac{N_{\max}}{2^k} \right)^{n-\nu}
\]

Supposed \( N_{\min} = N_{\max} := N \) we have:

\[
|\Sigma_k^\nu| = \binom{n}{\nu} \left( \frac{N}{2^k} + 1 \right) \nu \left( N - \frac{N}{2^k} \right)^{n-\nu}
\]  (45)

Looking at formula (45) as a binomial expansion where the number of \( 2^k \)-multiples, i.e. \( \nu \), varies between 1 and \( n \) we arrive back at (44) by checking that:

\[
\sum_{\nu=1}^{n} |\Sigma_k^\nu| = \sum_{\nu=0}^{n} \binom{n}{\nu} \left( \frac{N}{2^k} + 1 \right) \nu \left( N - \frac{N}{2^k} \right)^{n-\nu} = (1 + N)^n - \left( N - \frac{N}{2^k} \right)^n \]
A.II  Size of the set  $\hat{\Sigma}_k$

To get the number of elements in the set $\hat{\Sigma}_k$, defined in section 2.1.1, we start with the set of $n$-tuples of integers that are less than or equal to $2^k$ in absolute value. By (41) there are $2^{2k}+1$ such integers. Although since we want the equality $|j_i| = 2^k$ fulfilled in at least one component we remove the set of integer $n$-tuples that are strictly less in all components. By (42) there are $2^{2k}−1$ such integers. The size of the set is therefore:

$$\left|\hat{\Sigma}_k\right| = (2^{2k}+1)^n - (2^{2k}−1)^n$$  \hspace{1cm} (46)

Particularly, in level zero

$$\left|\hat{\Sigma}_0\right| = 3^n − 1$$  \hspace{1cm} (47)

Size of the set  $\hat{\Sigma}_\nu^k$

We shall also investigate the size of the subset $\hat{\Sigma}_\nu^k$ of $\hat{\Sigma}_k$ where every member has $\nu$ exactly $2^k$-components. There are $\binom{n}{\nu}$ ways to pick those components. They can each be initiated in 2 ways and the remaining $n−\nu$ components can each be initiated in $(2^{k+1}−1)$ different ways. Therefore:

$$\left|\hat{\Sigma}_\nu^k\right| = \binom{n}{\nu} 2^\nu (2^{k+1}−1)^{n−\nu}$$  \hspace{1cm} (48)

Every member of $\hat{\Sigma}_k$ has at least one and at most $n$ $2^k$-components and we arrive back at (46) by the sum:

$$\sum_{\nu=0}^{\nu=n} \binom{n}{\nu} 2^\nu (2^{k+1}−1)^{n−\nu} = (2^{k+1}+1)^n - (2^{k+1}−1)^n = \left|\hat{\Sigma}_k\right|$$

A.III  Size of the set  $\mathcal{L}_k$

A formula can now be derived for the size of the set $\mathcal{L}_k$. In doing so we need to involve the set $\hat{\Sigma}_k \subset \mathbb{R}^n$, as defined in 2.1.1 but where $m \leq n$. Therefore a small addition to the notations of section 2.1.1 will be made to denote the dimensions of the vectors in these sets and we write:

$$\Sigma_k := \Sigma_k \subset \mathbb{R}^n, \Sigma_k := \hat{\Sigma}_k \subset \mathbb{R}^m$$

For $k > 0$ the set $\mathcal{L}_k$ is a proper subset of $\Sigma_k \times \hat{\Sigma}_k$ and $|\mathcal{L}_k| < |\Sigma_k||\hat{\Sigma}_k|$. A member $x_l$ of $\Sigma_k$ can be combined with $\xi_j \in \hat{\Sigma}_k$, to a line segment in $\mathcal{L}_k$ if and only if at least one $2^k$-component of $\xi_j$ coincides with a $2^k$-multiple component of $x_l$, see 2.1.3. Let’s assume that $x_l \in \Sigma_k^\nu$ that is $x_l$ has $\nu$ exactly $2^k$-multiples.
Then, corresponding to those \( \nu \) components, we choose a vector from \( \hat{\Sigma}_k \) to constitute a part of \( \xi_j \). The remaining \( n - \nu \) components of \( \xi_j \) can be initiated (with or without \( 2^k \)-components) in \( (2^k + 1)^{n-\nu} \) ways. This gives the formula:

\[
|\mathcal{L}_k| = \sum_{\nu=1}^{n} \left| \sum_{\nu}^{n} \hat{\Sigma}_k \right| (2^{k+1} + 1)^{n-\nu} \tag{49}
\]

By (49) and the formulae of section A.II we can now verify the relation

\[
|\mathcal{L}_k| \leq \left| \sum_{k}^{n} \hat{\Sigma}_k \right|
\]

by substituting formula (46) in (49):

\[
|\mathcal{L}_k| = \sum_{\nu=1}^{n} \left| \sum_{\nu}^{n} \hat{\Sigma}_k \right| [(2^{k+1} + 1)^{\nu} - (2^{k+1} - 1)^{\nu}] (2^{k+1} + 1)^{n-\nu} \\
\leq \sum_{\nu=1}^{n} \left| \sum_{\nu}^{n} \hat{\Sigma}_k \right| [(2^{k+1} + 1)^{n} - (2^{k+1} - 1)^{n}] \\
= \sum_{\nu=1}^{n} \left| \sum_{\nu}^{n} \hat{\Sigma}_k \right| \\
= \left| \sum_{k}^{n} \hat{\Sigma}_k \right|
\]

Finally, in the case that \( N_{\min} = N_{\max} := N \), by substituting (45) and (46) into (49), and then using (44) a closed form expression for \( |\mathcal{L}_k| \) can be derived:

\[
|\mathcal{L}_k| = \sum_{\nu=1}^{n} \left| \sum_{\nu}^{n} \hat{\Sigma}_k \right| [(2^{k+1} + 1)^{\nu} - (2^{k+1} - 1)^{\nu}] (2^{k+1} + 1)^{n-\nu} \\
= (2^{k+1} + 1)^{n} \left| \sum_{k}^{n} \right| \\
- \sum_{\nu=1}^{n} \binom{n}{\nu} \left[ \left( \frac{N}{2^k} + 1 \right) (2^{k+1} - 1) \right]^{\nu} \left[ \left( N - \frac{N}{2^k} \right) (2^{k+1} + 1) \right]^{n-\nu} \\
= [N(2^{k+1} + 1) + 2^{k+1} + 1]^{n} - [N(2^{k+1} + 1 - 2^{1-k}) + 2^{k+1} - 1]^{n} \\
= [(2^{k+1} + 1)(N + 1)]^{n} - [(2^{k+1} + 1)(N + 1) - 2(2^{-k}N + 1)]^{n} \tag{50}
\]

In (50) it can be checked that \( k = 0 \) yields \( |\mathcal{L}_0| = (1 + N)^{n}(3^{n} - 1) \) which is consistent with formulae (43) and (47).
Appendix B Accuracy analysis

In this appendix the errors arising in the 2-dimensional algorithm of section 3 shall be thoroughly analyzed. The error at level zero remains unknown, obviously depending on the method that level zero has been obtained. This unknown error will also be taken into account however. In section B.I we shall consider the error arising in a single upgrade of level and in B.II all errors are propagated through all levels and a closed form expression bounding the error at level \( k \) is developed. We then make a few conclusions concerning the shrink-rate of the error with decreasing \( h \) and the correlation of interpolation errors. Results of the calculations in this entire appendix are given in the main text in section 3.3 and they are also summed-up in the final section B.VII of the appendix. The most efficient way to study this appendix may actually be to start in the final section and follow the references back into the text.

B.I Error analysis at a single upcount of level

In this section we will propagate errors from level \( k \) to level \( k + 1 \). The big interest is to investigate the error arising by using \( \tilde{J} \) (sections 3.2.1-3.2.2) as an approximation for \( \tilde{I} \). The following well known theorem shall be used:

**Theorem B.1 (Linear interpolation error).** Let \( f \in C^2 \) and \( p_1(x) \) be the linear polynomial that interpolates \( f \) at \( x_0 \) and \( x_1 \). Then, for all \( x \in [x_0, x_1] \)

\[
|f(x) - p_1(x)| \leq \frac{1}{8}(x_1 - x_0)^2 \max_{x \in [x_0, x_1]} |f''(x)|
\]

*Prooving theorem B.1.* The proof follows from Rolles theorem, which is a special case of the mean value theorem. See e.g. [2] for the proof.

**Error of \( \tilde{I}^{k} \)**

First we consider the error of \( \tilde{I}^{k}(x_l, \xi_j) \). By the error of \( \tilde{I}^{k}(x_l, \xi_j) \) we mean a number \( R_{\tilde{I}}^{k}(\xi_j) \) such that

\[
\left| \tilde{I}(x_l, \xi_j) - I(f)(x_l, \xi_j) \right| \leq R_{\tilde{I}}^{k}(\xi_j) \quad (51)
\]

for fixed \( \xi_j \in \Sigma^k \) and all \( x_l \) such that \( (x_l, \xi_j) \in \mathcal{L}_k \). In other words we allow the error of level \( k \) to be dependent of the angular variable. For now, no further assumptions are made about this error \( R_{\tilde{I}}^{k}(\xi_j) \).

**Error of \( J^{k} \)**

Now to the error of \( J^{k}(x_l, \frac{\xi_j}{2}) \). By this we mean analogously a number \( R_{J}^{k}(\frac{\xi_j}{2}) \) such that

\[
\left| J \left( x_l, \frac{\xi_j}{2} \right) - I(f) \left( x_l, \frac{\xi_j}{2} \right) \right| \leq R_{J}^{k} \left( \frac{\xi_j}{2} \right)
\]
although this time \((x_l, \xi_j) \in \mathcal{L}_{k+1}\). Since there are two different ways to compute \(J^k\) we shall look at two cases:

**Case 1:** \(j^{k+1}\) has no component that is an odd number.

This case is simple since \(\xi_j(0) = \xi_j(2)\) (section 3.2.2) and \(J^k\) is just a copy of a member \(\tilde{I} \left(x_l, \xi_j(2)\right)\) of level \(k\). The error is propagated straight from that member, i.e:

\[
R^k_{J} \left(\frac{\xi_j}{2}\right) := R^k_{\tilde{I}} \left(\frac{\xi_j}{2}\right) \quad (52)
\]

**Case 2:** \(j^{k+1}\) has one component that is an odd number. Now \(\xi_j(0) \neq \xi_j(2)\) and \(J^k\) is computed as a mean of two level-\(k\) members \(\tilde{I} \left(x_l, \xi_j(0)\right)\) and \(\tilde{I} \left(x_l, \xi_j(2)\right)\). These two have errors \(R^k_{\tilde{I}} \left(\frac{\xi_j}{2}\right)\) and \(R^k_{\tilde{I}} \left(\frac{\xi_j}{2}\right)\) respectively.

Before we continue we shall investigate the most important source of error, obviously the interpolation.

**Definition B.I.1.** The interpolation error of \(I\) is a number \(\mathcal{R}_I\) such that:

\[
\left| I(f) \left(x_l, \frac{\xi_j}{2}\right) - I(f) \left(x_l, \frac{\xi_j(0)}{2}\right) + I(f) \left(x_l, \frac{\xi_j(2)}{2}\right) \right| \leq \mathcal{R}_I
\]

The interpolation error constitutes part of the error \(R^k_{J}\) and it will be shown that it is independent of level. By theorem B.1 the size of \(\mathcal{R}_I\) is:

\[
\frac{1}{8} 2^2 \max_{s \in [0,2]} \left| I''_s(f) \left(x_l, \frac{\xi_j(s)}{2}\right) \right|
\]

where

\[
I''_s(f) \left(x_l, \frac{\xi_j(s)}{2}\right) = I''_s(f) \left(x_l, \frac{1}{2}(j_e + j_o(s)) \odot h\right)
\]

\[
= \frac{\partial^2}{\partial s^2} \int_0^1 f(x_l + \frac{\gamma}{2}(j_e + j_o(s)) \odot h) d\gamma
\]

\[
= \int_0^1 \frac{\partial^2}{\partial s^2} f(x_l + \frac{\gamma}{2}(j_e + j_o(s)) \odot h) d\gamma
\]

\[
= \int_0^1 \frac{\partial^2}{\partial x_l^2} f(x_l + \frac{\gamma}{2}(j_e + j_o(s)) \odot h) \frac{\gamma^2 h^2}{4} d\gamma
\]

\[
= \int_0^1 \frac{\partial^2}{\partial x_l^2} f(x_l + \frac{\gamma}{2}(j_e + j_o(s)) \odot h) \frac{h^2}{12} d\gamma^3
\]

and

\[
\left| I''_s(f) \left(x_l, \frac{\xi_j(s)}{2}\right) \right| \leq \frac{1}{12} \max \left\{ \frac{\partial^2 f}{\partial x^2} \right\}_{\text{max}} \left| h_i \right|^2
\]

Suppose now that there is a number \(\left| \frac{\partial^2 f}{\partial x^2} \right|_{\text{max}}\) which is a bound for the absolute value of the second partial derivatives of \(f\) and that the maximal component of
**APPENDIX B  ACCURACY ANALYSIS**

$h$ is denoted $h_{\text{max}}$. Then the following bound for the interpolation error can be used:

$$R^I := \frac{1}{24} \left| \frac{\partial^2 f}{\partial x^2} \right|_{\text{max}} h_{\text{max}}^2$$

(53)

Note that the interpolation error $R^I$ is independent of level. In order to express the error $R^k_J$ in terms of the interpolation error, we first note that:

$$\left| \mathcal{J}\left( x_l, \frac{\xi_j}{2} \right) - \mathcal{I}(f)\left( x_l, \frac{\xi_j}{2} \right) \right| \leq \frac{R^k_J\left( \frac{\xi_j(0)}{2} \right) + R^k_J\left( \frac{\xi_j(2)}{2} \right)}{2} + R^I$$

Thus, by triangular inequality:

$$\left| \mathcal{J}\left( x_l, \frac{\xi_j}{2} \right) - \mathcal{I}(f)\left( x_l, \frac{\xi_j}{2} \right) \right| \leq \frac{R^k_J\left( \frac{\xi_j(0)}{2} \right) + R^k_J\left( \frac{\xi_j(2)}{2} \right)}{2} + R^I$$

Where $R^I$ is the interpolation error. $\mathcal{I}(f)\left( x_l, \frac{\xi_j}{2} \right)$. We can sum these results to the following formula for the error of $\mathcal{J}^k$:

$$R^k_J\left( \frac{\xi_j}{2} \right) = \begin{cases} R^k_J\left( \frac{\xi_j}{2} \right) & \text{if } \frac{\xi_j}{2} \in \Sigma^k \\ \frac{R^k_J\left( \frac{\xi_j(0)}{2} \right) + R^k_J\left( \frac{\xi_j(2)}{2} \right)}{2} + R^I & \text{otherwise} \end{cases}$$

(54)

Where $R^I$ is the bound given by equation (53).

**Error of $\tilde{I}^{k+1}$**

Finally the error of $\mathcal{J}^k$ shall be propagated to $\tilde{I}^{k+1}$. $\tilde{I}^{k+1}\left( x_l, \xi_j \right)$ is computed as the mean of $\mathcal{J}^k\left( x_l, \frac{\xi_j}{2} \right)$ and $\mathcal{J}^k\left( x_l + \frac{\xi_j}{2}; \frac{\xi_j}{2} \right)$ and the following claim shall be made.

**Claim B.1.** The error $R^k_J$ derived for $\mathcal{J}^k\left( x_l, \frac{\xi_j}{2} \right)$ applies also to $\mathcal{J}^k\left( x_l + \frac{\xi_j}{2}, \frac{\xi_j}{2} \right)$.

This claim can be made because in the case that $\mathcal{J}^k$ is computed as a mean, see (9), the same angular variables, $\xi_j(0)/2$ and $\xi_j(2)/2$ are used in both terms, thus the errors propagated from level $k$-members are the same. also the interpolation errors are the same in both cases because of symmetry (equations (7) and (9)).

Because of claim B.1 it is now shown that

$$R^{k+1}_{\tilde{I}}(\xi_j) = R^k_J\left( \frac{\xi_j}{2} \right)$$

(55)

where $R^k_J\left( \frac{\xi_j}{2} \right)$ is as given by equation (54).
Now equations (53), (54) and (55) provide means to propagate errors through an upcount of level. The prerequisite for doing so is the assumption that the error of level \( k \) is depending only on the angular variable. This can be assured by assigning a global bound of the errors at level zero, which will be done in the following section.

The next step in the error analysis is by recursive formulae to investigate the angular error distribution at level \( k \).

**B.II Angular distribution of total error at level \( k \)**

In this section errors shall be propagated from level zero to any level \( k \) and all the interpolation errors arising on the way shall be taken into account. The error in the zeroth level is undetermined and it depends on the way the the zeroth level is calculated.

**Error of level zero**

About the error at level zero, the only assumption to be made here is that there is a number \( \mathcal{R}^0 \) such that

\[
\left| \tilde{I}^0(f_D) - I(f) \right| \leq \mathcal{R}^0 \text{ for all } (x_l, \xi_j) \in \mathcal{L}_0
\]

**Tabulation of angular error distribution**

By equations (53), (54) and (55) we can now tabulate the error distribution over the angular variable. See table 1. \( \mathcal{R}_k^j(\xi_j) \) is tabulated from level one to four in terms of interpolation error \( \mathcal{R}^l \) and level-zero error \( \mathcal{R}^0 \). The table treats the angular interval given by

\[
\frac{j_2}{j_1} \in [0, 1] \quad \text{and} \quad j_1, j_2 \geq 0
\]

(\( \xi_j = j \circ h \)). The level-zero error always propagates as one unit-\( \mathcal{R}^0 \) over the whole distribution. In figure 9 the angular distribution of \( \mathcal{R}_k^j(\xi_j) \) excluding \( \mathcal{R}^0 \) is depicted for some selected levels up to 15. Level 15 means that the pathlength is doubled 15 times, thus the integration paths are \( 2^{15} = 32768 \) times long as the paths of level zero.

**Remark B.II.1.** An alternative way to define the error instead of the inequality (51) is as a number number \( \tilde{\mathcal{R}}_k^j(\xi_j) \) such that

\[
\left| \tilde{I}(x_l, \xi_j) - \tilde{I}(f)(x_l, \xi_j) \right| \leq \tilde{\mathcal{R}}_k^j(\xi_j)
\]

By this definition\(^6\) there is no error in level zero (section 3.1). This is the error depicted as the distribution in figure 9.

---

\(^6\)This definition is particularly useful in section 6.2 where it will be shown that it can be propagated to the adjoint operator.
When looking at the angular error-distribution it has to be kept in mind that we are working with weighted projection and the error has to be multiplied by pathlength for obtaining the error of $P(x_1, \xi_j)$ (equation (1)). This distribution shall be investigated further. The mean and an upper bound of the distribution in figure 9 shall be calculated. We shall then remove one endpoint and consider the semi-open angular interval given by

\[ j_2 \in [0, 1[ \quad \text{and} \quad j_1, j_2 \geq 0 \quad (57) \]

(where $\xi_j = j \circ h$). This interval is, for reasons of symmetry, representative of the entire distribution over $\xi_j$.

**Mean of distribution**

Let $n_k$ denote the size of the subset of $\hat{\Sigma}^k$ that fulfills the criterion of (57). In other words $n_k$ is the number of elements in the $k$'th row of table 1 excluding one of the outer columns according to (57). In level zero $n_0 = 1$ (since only $\xi_j = (h_1, 0)$ is in the interval in question), and in every upcount of level $n_k$ is doubled. Thus

\[ n_k = 2^k \]

Now let $s_k$ denote the sum of all errors arising by interpolation for members of the interval, in level $k$. That is, $s_k$ is the sum of all coefficients for $R^I$ in the $k$'th row of table 1, although one of the outer columns has been excluded. It follows from formula (54) that each error present in level $k - 1$ propagates times two into level $k$ (since it propagates $1 + 2^\frac{1}{2}$ times). Also, in level $k$ there are $n_{k-1}$ new interpolation errors. For $s_k$ therefore there is the following formula:

\[ s_k = 2s_{k-1} + 2^{k-1} \quad \text{and} \quad s_0 = 0 \]

\[ \implies s_1 = 1, \ s_2 = 2 + 2, \ s_3 = 2^2 + 2^2 + 2^2, \ldots \ s_k = k \ 2^{k-1} \]
Figure 9: Angular distribution of error $R_k^i(\xi_j)$ excluding the initial error $R^0$. 
And hence if we denote by \( \bar{R}^k \), the mean of the error-distribution of figure 9 and table 1 (excluding one endpoint), including the level-zero error there is the formula:

\[
\bar{R}^k = \frac{s_k}{n_k} R^I + R^0 = \frac{k^{2^{k-1}}}{2k} R^I + R^0 = \frac{k}{2} R^I + R^0 \tag{58}
\]

**Bound of distribution**

The maximum of the distribution does not come in closed form as easy as the mean (58). However, an upper bound which is very close to the maximum can be derived. Looking at the way the interpolation error (coefficient for \( R^I \)) propagates in table 1 gives the method. First, in level 1 the only interpolation error is one unit \( R^I \) for \( j_2/j_1 = 1/2 \). In the second level, interpolating in either direction from this angle will yield a bigger error \( 3/2 R^I \). In the third level, interpolating between the two previous maxima will yield again bigger \( 9/4 R^I \). If we denote the coefficient of \( R^I \) at level \( k \) by \( a_k \) then the recursion and initial conditions for the sequence \( \{a_k\} \) are:

\[
a_k = \frac{a_{k-1} + a_{k-2}}{2} + 1, \quad a_0 = 0, \quad a_1 = 1 \tag{59}
\]

Continuing this sequence corresponds to the error arising by interpolating in each level between the position of previous maximal error and the one before previous. This recursive formula gives a bound for the maximal error of level \( k \) (a bound expressed as \( a_k R^I + R^0 \)). In order to prove this statement it has to be certain that the maximum of the distribution in every level can be found between the two previous maxima.

**Claim B.2.** If the maxima \( R_{\text{max}}^k \) and \( R_{\text{max}}^{k-1} \) of levels \( k \) and \( k-1 \) are neighbors in the distribution of table 1, then the maximum of level \( k+1 \) will be found in between.

**Proof of claim.** This claim is true because of the fact that for every two neighboring errors of the distribution one of them has remained unchanged since the previous level (since only one is the result of a new interpolation). Therefore, in order for the claim to be false there has to be errors \( R^k \) and \( R^{k-1} \) of levels \( k \) and \( k-1 \) such that either

\[
\frac{R^{k-1} + R^k}{2} + 1 > \frac{R_{\text{max}}^{k-1} + R_{\text{max}}^k}{2} + 1
\]

or

\[
R^k > \frac{R_{\text{max}}^{k-1} + R_{\text{max}}^k}{2} + 1 \Rightarrow R_{\text{max}}^k > R_{\text{max}}^{k-1} + 2
\]

which both are impossibilities.

Formula (59) has the recursion of the linear function \( a_k = C + \frac{3}{2} k \), although the initial conditions \( a_0 = 0 \) and \( a_1 = 1 \) do not apply to that function. However
APPENDIX B  ACCURACY ANALYSIS

the sequence \( \{a_k\} \) can be trapped between two such functions:

\[
\begin{align*}
  a_k & \geq \frac{2k}{3} & \text{where equality holds only for } k = 0 \\
  a_k & \leq \frac{2k+1}{3} & \text{with equality only for } k = 1
\end{align*}
\]

This is proved by induction.

**Proof.** Suppose that \( a_{k-1} > \frac{2(k-1)}{3} \) and \( a_{k-2} > \frac{2(k-2)}{3} \). These conditions can be verified from table 1 when \( k = 3 \).

Then \( a_k = \frac{a_{k-1} + a_{k-2}}{2} + 1 > \frac{4k-6}{6} + 1 = \frac{2k}{3} \).

The proof of \( a_k < \frac{2k+1}{3} \) is analogous with initial conditions holding for \( k = 4 \). \( \square \)

Maximum of the angular error-distribution can now be trapped between linear functions of \( k \).

\[
\frac{2k}{3} R^I + R^0 \leq \max_{\xi_j \in \hat{\Sigma}} \{ R^I_k(\xi_j) \} \leq \frac{2k+1}{3} R^I + R^0
\]

Combining the upper bound of the angular error distribution just derived with the bound of the interpolation error (53), one finds that there is a number

\[
R^k := \frac{2k+1}{72} \left| \frac{\partial^2 f}{\partial x^2} \right|_{\max} h_{\max}^2 + R^0
\]

such that \( \left| \tilde{T}^k - I(f) \right| \leq R^k \) for all \( (x_l, \xi_j) \in \mathcal{L}_k \). However useful formula (60) may be for estimating error it is somewhat insufficient for shrinking the error by shrinking \( h_{\max} \). The reason is the \( k \)-dependence of the error which has to be taken into account in what ever application where integrals exceeding a certain pathlength shall be computed in the final level (e.g. tomographic applications). This needs to be investigated a little further.

**B.III  \( h \)-dependence of \( k \)**

When looking at an error estimate depending on both \( h_{\max} \) and \( k_{\max} \) it has to be kept in mind that in order to achieve a certain integration pathlength \( L = |\xi| \), in the final level \( k_{\max} \), the mesh size \( (h_{\max}) \) may not be shrunken arbitrarily without increasing \( k_{\max} \). This dependence occurs whenever \( h_{\max} \) is shrunken so that \( h_{\max} = h_{\min} \). This dependence shall now be made clear.

For pathlength \( L = |\xi| \) of level \( k \) there are the limitations:

\[
h_{\min} 2^k \leq L \leq \sqrt{n} h_{\max} 2^k
\]

If we let \( L^k_{\min} \) be the minimal pathlength in level \( k \), then

\[
2^k = \frac{L^k_{\min}}{h_{\min}} \Rightarrow k = \log_2 \left( \frac{L^k_{\min}}{h_{\min}} \right)
\]
We therefore define the function $k_L(h_{\min})$ which is the lowest level that guarantees integral paths greater than or equal to $L$ for a certain size of $h_{\min}$

$$k_L(h_{\min}) = \sup_n \left\{ \log_2 \left( \frac{L}{h_{\min}} \right) \right\}$$

which is bounded by:

$$k_L(h_{\min}) < \log_2 \left( \frac{L}{h_{\min}} \right) + 1$$

Summing-up, this concerns having a $(h_{\max}, k_{\max})$-dependent error estimate that grows with increasing $k_{\max}$ as well as with increasing $h_{\max}$. Suppose then that the error shall be made smaller by shrinking $h_{\max}$ (so that eventually $h_{\max} = h_{\min}$). Then $k_{\max}$ needs to be substituted with one of the expressions given by (61) or (62). This is under the assumption that we desire integrals with pathlength equal to or exceeding a certain number $L$ in the final level $k_{\max}$.

B.IV Shrinkage of error

It can now be shown that the error can be made arbitrarily little by shrinking $h_{\max}$. This is under the assumption that the error $R^0$ in level zero is either negligible or also decreases with shrinking $h_{\max}$. If $k_{\max}$ is a fixed number then the error of any level $k \in \{0, \ldots, k_{\max}\}$ is bounded by the expression given by inserting $k$ in formula (60). However if integrals of pathlength equal to or exceeding some number $L$ shall be computed in the final level, then $L$ rather than $k_{\max}$ is a fixed number. In that case (60) shall have $k$ substituted with the bound for the function $k_L(h_{\min})$ given by (62). For sufficient shrinkage of $h_{\max}$ then $h_{\max} = h_{\min} := h$ and the error in the final level is bounded by:

$$R^k_{L,(h)} := \frac{2k_L(h) + 1}{72} \left| \frac{\partial^2 f}{\partial x^2_{\max}} \right| h^2 + R^0$$

$$= \frac{(2 \log_2 L + 3 + 2 \log_2 \frac{1}{h})}{72} \left| \frac{\partial^2 f}{\partial x^2_{\max}} \right| h^2 + R^0$$

If $R^0$ is neglected we see that the error $R^k_{L,(h)}$ decreases with $h$ like $\log \left( \frac{1}{h} \right) h^2$.

B.V Angle of greatest error

In section B.II a recursive sequence of errors lead to formula (60), the maximal error at level $k$. The values of $\xi_j$ corresponding to the errors in this sequence are in each level the closest approximations to

$$\left| \frac{j_2}{j_1} \right| = \frac{1}{3} \text{ or } \frac{2}{3}$$
This is realized by looking at the sequence of $\frac{j^2}{j^1}$ corresponding to the sequence of maximal errors (see table 1):

$$\frac{1}{2} \pm \frac{1}{4} \mp \frac{1}{8} \pm \frac{1}{16} \mp \ldots = \frac{1}{2} \pm \frac{1}{4} \sum_{i=1}^{\infty} \frac{1}{2^i} \pm \frac{1}{8} \sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1}{3} \text{ or } \frac{2}{3}$$

### B.VI Correlation of interpolation errors

Finally some observation shall be made about the correlation of interpolation errors. The question of correlation of several errors affecting a computation is basically the question of whether or not the errors tend to be of the same sign and in the worst case cause a resulting error to grow close to an upper bound. The subject of correlation shall not be entered deeply into. The aim is to point out what condition makes two or more interpolation errors of neighboring interpolation sites likely to strike in the same direction. In Figure (10) a couple of paths of $\mathcal{L}_{k-2}$, $\mathcal{L}_{k-1}$ and $\mathcal{L}_{k}$ are illustrated. In the level $k-2$ image there are also level curves of the underlying function $f$. The projections $\tilde{I}^{k-1}(x, \xi)$ on the level $k-1$ paths in the illustration has to be computed by interpolation, each one as the mean of four integrals. The right one of these in the figure is an underestimation. This is realized by noting that the path is closer in terms of level curves to those two paths of level $k-2$ where $f$ has a higher value. In other words, the underestimation can be blamed on the increasing denseness of level curves in direction

$$\omega^\perp := (-\omega_2, \omega_1) = \frac{1}{|\xi|}(-\xi_2, \xi_1)$$

which is a direction orthogonally to the angular variable. The denseness of level curves in this direction is given by $\omega^\perp \cdot \nabla f$. For the increase of denseness in this direction there is the second derivative:

$$\omega^\perp \cdot \nabla |\omega^\perp \cdot \nabla f| \neq 0$$
APPENDIX B  ACCURACY ANALYSIS

Up to some approximation (e.g. making a local second order approximation of \( f \)) and for sufficiently short paths this expression can be assumed to be significantly nonzero on the integration path whenever there is a significant interpolation error. The function (64) is a linear combination of the partial second derivatives of \( f \) and it is of a sign and magnitude relevant for the error. Continuity of the partial second derivatives will therefore likely yield the same sign of error on neighboring interpolation sites if they are sufficiently close and has the same or similar orientation (\( \omega^k \)).

By this argumentation it can be assumed that the interpolation errors are positively correlated provided that

\[
f \in C^2
\]

particularly for for sufficiently short integral paths.

B.VII  Conclusions

The most important results of section Appendix B can be summed-up as follows. In B.I formula (53) gives a bound for the error caused by interpolation. It is independent of level. Provided that there is a global bound of the error at level zero as well as a global bound of the partial second derivatives a bound for the error at any level will remain independent of the spatial variable but depend on the angular variable. In B.II the mean of the angular error distribution at level \( k \) is given by formula (58) and the maximum is bounded by (60). Both of these expressions are:

\[
O(kh^2)
\]

If, however, integrals on paths equal to or exceeding a certain length \( L \) are to be computed in the final level then \( k \) in the formulae shall be substituted with the function \( k_L(h) \) of (62). Making this substitution in the bound of the maximal error results in the \( k \)-independent formula (63) bounding the error. The error decreases with shrinking \( h \) as

\[
\log \left( \frac{1}{h} \right) h^2
\]

In B.V it is shown that the values of the angular variable corresponding to the maximal error are in each level the ones closest approximating

\[
\left| \frac{j_2}{j_1} \right| = \frac{1}{3} \text{ or } \frac{2}{3}
\]

In B.VI it is concluded that interpolation errors, of integration paths that are sufficiently close to each other and with similar orientation, can be assumed to be positively correlated provided that

\[
f \in C^2
\]

particularly for sufficiently short paths.
Appendix C  Accuracy analysis

We have in this Appendix generalized parts of the calculations in Appendix B. The results are stated in the main text in section 5.4, where references are made to proofs in this appendix. In two dimensions and with zero attenuation we derived the closed form formula (58) for the mean of the angular error distribution. The maximum of the same distribution given by (60) and (63) where based on the recursive formula (59). The general case is a great deal more complicated. We shall however derive the interpolation errors corresponding to formula (53) in section C.I.

In section C.II it shall be investigated how the errors of the operator $\tilde{E}$ propagates from level $k$ to $k+1$ and in C.III the same investigations are made for the operator $\tilde{T}_\mu$.

In section C.IV it is shown why the second approach (5.3) decreases the errors of $\mathcal{F}$ (5.2), particularly for short integral paths.

In C.V we attempt to use the previous results to make estimates of the total error in level $k$.

Finally in section C.VI the important results of Appendix C are summed-up.

An efficient way to study this appendix may be to start in the final section and follow the references back. Again, section 5.4 in the main text is a more detailed review of the results.

C.I  Interpolation errors

In definition B.I.1 of Appendix B we defined the interpolation error $R^I$ of $I(f)$. Analogous definitions shall be made for the interpolation errors of $E(\mu)$ and $I(\mu,f)$ although the upper bounds that shall now be derived will be depending on level. We shall derive these results for univariate parameterization (equations (20) and (21)) of $\xi_j$. Error-bounds at multivariate parameterization (equations (22) and (23)) has to be the same or better.

**Definition C.I.1.** The interpolation errors $kR^I_E$ and $kR^I_I$ are numbers such that:

$$|E\left(\frac{x_i^k}{2}, \frac{\xi_j^k}{2}\right) - \frac{E\left(\frac{x_i^k}{2}, \frac{\xi_j^k(0)}{2}\right) + E\left(\frac{x_i^k}{2}, \frac{\xi_j^k(2)}{2}\right)}{2}| \leq kR^I_E$$

$$|I_\mu(f)\left(\frac{x_i^k}{2}, \frac{\xi_j^k}{2}\right) - \frac{I_\mu(f)\left(\frac{x_i^k}{2}, \frac{\xi_j^k(0)}{2}\right) + I_\mu(f)\left(\frac{x_i^k}{2}, \frac{\xi_j^k(2)}{2}\right)}{2}| \leq kR^I_I$$

These errors constitute parts of the errors of $\mathcal{F}^{k-1}$ and $\mathcal{J}_\mu^{k-1}$ (formulae (71) and (75)).
The interpolation error $R_E^I$

By theorem B.1 the size of the error $kR_E^I$ is:

$$ \frac{1}{8} 2^2 \max_{s \in [0,1]} \left| \mathcal{E}''(\mu) \left( x_{t \pm} \frac{\xi_j(s)}{2} \right) \right| $$

To derive a bound for the second derivative w.r.t $s$ we shall work with the functions $\mathcal{X}(s)$ and $\mathcal{M}(s)$ defined as:

$$ \mathcal{E}(\mu) \left( x_{t \pm} \frac{\xi_j(s)}{2} \right) = \exp \left\{ - \int_0^1 \frac{\xi_j(s)}{2} \left| \frac{d}{dx} \left( x_{t \pm} + t \frac{\xi_j(s)}{2} \right) \right| dt \right\} $$

For differentiation w.r.t. $s$ the following formula shall be used:

$$ \frac{d^2}{ds^2} \exp \{-\mathcal{X}(s)\mathcal{M}(s)\} = \exp \{-\mathcal{X} \mathcal{M}\} \left[ (\mathcal{X}' \mathcal{M} + \mathcal{M}' \mathcal{X})^2 - (\mathcal{X}'' \mathcal{M} + 2 \mathcal{X}' \mathcal{M}' + \mathcal{M}'' \mathcal{X}) \right] $$

Before differentiating $\mathcal{X}(s)$ and $\mathcal{M}(s)$ a few preparations will be made considering the derivatives of $\mathcal{X}(s)$.

Since $\mathcal{X}(s) = |\xi(s)|/2$ and $|\xi(s)| = \sqrt{\sum \xi_i^2}$, the following identities will be used:

if $|\xi(s)| > 0$ :

$$ \frac{\partial}{\partial s} |\xi(s)| = \sum \xi_i \frac{\partial \xi_i}{\partial s} $$

$$ \frac{\partial^2}{\partial s^2} |\xi(s)| = \sum \left[ \left( \frac{\partial \xi_i}{\partial s} \right)^2 + \xi_i \frac{\partial^2 \xi_i}{\partial s^2} \right] - \left( \frac{\sum \xi_i \frac{\partial \xi_i}{\partial s}}{|\xi(s)|} \right)^2 $$

Also we shall use the following bounds for absolute value and a single component of $\xi_j^k(s)$:

$$ 2^k h_{\min} \leq |\xi_j^k(s)| \leq \sqrt{n} 2^k h_{\max} $$

$$ 0 \leq |\xi_j^k(s)| \leq 2^k h_{\max} $$

The univariate parameterization of $\xi_j$ (section 5.2.2) affects the $m$ components in $j$ that are odd numbers. The derivatives w.r.t $s$ will therefore be nonzero for the $m$ components corresponding to the sequence $\{\alpha_i\}_j$ (definition 5.2.1). Differentiating such a component $\xi(s)_{\alpha_1}$ once w.r.t. $s$ yields $\pm h_{\alpha_1}$ while the second derivative is zero. We get the following bounds for the derivatives of $\mathcal{X}(s)$:

$$ \mathcal{X}' = \left| \frac{\xi(s)}{2} \right|, \mathcal{X}'' = \frac{1}{2} \sum_{\alpha_{i \pm}} h_{\alpha_{i \pm}} |\xi(s)|, \mathcal{X}''' = \frac{1}{2} \left[ \frac{\sum_{\alpha_{i \pm}} h_{\alpha_{i \pm}}^2 |\xi(s)|}{|\xi(s)|} - \frac{\left( \sum_{\alpha_{i \pm}} h_{\alpha_{i \pm}}^2 |\xi(s)| \right)^2}{|\xi(s)|^3} \right] $$

$$ \left| \mathcal{X}' \right| \leq \sqrt{n} 2^{k-1} h_{\max}, \left| \mathcal{X}'' \right| \leq \frac{1}{2} \left( \frac{m h_{\max}^2}{h_{\min}} + \frac{(m h_{\max}^2)^2}{(h_{\min})^3} \right) $$
If there are numbers $\partial \mu_{\max}$ and $\partial^2 \mu_{\max}$ bounding the absolute value of the first and second partial derivatives of $\mu$, then the derivatives of the function $M(s)$ has the bounds:

$$
M = \int_0^1 \mu(x + t\xi(s)/2)\,dt \quad |M| \leq \mu_{\max} \\
M' = \int_0^1 \sum_{i=1}^m \pm h_{\alpha_i} \frac{\partial \mu}{\partial \alpha_i} \,dt^2 \quad |M'| \leq \frac{1}{2}mh_{\max} \partial \mu_{\max} \\
M'' = \int_0^1 \sum_{i=1}^m \sum_{l=1}^m \pm h_{\alpha_i} h_{\alpha_l} \frac{\partial^2 \mu}{\partial \alpha_i \partial \alpha_l} \,dt^3 \quad |M''| \leq \frac{1}{12}m^2h_{\max}^2 \partial^2 \mu_{\max}
$$

Now all the terms in formula (65) can be bounded with the expressions derived.

First, since $X(s)$ and $M(s)$ are positive functions the exponential is bounded by one:

$$
\exp \{ -X(s)M(s) \} \leq 1
$$

In order to simplify expressions bounding the terms of (65), and later to derive a formula for the interpolation error of $I\mu$, we shall make the following definition:

**Definition C.I.2.** If the functions $\mu$ and $f$ are bounded, as well as their partial first and second derivatives, we shall denote by $D$ a positive number bounding the absolute value of zeroth, first and second derivatives of $\mu$ and $f$.

Suppose $h_{\max} = h_{\min} := h$ and that the necessary condition for definition C.I.2 is fulfilled. Then

$$
\begin{align*}
(M'X)^2 &\leq \frac{m^2h^2}{2^4}h^42^k \quad (X'M)^2 \leq \frac{m^2h^2}{2^4}h^42^k \\
|2M'X' - |M'|^2 &\leq \frac{m^2h^2}{2^4}h^42^k \quad |M'X' - |M''|^2 \leq \frac{m^2h^2}{2^4}h^42^k
\end{align*}
$$

and if $D \geq 1$ the following inequality holds:

$$
\left| \frac{\partial^2}{\partial s^2} \mathcal{E}(\mu) \left( x_k^k, \frac{\xi_j^{k}(s)}{2} \right) \right| \leq m^2D^2 \left[ \frac{m^2h^2}{2^4}2^k + \frac{\sqrt{\pi}}{\pi}2^k + \frac{1}{2}h^2 + h2^{-k} \right]
$$

(66)

**Remark C.I.1.** When analyzing the terms of (66) we have to take the $k$-dependence into consideration. If one wants to look at the interpolation error of paths exceeding a certain pathlength $L$, the $k$-dependence has to be substituted with $k = \log_2 \left( \frac{L}{h} \right) + 1$ which is the upper bound of the function $k_L(h)$ defined in equation (61). In that case, for fixed $L$, we have:

$$
2^k \propto \frac{L^2}{h^2} \quad 2^k \propto \frac{L}{h} \quad \frac{1}{2^k} \propto \frac{h}{L}
$$

If we hold $L$ fixed instead of $k$ all of the terms of (66) are proportional to $h^2$.

The interpolation error is now given according to theorem B.1 by substituting (66) in the formula:

$$
R^k_{\mathcal{E}} = \frac{1}{2} \max \left| \frac{\partial^2}{\partial s^2} \mathcal{E}(\mu) \left( x_k^k, \frac{\xi_j^{k}(s)}{2} \right) \right| \leq m^2D^2 \left[ \frac{m^2h^2}{2^4}2^k + \frac{\sqrt{\pi}}{\pi}2^k + \frac{1}{2}h^2 + h2^{-k} \right]
$$

(67)
Application of these results will be discussed in section C.V and section 5.4 of the main text.

The interpolation error \( R^I_\mu \)

By theorem B.1 the size of the error \( R^I_\mu \) is:

\[
\frac{1}{8} \cdot 2^2 \max_{s \in [0,2]} \left| J''_\mu \left( x_1, \frac{\xi_j(s)}{2} \right) \right|
\]

To derive a bound for the second derivative w.r.t \( s \) we shall again work with the functions \( X(s) \) and now \( M_\gamma(s) \) defined as:

\[
I_\mu \left( f \right) \left( x_l, \frac{\xi_j(s)}{2} \right) = \int_0^1 f \left( x_l + \gamma \frac{\xi_j(s)}{2} \right) \exp \left\{ -X(s) \right\} \left\{ \int_0^1 \mu \left( x_l + \frac{t \xi_j(s)}{2} \right) \right\} d\gamma
\]

For differentiation of the integrand w.r.t. \( s \) the following formula shall be used:

\[
\frac{d^2}{ds^2} f(s) \exp \left\{ -X(s)M_\gamma(s) \right\} = \exp \left\{ -X \right\} \left[ f'' - 2f' \cdot (X'M_\gamma + M'_\gamma X) \right]
\]

To find bounds of the terms of (68) we start by noting that since \( \gamma \in [0,1] \) \( M_\gamma \leq M \) (where \( M \) is the function of equation (65)) thus the same bound can be used. Also the bounds of the derivatives of \( M \) applies to \( M_\gamma \) as well for the same reason. Bounds for the derivatives of \( X \) has also been derived following equation (65) hence we only have to differentiate \( f \) w.r.t \( s \) and integrate:

\[
f'_s = \frac{\gamma}{2} \sum_{i=1}^{m} \pm h_{\alpha_i} \frac{\partial f}{\partial x_{\alpha_i}}
\]

\[
f''_s = \frac{\gamma^2}{4} \sum_{i=1}^{m} \sum_{l=1}^{m} \pm h_{\alpha_i} h_{\alpha_l} \frac{\partial^2 f}{\partial x_{\alpha_i} \partial x_{\alpha_l}}
\]

\[
\int_0^1 |f||f'\gamma| \leq |f|_{\max} \frac{\partial f_{\max}}{2h_{\max}}
\]

\[
\int_0^1 |f'|d\gamma = \int_0^1 |f'| \frac{d\gamma^2}{2d\gamma} \leq \frac{1}{2} m h_{\max} \frac{\partial f_{\max}}{\partial x_{\max}^2}
\]

\[
\int_0^1 |f''|d\gamma = \int_0^1 |f''| \frac{d\gamma^3}{3d\gamma^2} \leq \frac{1}{12} m^2 h_{\max}^2 \frac{\partial^2 f_{\max}}{\partial x_{\max}^4}
\]

Where \( m \) is the number of components in \( \xi_j \) that are odd. We shall again use the number \( D \) of definition C.I.2, provided that the necessary conditions are fulfilled. To derive an inequality bounding the derivative of equation (68) all the terms of (65) multiplied by \( D \) can be used (since \( |f| \leq D \)), and in addition:

\[
2f'X'M \leq \frac{m^2 \sqrt{\pi h^2}}{12} 2^k h^3
\]

\[
2f'X'M \leq \frac{m^2 \sqrt{\pi h^2}}{12} 2^k h^3
\]

\[
f'' \leq \frac{m^2 \sqrt{\pi h^2}}{12} 2^k h^3
\]
If \( D \geq 1 \) the derivative (68) is bounded by
\[
\left| \frac{\partial^2}{\partial s^2} I_\mu (f) \left( x^k, \frac{\xi^k(s)}{2} \right) \right| \leq m^2 D^2 \left[ \frac{\pi^2}{64} 2^k h^4 + \frac{\sqrt{\pi}}{4} 2^k h^3 + \frac{5}{6} h^2 + h 2^{-k} \right] \tag{69}
\]
By the \( k \)-dependence the terms are of varying importance depending on the length \( L \) of the integral paths (see remark C.I.1). The expression (69) can now be substituted according to theorem B.1 in the formula:
\[
k^R I = \frac{1}{2} \max \left| \frac{\partial^2}{\partial s^2} I_\mu (f) \left( x^k, \frac{\xi^k(s)}{2} \right) \right| \tag{70}
\]
Application of these results will be discussed in section C.V and section 5.4 of the main text.

C.II Propagation of the error of \( \tilde{E}^k \) to \( \tilde{E}^{k+1} \)

Analogously to the error analysis of section B.II we shall first look at the error of the approximation \( \mathcal{F}^k \) when defined by linear interpolation. The error of \( \mathcal{F}^k \) shall then be propagated to \( \tilde{E}^{k+1} \). We shall also assume that a bound of the error in level \( k \) depends only on the angular variable \( \xi^k \) and not on the spatial \( x^k \). This is assured by assigning a global bound of the error at level zero. First, if \( \mathcal{R}_k^F (\xi_j) \) is a number such that:
\[
\left| \tilde{E}^k(x_t, \xi_j) - \mathcal{E}^k(\mu)(x_t, \xi_j) \right| \leq \mathcal{R}_k^E (\xi_j)
\]
for all \( x_t \) such that \( (x_t, \xi^k) \in \mathcal{I}_k \). And if \( \mathcal{R}_F^k \left( \xi_j \right) \) is a number such that:
\[
\left| \mathcal{F} \left( x_t, \frac{\xi_j^{k+1}}{2} \right) - \mathcal{E}^k(\mu) \left( x_t, \frac{\xi_j^{k+1}}{2} \right) \right| \leq \mathcal{R}_F^k \left( \frac{\xi_j}{2} \right)
\]
then \( \mathcal{R}_F^k \) can be chosen as:
\[
\mathcal{R}_F^k \left( \frac{\xi_j^{k+1}}{2} \right) := \frac{\mathcal{R}_E^k \left( \frac{\xi_j(0)}{2} \right) + \mathcal{R}_E^k \left( \frac{\xi_j(2)}{2} \right)}{2} + k^+ \mathcal{R}_F^I
\tag{71}
\]
This follows from triangular inequality (as was shown following definition B.I.1). We can now make the equivalence to claim B.1 and state that the error \( \mathcal{R}_F^k \left( \frac{\xi_j}{2} \right) \) derived for \( \mathcal{F} \left( x_t, \frac{\xi_j^{k+1}}{2} \right) \) applies also to \( \mathcal{F} \left( x_t + \frac{\xi_j^{k+1}}{2}, \frac{\xi_j^{k+1}}{2} \right) \). So far there is no difference from the analysis made in section B.I except that we now allow the
interpolation error to be depending of level. The errors shall now be propagated to $\tilde{E}^{k+1}$, by looking at formula (18). Since $\mathcal{F} = (\tilde{E} + \mathcal{R}_f)$ we get:

\[
\mathcal{R}_{\tilde{E}}^{k+1}(\xi_j) \leq \left( \mathcal{R}_f^k \left( \frac{\xi_j}{2} \right) \right)^2 + \mathcal{E}(\mu) \left( x_l + \frac{\xi_j}{2}, x_l + \frac{\xi_j}{2} \right) \mathcal{R}_f^k \left( \frac{\xi_j}{2} \right) \tag{72}
\]

Alternatively we can propagate the errors by the formula

\[
\mathcal{R}_f \approx \sum \left| \frac{\partial f}{\partial x_i} \right| \mathcal{R}_{x_i} \tag{73}
\]

(where $\mathcal{R}_f$ is the error of a computed function $f$ and $x_i$ are the variables upon which $f$ depends and which have errors $\mathcal{R}_{x_i}$) that is a first order approximation commonly used for error estimation, see i.e. [3]. Applied on (18) the formula has the form

\[
\mathcal{R}_{\tilde{E}} \approx \sum \left| \frac{\partial \tilde{E}}{\partial x_i} \right| \mathcal{R}_{x_i}
\]

which yields:

\[
\mathcal{R}_{\tilde{E}}^{k+1}(\xi_j) \leq \left( \mathcal{R}_f^k \left( \frac{\xi_j}{2} \right) \right)^2 + 2\mathcal{R}_f^k \left( \frac{\xi_j}{2} \right) \tag{74}
\]

By (71) and (72) or (74) the errors of $\tilde{E}$ can be propagated from level $k$ to $k+1$. This way errors can be tabulated or estimates can be computed numerically. We shall get back to applying these results in section C.V.

C.III Propagation of the error of $\tilde{I}^k_\mu$ to $\tilde{I}^{k+1}_\mu$

In this section the same analysis will be made as in C.II but for the error of $\tilde{I}_\mu$. This error however depends also on the error of $\mathcal{F}$. We shall first look at the error of the approximation $\mathcal{J}_\mu^k$ when defined by linear interpolation. The error of $\mathcal{J}_\mu^k$ shall then be propagated together with the error of $\mathcal{F}$ to $\tilde{I}^{k+1}_\mu$. This way recursively the error at level $k$ may be estimated. We still assume that a bound of the error in level $k$ depends only on the angular variable $\xi_j$ and not on the spatial $x_l^k$. This is assured by assigning a global bound of the error at level zero. If $\mathcal{R}_{\tilde{I}_\mu}^k(\xi_j)$ is a number such that

\[
\left| \tilde{I}^k_\mu(x_l, \xi_j) - \tilde{I}^k_\mu(\mu)(x_l, \xi_j) \right| \leq \mathcal{R}_{\tilde{I}_\mu}^k(\xi_j)
\]

The formula implies approximating a function with its tangent hyperplanes.
for all $x_l^k$ such that $(x_l^k, \xi_j^k) \in \mathcal{L}_k$. And if $R_{\mathcal{J}_\mu}^k \left( \frac{\xi_j}{2} \right)$ is a number such that:

$$
\left| J_\mu \left( x_l, \frac{\xi_j^{k+1}}{2} \right) - I_\mu^k (\mu) \left( x_l, \frac{\xi_j^{k+1}}{2} \right) \right| \leq R_{\mathcal{J}_\mu}^k \left( \frac{\xi_j}{2} \right)
$$

then

$$
R_{\mathcal{J}_\mu}^k \left( \frac{\xi_j^{k+1}}{2} \right) := \frac{R_{\mathcal{J}_\mu}^k \left( \frac{\xi_j^{(0)}}{2} \right) + R_{\mathcal{F}_\mu}^k \left( \frac{\xi_j^{(2)}}{2} \right)}{2} + k^{+1} R_{\mathcal{F}_\mu}^k
$$

(75)

This follows from triangular inequality (as was shown following definition B.I.1).

We can now make the equivalence to claim B.1 and state that the error $R_{\mathcal{J}_\mu}^k (\xi_j^2)$ derived for $J_\mu \left( x_l, \frac{\xi_j^{k+1}}{2} \right)$ applies also to $J_\mu \left( x_l + \frac{\xi_j^{k+1}}{2}, \frac{\xi_j^{k+1}}{2} \right)$. The errors of $J_\mu^k$ and $F^k$ are now propagated to $\tilde{\mathcal{E}}_{\mu}^k$ by expanding the terms of formula (19) when written as $F = (E + R_F^k)$ and $J_\mu = (I_\mu + R_{\mathcal{J}_\mu})$:

$$
R_{\mathcal{J}_\mu}^k (\xi_j) \leq \frac{1}{2} \left[ R_{\mathcal{J}_\mu}^k \left( \frac{\xi_j}{2} \right) \cdot \left( 1 + E^k (\mu) \left( x_l + \frac{\xi_j}{2}, \frac{\xi_j}{2} \right) \right) 
+ I_\mu^k (f) \left( x_l, \frac{\xi_j}{2} \right) R_{\mathcal{J}_\mu}^k \left( \frac{\xi_j}{2} \right) + R_{\mathcal{F}_\mu}^k \left( \frac{\xi_j}{2} \right) R_{\mathcal{F}_\mu}^k \left( \frac{\xi_j}{2} \right) \right]
$$

(76)

While using the approximative formula (73) we have:

$$
R_{\mathcal{J}_\mu}^k (\xi_j) \approx \frac{1}{2} \left[ R_{\mathcal{J}_\mu}^k \left( \frac{\xi_j}{2} \right) \cdot \left( 1 + E^k \left( x_l + \frac{\xi_j}{2}, \frac{\xi_j}{2} \right) \right) 
+ I_\mu^k (f) \left( x_l, \frac{\xi_j}{2} \right) R_{\mathcal{J}_\mu}^k \left( \frac{\xi_j}{2} \right) \right]
$$

(77)

By (75) and (76) or (77) the errors of $\tilde{\mathcal{E}}$ can be propagated from level $k$ to $k+1$. Note that the presence of $R_{\mathcal{F}_\mu}^k$ makes necessary to also use consult formula (71).

Application of these results will be discussed further in C.V

C.IV Improved accuracy by the second approach in computation of $\mathcal{F}$

Using the terms of (65) in formula (66) we see that the term $|\chi''(s)\mathcal{M}|$ may have great influence on the interpolation error of $\tilde{\mathcal{E}}$ particularly for short integral paths (since it is $\propto 2^{-k}$). The term is a result of interpolating the function $e^{-\chi(s)M(s)}$
in variable $s$, where $X(s) = \left| \frac{\xi_j(s)}{2} \right|$, see (65). This is not necessary however since $X(1)$ is known.

$$X(1) = \left| \frac{\xi_j}{2} \right|$$

This is the reason why the second approach of section 5.3 in the main text in an improvement. The interpolation error of the operator $M(\mu) \left( x_l \frac{\xi_j}{2} \right)$ when approximated by

$$\frac{M(\mu) \left( x_l \frac{\xi_j^{(0)}}{2} \right) + M(\mu) \left( x_l \frac{\xi_j^{(2)}}{2} \right)}{2}$$

is by theorem B.1 bounded by

$$|M'_s| \leq \frac{1}{12} m^2 h^2 \max \frac{\partial \mu^2}{\partial x^2}$$

(78)

Propagated to $F = e^{-XM}$ by the error propagation formula (73) the error of $F$ has the size

$$\frac{1}{2} e^{-XM} |X'| |M''|_{\text{max}}$$

(79)

to compare with inserting all the terms of (65) in formula (67). Basically this means that when interpolating $M$ in variable $s$ and propagating the error to $F$ instead of interpolating $F$ itself in $s$ (see equation (20)) all the terms of formula (65) can be replaced\(^8\) with the single term $e^{-XM}|M''X|$. The term $|M''X|$ is proportional to $2^k h^3$ and by remark C.I.1 one of the terms that has greater influence on the error when interpolating on longer paths. Thus we conclude that using this method is most worth the effort for shorter integral paths\(^9\). By the following remark, the method however may not give the great accuracy improvement in $I^k_\mu$ as the comparison of the errors (79) and (65) indicates.

**Remark C.IV.1.** The method described in section 5.3 is an improvement over the method in section 5.2 since we interpolate only the function $M(s)$ (24) in $s$. Unfortunately using this method does not alone eliminate the problem of interpolating $e^{-X(s)M(s)}$ in $s$, as long as $J^s_\mu$ is computed as described in section 5.2. This is since we then interpolate the function $\int_0^1 f(x, \xi(s)/2) e^{-X(s)M(s)} d\gamma$ (section C.I) in $s$. This can also be realized by looking at the terms of the error of equation (68).

\(^8\)Note that the interpolation error only constitutes part of the error of $F$, and that the way errors propagate from previous levels may also differ between the methods.

\(^9\)This can also be realized by noting that the relative difference in pathlength between neighboring paths is less for longer paths as a consequence of the increased angular resolution thus an accurate value of $X(1)$ is of greater gain for computation on shorter paths.
C.V Estimating total error at level $k$

We shall now investigate the error at level $k$ by means of the results of sections C.I through C.IV.

First to the error of $\hat{E}^k$. We shall first suppose that $\hat{E}^k$ and $\mathcal{F}^k$ are computed by the first algorithm (described in section 5.2). A loose estimate of the growth-rate with $k$ of the error can be made by the results of section C.II as follows. Now let $R^k_{\hat{E}}$ be an upper bound of the error of $\hat{E}^k$ and correspondingly $R^k_{\mathcal{F}}$ be an upper bound of the error of the approximation $\mathcal{F}^k$. By formula (71)

$$R^k_{\mathcal{F}} = C^k R^k_{\hat{E}} + k+1 R^I_{\hat{E}}$$

where $C^k \in [\frac{1}{2}, 1]$ and $k+1 R^I_{\hat{E}}$ is the interpolation error. And by (74):

$$R^{k+1}_{\hat{E}} \lesssim 2 R^k_{\mathcal{F}} = 2 \left( C^k R^k_{\hat{E}} + k+1 R^I_{\hat{E}} \right)$$

Since $R^{k+1}_{\hat{E}} = 2 C^k R^k_{\hat{E}}$ alone is a recursion for an exponential function of $k$ if $C > \frac{1}{2}$ we shall not expect the growth-rate with $k$ to be less than exponential. The alternative computation method of section C.IV provides better conditions for precisely estimating the error.

Error of $\hat{M}^k$ - generalization of (60)

Suppose that we compute $\hat{M}$ (equation (24)) as suggested in remark 5.3.1, as the weighted projection in $\mathbb{R}^n$ where $n$ may be larger than 2 and the attenuation is zero. Then we can estimate the error by generalizing the sequence of maximal error at level $k$ which led to the error estimates (60) and (63), for two dimensions, in section Appendix B. In $\mathbb{R}^n$ the interpolation error depends on in how many dimensions we interpolate. It has the factor $m^2$ from the second derivative w.r.t. $s$ (see equation (78)). By setting $m$ to its maximal value $m := n - 1$ (since one component is $\pm 2^k$) there is the formula

$$R^k_{\hat{M}} = \frac{2k+1}{72} (n-1)^2 h^2_{\text{max}} \left| \frac{\partial^2 \mu}{\partial x^2} \right|_{\text{max}} + R^0_{\hat{M}}$$

(80)

where $R^0_{\hat{M}}$ is the global error of $\hat{M}$ in level zero. Formula (80) is the generalization of formula (60) to $\mathbb{R}^n$.

Remark C.V.1. We have only proved that (80) bounds the maximal error for the case $n = 2$ (by proving claim B.2). However (80) bounds the error of $\hat{M}$ corresponding to the sequence of $\xi_j^k$ where in every level we choose the angular variable by interpolating in $(n-1)$ dimensions between the two previous choices.
APPENDIX C  ACCURACY ANALYSIS

Propagation of error to $\mathcal{F}^k$

The error of $\tilde{M}$ (80) of level $k$ shall now be propagated to

$$\mathcal{F}^k = e^{-k\tilde{M}^k}$$

(equation (25)). If $R_k^k$ is a number bounding the error of $\mathcal{F}^k$, then by the first order approximative formula (73)

$$R_k^k \approx \left| \frac{\partial \mathcal{F}^k}{\partial \tilde{M}^k} \right| R_k^{\tilde{M}^k} = \left| \frac{\xi_j^{k+1}}{2} \right| \exp \left\{ -\left| \frac{\xi_j^{k+1}}{2} \right| \tilde{M}^k \right\} R_k^{\tilde{M}^k}$$

Hence $R_k^k$ can be estimated:

$$R_k^k \approx \sqrt{n^2 k R_k^{\tilde{M}^k}}$$  (81)

where $R_k^{\tilde{M}^k}$ is as given in (80). The growth-rate with $k$ is $k^2$. Finally a closed form $k$-independent formula can be given by substituting the expression (62) for $k_L(h)$ into (81). The error of $\mathcal{F}$ computed on paths equal to or exceeding a certain length $L$ is approximately bounded by:

$$R_k^{\mathcal{F}} \approx \sqrt{n^2 k R_k^{\tilde{M}^k}} R_k^{kL(h)}$$  (82)

Apart from the term with the level zero error $R_0^{\tilde{M}^k}$, the error is $O(h \log \frac{1}{h})$ which is under control by shrinking $h$. It shall be emphasized that this error concerns the second approach for computing $\mathcal{F}$, given in section 5.3.

Propagation to $\tilde{\mathcal{I}}_\mu^k$

We shall propagate the error one step further to $\tilde{\mathcal{I}}_\mu^k$. Let $R_{\mathcal{J}_\mu}^k$ be a bound for the error of $\mathcal{J}_\mu^k$, and $R_{\tilde{\mathcal{I}}_\mu}^k$ bound the error of $\tilde{\mathcal{I}}_\mu^k$. Then by (75)

$$R_{\mathcal{J}_\mu}^k = C_k R_{\tilde{\mathcal{I}}_\mu}^k + kR_{\mathcal{I}_\mu}^L$$

where $C_k \in [\frac{1}{2}, 1]$ and $kR_{\mathcal{I}_\mu}^L$ is the interpolation error given by (70). Now, by first order approximation and by formula (77) there is the following recursive expression:

$$R_{\tilde{\mathcal{I}}_\mu}^{k+1} \approx \frac{f_{\max} R_{\mathcal{F}}^k}{2} + R_{\mathcal{J}_\mu}^k = \frac{f_{\max} R_{\mathcal{F}}^k}{2} + C_k R_{\tilde{\mathcal{I}}_\mu}^k + kR_{\mathcal{I}_\mu}^L$$  (83)

The terms of the error (83) are given as follows: $R_{\mathcal{F}}^k$ is given by (81). A closed form $k$-$k$-independent expression is given in (82) where $k$ has been substituted with $k_L(h)$ (62), $C_k \in [\frac{1}{2}, 1]$

The interpolation error $kR_{\mathcal{I}_\mu}^L$ is given by formula (70) and written in $k$-independent form it is proportional to $h^2$ (see remark C.I.1).

The approximation made is using formula (73). See also remark C.V.1
C.VI Conclusions

The important results of section Appendix C can be summed-up as follows (a more detailed summary of the results is given in section 5.4 in the main text):

The interpolation errors for $\mathcal{E}$ and $\mathcal{I}_\mu$ are investigated in C.I. The terms of the interpolation errors are given in formulae (65) and (68). These terms are level-dependent. For fixed pathlength $L$ all the terms of the errors are proportional to either one of

$$\frac{h^2}{L}, \quad h^2, \quad h^2 L, \quad h^2 L^2$$

by remark C.I.1. There is also a $m^2$-dependence of the terms, where $0 < m < n - 1$ is the number of dimensions over which to interpolate. The errors has been derived for univariate parameterization of $\xi_j$ (section 5.2.2). Multivariate parameterization (5.2.3) should yield same or better error-estimates.

In sections C.II and C.III we investigate how error estimates can be propagated from level $k$ to level $k + 1$. Formulae (71), (72), (75) and (76) describes this process depending on the angular variable and hence they are a generalization of the angular error-distribution of section B.II.

In section C.IV the alternative way of computing the approximation $\mathcal{F}$ (the second approach, section 5.3) is investigated as way of avoiding to interpolate for the pathlength $|\xi_j|$. All of the terms of the interpolation error of the previously described method (65) except one, will be eliminated using this method. This is under first order approximation. The one term remaining (79) is proportional to

$$h^2 L$$

and hence the method is particularly an improvement for short paths. There is by remark C.IV.1 a predicament concerning how big improvement of the algorithm that this method provides.

In section C.V we estimate the error at level $k$. The formula (60) bounding the error of the weighted projection with no attenuation, at level $k$, is generalized to $\mathbb{R}^n$ in formula (80). The 2-dimensional formula (60) is multiplied with the factor

$$(n - 1)^2$$

The formula has by remark C.V.1 only been proved for $n = 2$. This error, when propagated (by first order approximation) gives the error-estimate for $\mathcal{F}^k$ (when computed by the method of 5.3) which is given in the formula (81). A $k$-independent closed form of this error is given as formula (82). If the level-zero error is neglected, the remaining terms of the formula are

$$O \left( h \log \frac{1}{h} \right)$$
The error $R^k_F$ of $F$ is finally propagated to $\tilde{I}^k_\mu$ in the recursive formula (83). The error $R^{k+1}_{\tilde{I}_\mu}$ has three terms:

$$R^{k+1}_{\tilde{I}_\mu} = \frac{f_{\text{max}} R^k_F}{2} + C_k R^k_{\tilde{I}_\mu} + k R^l_{\tilde{I}_\mu}$$

Where $C \in [\frac{1}{2}, 1]$. By section C.I, the interpolation error $k R^l_{\tilde{I}_\mu}$ is, for fixed pathlength, $O(h^2)$. 