$H^2$-Matrix Method vs. FFT in Thin-Film

Stray-Field Computations

Jörg Drwenski, Felix Otto

no. 161

Diese Arbeit ist mit Unterstützung des von der Deutschen Forschungsgemeinschaft getragenen Sonderforschungsbereiches 611 an der Universität Bonn entstanden und als Manuskript vervielfältigt worden.

Bonn, August 2004
\textbf{H}^2\text{-Matrix Method vs. FFT in Thin-Film Stray-Field Computations}

Jörg Drwenski, Felix Otto\textsuperscript{*}

August 26, 2004

\textbf{Abstract}

We consider a magnetostatic problem arising in thin-film micromagnetics. For weak external field it is a variational formulation of the Dirichlet screen problem. Its solution amounts to the inversion of the single layer potential on a polygonal surface. The charge density on the surface has characteristic singularities near the edges and corners. We first find that establishing the regularity theory needed for the numerical analysis can be based on rather elementary arguments. A Galerkin ansatz with a piecewise constant charge distribution on a triangulation of the surface leads to a linear problem $Kz + b = 0$ with a fully populated matrix $K$. An iterative solver like the conjugate gradient method thus needs $O(N^2)$ operations per iteration when based on an explicit representation of $K$. We consider two different methods to reduce computational cost of the matrix vector multiplication: On a uniform triangulation we apply FFT and on a triangulation with a priori refinement we use the $\mathcal{H}^2$-matrix method developed by Hackbusch et al. For the latter case we derive the necessary refinement rate to retain the optimal rate of convergence in the energy norm. In particular, we show that it is sufficient to refine the triangulation towards the edges. We compare the CPU-time of the FFT-based conjugate gradient method with that of the $\mathcal{H}^2$-based method. We stop the conjugate gradient iteration once the relative error in the energy norm falls below a given tolerance. We find that, in terms of CPU time, the latter outperforms the first if an error tolerance in the energy norm of 5 percent or less is required.

Mathematics Subject Classification: 65R20, 65N38, 65N50, 65N15

1 Introduction & Summary

In Section 2, we briefly recall how the considered magnetostatic problem comes up in thin-film micromagnetics. We consider a thin ferromagnetic sample described by its cross-section $\Omega' \subset \mathbb{R}^2$ under an in-plane external field $H_{\text{ext}}$. In a certain regime, the in-plane magnetization $m'$ arranges itself in such a way that the stray field

\textsuperscript{*}Institute of Applied Mathematics, University of Bonn, Germany
$-\nabla u$ compensates $H'_\text{ext}$ in the sample. This can be reformulated as a minimization problem for the “magnetic charge density” $\sigma$, which is the in–plane divergence $-\nabla' \cdot m'$ of $m'$.

In Section 3, we point out that the problem for $\sigma$ is a variational formulation of the Dirichlet screen problem. More precisely, it can be seen as a boundary integral ansatz for the magnetostatic potential $u$ in $\mathbb{R}^3$ with prescribed Dirichlet data $H'_\text{ext} \cdot x'$ on $\Omega' \times \{0\}$. The potential $u$ is written as the convolution of $\sigma$ with the single layer potential on $\Omega' \times \{0\}$. The bilinear form in the variational formulation is the homogeneous part of the $H^{-1/2}(\mathbb{R}^2)$–norm of $\sigma$ (hereafter called “energy norm”, since it is the energy of the stray–field $-\nabla u$).

We think of $\Omega'$ as being polygonal; $\Omega' = (0,1)^2$ to fix ideas. The charge distribution $\sigma$ is known to have characteristic singularities near the edges and corners of $\Omega'$. In Section 4, we establish the minimal regularity theory we later require. We only use elementary arguments and express the regularity by pointwise estimates in Theorem 1.

In Section 5, we introduce the Galerkin ansatz. It is based on a uniform triangulation $T_h$ of $\Omega'$. The ansatz functions for $\sigma$ are piecewise constant. If $\sigma$ did not have singularities, the error $\epsilon$ in the Galerkin method would scale like $h^{3/2}$, where $h$ denotes the maximal diameter of the triangles. Here and in the sequel, the discretization error $\epsilon$ is measured in terms of the energy norm, which is the natural norm in this context of micromagnetics. Because of the singularities of $\sigma$, the order $3/2$ can only be preserved under a particular, but a priori known refinement of $T_h$ towards the edges. We characterize this refinement of $T_h$ by an exponent $\alpha$. In Theorem 2, we give an error estimate in terms of the diameter $h$ of the triangles in the bulk. Naturally, it depends on $\alpha$ and yields the optimal scaling $h^{3/2}$ provided $\alpha > 2/3$. This estimate of the discretization error $\epsilon$ in the energy norm reduces to an estimate of the approximation error and is based on the regularity theory from Section 4. The corner singularities do not affect the scaling.

In Section 6, we explain how we numerically generate an a priori refined triangulation. The two parameters are the triangle diameter $h$ in the bulk and the refinement exponent $\alpha$. We can ensure numerical stability of the triangulations if we work with regular triangles in the sense that no small angles occur. We then argue how the number of triangles $N$ scales in $h$, depending on $\alpha$. In consequence, we find that the best relation between discretization error $\epsilon$ and $N$ is given for $\alpha$ in the range $[1/2, 2/3]$. This finding is based on the error estimates from Section 5.

In Section 7, we investigate how to numerically solve the discrete linear system derived on such triangulations. It amounts to solving a linear problem of the form $Kz + b = 0$. The non–sparse matrix $K$ comes from the convolution operator with the single layer potential. Since $K$ turns out to be symmetric positive definite, we can employ the conjugate gradient method for the solution of $Kz + b = 0$. In every step of the conjugate gradient method, a matrix vector multiplication with $K$ has to be carried out. For a uniform triangulation, the operator $K$ preserves the convolution
structure. Hence the matrix vector multiplication can be efficiently carried out with help of the Fast Fourier Transform (FFT). We recall complexity results of the FFT, which is known to be of order $O(N \log N)$ in the number $N$ of triangles.

In Section 8, we give a brief introduction into the method of $\mathcal{H}^2$–matrices [17]. It employs an approximation of $K$ based on a hierarchical organization of the data structure. We restrict ourselves to our special setup. We discuss its complexity, which roughly speaking is $O(N)$.

Finally, in Section 9, we compare the performance of the FFT–method on a uniform triangulation (as parametrized by the diameter $h$ of the identical triangles) with that of the $\mathcal{H}^2$–matrices for the optimally refined triangulation ( i. e. $\alpha = 2/3$ and as parametrized by the diameter $h$ of the triangles in the bulk). We compare the methods in terms of the discretization error $\epsilon > 0$ in the energy norm, which we use as a stopping criterion in the conjugate gradient method. We measure the CPU–time, excluding the setup time (which is substantial for $\mathcal{H}^2$–matrices). We find that the $\mathcal{H}^2$–matrices beat FFT if one imposes an error tolerance of 5 percent or less.

## 2 A Reduced Model for Thin–Film Micromagnetics

In [7], a reduced 2–d model for a thin–film ferromagnetic element under external field was derived. The thin ferromagnetic sample is described by its cross section $\Omega' \subset \mathbb{R}^2$. Here and in the following pages the prime denotes the projection on the in–plane components. In the reduced 2–d model, the magnetization $m$ is in–plane, i. e. $m = m'$, and constant in the thickness direction, i. e. $m' = m'(x')$ for $x' \in \Omega'$. Hence $m': \Omega' \to \mathbb{R}^2$ is a 2–d vector field. Appropriately non–dimensionalized, the reduced energy takes the form

$$E(m') = \int_{\mathbb{R}^3} |\nabla u|^2 dx - 2 \int_{\Omega'} m' \cdot H'_{ext} dx'. \quad (2.1)$$

Here $H'_{ext}$ denotes the in–plane external field and $-\nabla u$ is the stray–field. The stray field potential $u: \mathbb{R}^3 \to \mathbb{R}$ is generated by the “magnetic charges”: $\sigma = -\nabla' \cdot m'$ on $\Omega'$ and $\nu' \cdot m'$ along $\partial \Omega'$. This is best formulated distributionally:

$$\int_{\mathbb{R}^3} \nabla u \cdot \nabla \zeta dx = \int_{\Omega'} m'(x') \cdot \nabla' \zeta(x',0) dx' \quad \text{for all test functions } \zeta: \mathbb{R}^3 \to \mathbb{R}.$$

As the exchange energy is neglected, the saturation constraint $|m'|^2 = 1$ relaxes to its convexification

$$|m'|^2 \leq 1. \quad (2.2)$$

This reduced model generalizes an ad hoc model proposed in [4]. Notice that the reduced model embodies the competition between stray–field energy and the effect of the external field. The Zeeman term $\int_{\Omega'} m' \cdot H'_{ext} dx'$ favors alignment, the stray
field favors pole avoidance. In particular, a finite stray-field energy imposes that $m'$ is tangential to $\partial \Omega'$:

$$\nu' \cdot m' = 0 \quad \text{on } \partial \Omega'.$$  

(2.3)

We notice that $E$ depends on $m'$ only via $\sigma$. Indeed, the stray-field energy can be easily expressed in terms of the Fourier transform $\hat{\sigma}$ of $\sigma$ (which for this purpose we extend trivially on $\mathbb{R}^2$):

$$\int_{\mathbb{R}^3} |\nabla u|^2 \, dx = \frac{1}{2} \int_{\mathbb{R}^2} |k'|^{-1} |\hat{\sigma}(k')|^2 \, dk'.$$

We interpret this expression as (the homogeneous part of) the $H^{-1/2}(\mathbb{R}^2)$–norm of $\sigma$:

$$\int_{\mathbb{R}^2} |k'|^{-1} |\hat{\sigma}(k')|^2 \, dk' =: \int_{\mathbb{R}^2} \left| (\nabla')^{1/2} \sigma \right|^2 \, dx'. 

(2.4)

Furthermore, (2.3) implies that the second term in (2.1) can be reformulated as

$$-2 \int_{\Omega'} \sigma H'_\text{ext} \cdot x' \, dx'.$$

Hence the reduced energy can be rewritten as

$$E(\sigma) = \frac{1}{2} \int_{\mathbb{R}^2} \left| (\nabla')^{-1/2} \sigma \right|^2 \, dx' - 2 \int_{\Omega'} \sigma H'_\text{ext} \cdot x' \, dx'. 

(2.5)

The Euler–Lagrange equation of the reduced problem is given by

$$\begin{align*}
-\nabla' u + H'_\text{ext} &= \lambda m \\
\lambda(|m| - 1) &= 0
\end{align*}$$

in $\Omega'$

where $\lambda(x) \geq 0$ is the Lagrange multiplier related to (2.2). For sufficiently weak external field $H'_\text{ext}$, the constraint (2.2) is not active leading to $\lambda \equiv 0$. Hence in this case, the magnetization adjusts itself such that the stray field $-\nabla' u$ compensates the external field $H'_\text{ext}$ in the samples cross section $\Omega'$.

First numerical simulations of the reduced model and its comparison to subsequently carried out experiments have been published in [8]. Simulations and experimental observations are in good agreement in the regime of sufficiently thick films where the wall type is the asymmetric Bloch wall. In thinner films, the repulsion of the Néel wall from the sample edge leads to some deviations with respect to wall expulsion.

3 The Dirichlet Screen Problem

For sufficiently weak external field $H'_\text{ext}$, the constraint (2.2) is not active. In view of (2.5), the reduced problem turns into the quadratic variational problem

Minimize

$$E(\sigma) = \frac{1}{2} \int_{\mathbb{R}^2} \left| (\nabla')^{-1/2} \sigma \right|^2 \, dx' - 2 \int_{\Omega'} H'_\text{ext} \cdot x' \sigma \, dx'$$

among all $\sigma$ with $\int_{\Omega'} \sigma \, dx' = 0$. 

(3.1)
We recall that the formal expression \( \int_{\mathbb{R}^2} |(\nabla')^{-1/2} \sigma|^2 \, dx' \) has many equivalent definitions:

- It can be expressed in terms of the Fourier transform \( \hat{\sigma}(k') \), see (2.4):
  \[
  \int_{\mathbb{R}^2} |(\nabla')^{-1/2} \sigma|^2 \, dx' = \int_{\mathbb{R}^2} |k'|^{-1} |\hat{\sigma}(k')|^2 \, dk'.
  \]

- It can be written as a two–dimensional convolution integral with the single layer potential as weakly singular kernel
  \[
  \frac{1}{2} \int_{\mathbb{R}^2} |(\nabla')^{-1/2} \sigma|^2 \, dx' = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \sigma(x') \frac{1}{4 \pi |x' - y'|} \sigma(y') \, dy' \, dx'.
  \tag{3.2}
  \]

- It can be expressed as a three–dimensional Dirichlet integral
  \[
  \frac{1}{2} \int_{\mathbb{R}^2} |(\nabla')^{-1/2} \sigma|^2 \, dx' = \int_{\mathbb{R}^3} |\nabla u|^2 \, dx,
  \]
  where the potential \( u \) is determined by
  \[
  -\Delta u = 0 \quad \text{in } \mathbb{R}^3 - (\Omega' \times \{0\}),
  \]
  \[
  \partial_3 u(x', 0+) - \partial_3 u(x', 0-) = \sigma(x') \quad \text{for } x' \in \Omega'.
  \tag{3.5}
  \]

- It can be written as a dual norm
  \[
  \left( \frac{1}{2} \int_{\mathbb{R}^2} |(\nabla')^{-1/2} \sigma|^2 \, dx' \right)^{1/2} = \sup \left\{ \frac{\int_{\mathbb{R}^2} \sigma u \, dx'}{(\int_{\mathbb{R}^3} |\nabla u|^2 \, dx)^{1/2}} \mid u : \mathbb{R}^3 \to \mathbb{R} \right\}, \tag{3.3}
  \]
  resp.
  \[
  \frac{1}{2} \int_{\mathbb{R}^2} |(\nabla')^{-1/2} \sigma|^2 \, dx' = \sup \left\{ -\int_{\mathbb{R}^3} |\nabla u|^2 \, dx - 2 \int_{\mathbb{R}^2} \sigma u \, dx' \mid u : \mathbb{R}^3 \to \mathbb{R} \right\}. \tag{3.4}
  \]

From (3.4) we see that (3.1) can be understood as a saddle point problem in \((\sigma, u)\) with solution \((\sigma^*, u^*)\). The first variation in \( u \) yields

\[
\Delta u^* = 0 \quad \text{in } \mathbb{R}^3 - (\Omega' \times \{0\}),
\]

\[
\partial_3 u^*(x', 0+) - \partial_3 u^*(x', 0-) = \sigma^*(x') \quad \text{for } x' \in \Omega'.
\tag{3.5}
\]

The first variation in \( \sigma \) yields, up to additive constants,

\[
u^*(x', 0) = H'_{\text{ext}} \cdot x' \quad \text{for } x' \in \Omega'. \tag{3.6}
\]

Hence solving the variational problem (3.1) amounts to evaluating the Dirichlet–Neumann–operator of the bounded surface \( \Omega' \). For constant fields \( H'_{\text{ext}} \), this representation shows that it is essentially the geometry of \( \Omega' \) which determines \( \sigma^* \).
Problem (3.1) is mathematically equivalent to the *Dirichlet screen problem*. This problem is usually encountered in an electrostatic instead of a magnetostatic context: Given the surface potential \( u \) find the charge density \( \sigma \) of a thin–film conductor described by \( \Omega' \). Mathematically, this amounts to the solution of a Dirichlet problem in \( \mathbb{R}^3 - (\Omega' \times \{0\}) \) with the help of a single–layer potential on \( \Omega' \times \{0\} \). In our reduced thin–film model, the prescribed potential is related to the external field via \( u(x') = H'_{\text{ext}} \cdot x' \).

### 4 Regularity Theory

It is well–known and easy to understand that \( \sigma \) diverges like \( r^{-1/2} \) where \( r \) denotes the distance to a smooth edge of \( \Omega' \). Corner singularities are at least numerically well–characterized: In [20], it is shown that \( \sigma \) diverges like \( R^{-0.7034} \) where \( R \) denotes the distance to a square corner, like the corner \( (-\frac{1}{2}, -\frac{1}{2}) \) of \( \Omega' = (-\frac{1}{2}, \frac{1}{2})^2 \). For our purpose the special form of the corner singularity is irrelevant. Roughly speaking, we only need the bound \( R^{-1} \). On the other hand, it is important to know how the \( r^{-1/2} \)–behavior near the middle of an edge degenerates as one approaches a corner.

We give here an elementary and self–contained proof of the statement our numerical analysis requires. In particular, we do not attempt to characterize the corner singularities. Our analysis is elementary in the sense that it is based exclusively on the maximum principle. For simplicity, we formulate the regularity result for \( \Omega' = (-\frac{1}{2}, \frac{1}{2})^2 \). We use the following language:

\[
\begin{align*}
\text{edges} & = \partial \Omega', \\
\text{corners} & = \{ \left( \frac{1}{2}, \frac{1}{2} \right), \left( \frac{1}{2}, -\frac{1}{2} \right), \left( -\frac{1}{2}, \frac{1}{2} \right), \left( -\frac{1}{2}, -\frac{1}{2} \right) \}.
\end{align*}
\]

In the sequel, we use the notation "\( \lesssim \)" to denote "\( \leq C \)" with a generic universal constant \( C \).

**Theorem 1.** The solution \( \sigma^* \) of (3.1) satisfies

\[
\begin{align*}
|\sigma^*(x')| & \lesssim |H'_{\text{ext}}| \text{dist}(x', \text{corners})^{-1/2} \text{dist}(x', \text{edges})^{-1/2}, \\
|\nabla \sigma^*(x')| & \lesssim |H'_{\text{ext}}| \text{dist}(x', \text{corners})^{-1/2} \text{dist}(x', \text{edges})^{-3/2}.
\end{align*}
\]

Theorem 1 uses the saddle point characterization (3.5) & (3.6) of (3.1). This allows to reduce the theorem to the following proposition:

**Proposition 1.** Let \( u \) be harmonic in \( \mathbb{R}^3 - (\Omega' \times \{0\}) \) and vanishing on \( \Omega' \times \{0\} \). Then for all \( x \in B_1 \)

\[
\begin{align*}
|u(x)| & \lesssim \left( \sup_{B_2} |u| \right) \text{dist}(x, \text{corners})^{-1/2} \text{dist}(x, \text{edges})^{1/2}, \\
|\nabla u(x)| & \lesssim \left( \sup_{B_2} |u| \right) \text{dist}(x, \text{corners})^{-1/2} \text{dist}(x, \text{edges})^{-1/2}, \\
|\nabla^2 u(x)| & \lesssim \left( \sup_{B_2} |u| \right) \text{dist}(x, \text{corners})^{-1/2} \text{dist}(x, \text{edges})^{-3/2}.
\end{align*}
\]

6
Proposition 1 in turn relies on

**Lemma 1.** Let \( u \) be harmonic in \( B_1 - \{x_3 = 0, x_2 \geq 0\} \) and vanishing on \( \{x_3 = 0, x_2 \geq 0\} \). Then we have

\[
|u(0, x_2, x_3)| \lesssim \left( \sqrt{x_2^2 + x_3^2} \right)^{1/2} \sup_{B_1} |u|.
\]

Lemma 1 is derived from the following two lemmas.

**Lemma 2.** Let \( x \in \mathbb{R}^3 \) be arbitrary and \( r > 0 \). Let \( u \) be harmonic in \( B_1(x) - \{x_3 = 0\} \) and vanishing on \( \{x_3 = 0\} \). Then we have

\[
r^2 \sup_{B_{r/2}(x)} |\nabla^2 u| + r \sup_{B_{r/2}(x)} |\nabla u| \lesssim \sup_{B_r(x)} |u|.
\]

**Lemma 3.** Let \( u \) be harmonic in \( B_1 - \{x_3 = 0, x_2 \geq 0\} \) and vanishing on \( \{x_3 = 0, x_2 \geq 0\} \). Then

\[
\sup_{B_{1/2}} |\partial_1 u| \lesssim \sup_{B_1} |u|.
\]

**Proof of Theorem 1.**

Let \( u^* \) be the solution of (3.5) & (3.6). The maximum principle on \( \mathbb{R}^3 - (\Omega' \times \{0\}) \) yields

\[
\sup_{\mathbb{R}^3} |u^*| \leq \sup_{x' \in \Omega'} |H_{\text{ext}}' \cdot x'| \lesssim |H_{\text{ext}}'|. \tag{4.4}
\]

We apply Proposition 1 to

\[
u(x) := u^*(x) - H_{\text{ext}}' \cdot x',
\]

which is indeed harmonic in \( \mathbb{R}^3 - (\Omega' \times \{0\}) \) and vanishes on \( \Omega' \times \{0\} \). In view of (4.4) we have

\[
\sup_{B_2} |u| \lesssim |H_{\text{ext}}'|.
\]

We infer from (3.5):

\[
\sigma^*(x') \equiv \partial_3 u^*(x', 0+ - \partial_3 u^*(x', 0-)) = \partial_3 u(x', 0+) - \partial_3 u(x', 0-),
\]

\[
\nabla' \sigma(x') = \nabla' \partial_3 u^* (x', 0+) - \nabla' \partial_3 u^* (x', 0-)) = \nabla' \partial_3 u (x', 0+) - \nabla' \partial_3 u (x', 0-).
\]

Hence the estimates of the theorem follow from those of the proposition. \( \square \)

**Proof of Proposition 1.**

We first address (4.1). Let \( x \in B_1 \) be arbitrary and set

\[
r := \text{dist}(x, \text{edges}) \leq \text{dist}(x, \text{corners}) =: R \leq 1.
\]

It suffices to consider points \( x \) with

\[
2r \leq R. \tag{4.5}
\]
Let
\[ y' \in \text{edges} \]
be such that
\[ r = |x - y'|. \tag{4.6} \]
Since
\[ \text{dist}(y', \text{corners}) \geq \text{dist}(x, \text{corners}) - |x - y'| \geq \frac{1}{2} R, \tag{4.5} \]
we have \( B_{R/2}(y') \cap \text{corners} = \emptyset \). Hence we may apply Lemma 1 (rescaled and translated) to the ball \( B_{R/2}(y') \). Since \( x \in B_r(y') \subset B_{R/2}(y') \), it yields
\[ |u(x)| \lesssim \sup_{B_{R/2}(y')} |u| \left( \frac{|x - y'|}{\frac{1}{2} R} \right)^{1/2} \]
\[ \quad \overset{(4.6)}{=} \left( \sup_{B_2} |u| \right) \left( \frac{\text{dist}(x, \text{edges})}{\text{dist}(x, \text{corners})} \right)^{1/2}. \]
This establishes (4.1).

We now argue that for any \( x \in B_1 \)
\[ |\nabla u(x)| \lesssim \left( \sup_{B_2} |u| \right) \text{dist}(x, \text{edges})^{-1}, \tag{4.7} \]
\[ |\nabla^2 u(x)| \lesssim \left( \sup_{B_2} |u| \right) \text{dist}(x, \text{edges})^{-2}. \tag{4.8} \]
Indeed, set
\[ r := \text{dist}(x, \text{edges}) \]
and notice that since \( B_r(x) \cap \text{edges} = \emptyset \) we must have
\[ u \text{ is harmonic in } B_r(x) - \{x_3 = 0\} \]
and vanishes on \( \{x_3 = 0\} \)
or
\[ u \text{ is harmonic in } B_r(x). \]
We now may apply Lemma 2 which yields
\[ r^2|\nabla^2 u(x)| + r|\nabla u(x)| \lesssim \sup_{B_r(x)} |u|. \]
Because of \( B_r(x) \subset B_2 \), this establishes (4.7) & (4.8).

We now derive (4.2) & (4.3). Let \( x \in B_1 \) be arbitrary and set
\[ r := \text{dist}(x, \text{edges}) \leq \text{dist}(x, \text{corners}) =: R \leq 1. \]
In view of (4.7) & (4.8), it again suffices to consider points \( x \) for which
\[
2r \leq R .
\]
(4.9)

We notice that for \( y \in B_r(x) \)
\[
\text{dist}(y, \text{edges}) \leq \text{dist}(x, \text{edges}) + |x - y| \leq 2r,
\]
dist(y, corners) \( \geq \text{dist}(x, \text{corners}) - |x - y| \geq \frac{1}{2} R, \quad \text{(4.9)} \)
so that (4.1) implies
\[
\sup_{B_r(x)} |u| \lesssim \left( \sup_{B_2} |u| \right) \left( \frac{r}{R} \right)^{1/2} .
\]
(4.10)

On the other hand, we have \( B_r(x) \cap \text{edges} = \emptyset \) by definition of \( r \), so that
\[
\text{u is harmonic in } B_r(x) - \{ x_3 = 0 \}
\]
and vanishes on \( \{ x_3 = 0 \} \)
or
\[
\text{u is harmonic in } B_r(x).
\]
Hence Lemma 2 yields
\[
r^2 |\nabla^2 u(x)| + r |\nabla u(x)| \lesssim \sup_{B_r(x)} |u|. \quad \text{(4.11)}
\]

The combination of (4.10) with (4.11) yields (4.2) with (4.3).

\[\square\]

**Proof of Lemma 2.**

Since \( u \) vanishes on \( \{ x_3 = 0 \} \), the odd reflections
\[
u_+(x_1, x_2, x_3) = \begin{cases} \begin{align*} u(x_1, x_2, x_3) & \text{ if } x_3 \geq 0 \\ -u(x_1, x_2, -x_3) & \text{ if } x_3 \leq 0 \end{align*} \end{cases},
\]
\[
u_-(x_1, x_2, x_3) = \begin{cases} \begin{align*} -u(x_1, x_2, -x_3) & \text{ if } x_3 \geq 0 \\ u(x_1, x_2, x_3) & \text{ if } x_3 \leq 0 \end{align*} \end{cases}
\]
are both harmonic in \( B_r(x) \cup B_r(-x) \). Now
\[
r^2 \sup_{B_{r/2}(\pm x)} |\nabla^2 u_{\pm}| + r \sup_{B_{r/2}(\pm x)} |\nabla u_{\pm}| \lesssim \sup_{B_r(\pm x)} |u_{\pm}| \leq \sup_{B_r} |u|
\]
is a standard estimate.

\[\square\]

**Proof of Lemma 3.**

We use that
\[
\partial_1 u \quad \text{is harmonic in} \quad B_1 - \{ x_3 = 0, x_2 \geq 0 \}, \quad \text{(4.12)}
\]
\[
\partial_1 u \quad \text{vanishes on} \quad \{ x_3 = 0, x_2 \geq 0 \}. \quad \text{(4.13)}
\]
We apply Bernstein’s argument. Fix a cut-off function \( \eta \in C_\infty^0(B_1) \) with \( \eta = 1 \) on \( B_{1/2} \) and consider for a constant \( \lambda \)

\[
w := \eta^2(\partial_1 u)^2 - \lambda u^2.
\]

By construction of \( \eta \) and by (4.13),

\[
w = 0 \quad \text{on} \quad \partial(B_1 - \{x_3 = 0, x_2 \geq 0\}). \tag{4.14}
\]

We claim that for \( \lambda = \lambda(\eta) \) sufficiently large

\[
-\Delta w \leq 0. \tag{4.15}
\]

Indeed

\[
-\Delta w = -\eta^2 \Delta (\partial_1 u)^2 - 4\eta \nabla \eta \cdot \nabla (\partial_1 u)^2 - \Delta \eta^2 (\partial_1 u)^2 - \lambda \Delta u^2 \tag{4.12}
\]

\[
\leq -2\eta^2 |\nabla \partial_1 u|^2 - 8\eta \partial_1 u \nabla \eta \cdot \nabla \partial_1 u - \Delta \eta^2 (\partial_1 u)^2 - 2\lambda |\nabla u|^2
\]

\[
\leq 8(\partial_1 u)^2 |\nabla \eta|^2 - \Delta \eta^2 (\partial_1 u)^2 - 2\lambda |\nabla u|^2
\]

so that (4.15) holds provided \( \lambda \geq 4|\nabla \eta|^2 + \frac{1}{2}|\Delta \eta|^2 \). By the maximum principle we obtain from (4.15) and (4.14)

\[
w \leq 0 \quad \text{in} \quad B_1,
\]

that is

\[
\eta^2 (\partial_1 u)^2 \leq \lambda u^2 \lesssim (\sup_{B_1} |u|)^2 \quad \text{in} \quad B_1.
\]

By construction of \( \eta \) this yields as desired

\[
(\partial_1 u)^2 \lesssim (\sup_{B_1} |u|)^2 \quad \text{in} \quad B_{1/2}.
\]

\[
\square
\]

**Proof of Lemma 1.**

We introduce the abbreviations

\[
u_0(x_2, x_3) := u(0, x_2, x_3) \quad \text{and} \quad r := \sqrt{x_2^2 + x_3^2}.
\]

According to Lemma 2 applied to balls with centers \( x_2 \geq \frac{1}{8} \) and radius \( = \frac{1}{16} \), we have

\[
\sup_{B_{1/4} \cap \{x_2 \geq \frac{1}{8}\}} |\nabla u| \lesssim \sup_{B_1} |u|.
\]

Since \( u \) vanishes on \( B_{1/2} \cap \{x_2 \geq \frac{1}{8}, x_3 = 0\} \), this implies

\[
u_0 \leq (\sup_{B_1} |u|)|x_3| \quad \text{for} \quad r \leq \frac{1}{4}, \quad x_2 \geq \frac{1}{8}. \tag{4.16}
\]
According to Lemma 3, applied twice, we have

\[ \sup_{B_{1/4}} |\partial_2^2 u| \lesssim \sup_{B_1} |u|. \]

This implies

\[ -(\partial_2^2 + \partial_3^2)u_0 \lesssim \sup_{B_1} |u| \quad \text{for } r \leq \frac{1}{4}, \]

(4.17)

We now compare \( u_0 \) to

\[ w(x_2, x_3) := (r - x_2)^{1/2} - (r - x_2) \quad \text{for } r \leq \frac{1}{4}. \]

An easy calculation shows

\[ -(\partial_2^2 + \partial_3^2)w = (\partial_2^2 + \partial_3^2)r = \frac{1}{r} > 1 \quad \text{for } r \leq \frac{1}{4}. \]

(4.18)

Notice that \( 0 \leq r - x_2 \leq 2r \leq \frac{1}{2} \) so that \( w \sim (r - x_2)^{1/2} \) for \( r \leq \frac{1}{4} \). Hence we have

\[ w = 0 \quad \text{for } x_2 \geq 0, \ x_3 = 0, \]

(4.19)

\[ w \gtrsim 1 \quad \text{for } x_2 \leq \frac{1}{8}, \ r = \frac{1}{4}, \]

(4.20)

\[ w \gtrsim |x_3| \quad \text{for } x_2 \geq \frac{1}{8}, \ r = \frac{1}{4}. \]

(4.21)

It follows from (4.17) & (4.18)

\[ -(\partial_2^2 + \partial_3^2)u_0 \lesssim (\sup_{B_1} |u|)(-(\partial_2^2 + \partial_3^2)w) \quad \text{for } r \leq \frac{1}{4} \]

and from (4.16) & (4.21)

\[ u_0 \lesssim (\sup_{B_1} |u|)w \quad \text{for } x_2 \geq \frac{1}{8}, \ r = \frac{1}{4}. \]

Moreover, (4.20) implies

\[ u_0 \leq \sup_{B_1} |u| \lesssim (\sup_{B_1} |u|)w \quad \text{for } x_2 \leq \frac{1}{8}, \ r = \frac{1}{4}. \]

Finally, we have thanks to (4.19)

\[ u_0 = 0 = (\sup_{B_1} |u|)w \quad \text{for } x_2 \geq 0, \ x_3 = 0. \]

Therefore, the maximum principle on \( \{r < \frac{1}{4}\} - \{x_2 \geq 0, x_3 = 0\} \) yields

\[ u_0 \lesssim (\sup_{B_1} |u|)w \leq (2r)^{1/2} \sup_{B_1} |u| \quad \text{for } r \leq \frac{1}{4}. \]

Likewise, we obtain

\[ -u_0 \lesssim r^{1/2} \sup_{B_1} |u| \quad \text{for } r \leq \frac{1}{4}, \]

so that

\[ |u(0, x_2, x_3)| \lesssim r^{1/2} \sup_{B_1} |u| \quad \text{for } r \leq \frac{1}{4}. \]

Since the estimate is trivial for \( r \geq \frac{1}{4} \), the lemma follows. \( \square \)
5 Galerkin Ansatz and Error Estimate

The singularities of $\sigma^*$ near edges and corners raises the question of how to choose a local refinement strategy for a triangulation of $\Omega'$. In the case of graded meshes and piecewise constant functions it is shown in [10, 23] that it suffices to refine the mesh only towards the edges to retain optimal convergence. We come to the same conclusion for our type of triangulations.

Consider a regular triangulation $\mathcal{T}_h$ of $\Omega'$ with typical triangle diameter $h$. A divergence–conforming finite element ansatz for the magnetization $m'$ with Raviart–Thomas elements [22] leads on the level of $\sigma$ to the space $X_h$ of all piecewise constant $\sigma_h$'s with vanishing mean $\int_{\Omega'} \sigma_h \, dx' = 0$ (which we extend trivially on $\mathbb{R}^2$). Hence we are investigating the following Galerkin ansatz for (3.1):

Minimize

$$E(\sigma_h) = \frac{1}{2} \int_{\mathbb{R}^2} |(\nabla')^{-1/2} \sigma_h|^2 \, dx' + 2 \int_{\Omega'} H'_\text{ext} \cdot x' \sigma_h \, dx'$$

among all $\sigma_h \in X_h$. (5.1)

We now formulate our estimate of the discretization error.

**Theorem 2.** Let $\Omega' = (-\frac{1}{2}, \frac{1}{2})^2$ and $0 < h, \alpha < 1$. Let the triangulation $\mathcal{T}_h$ of $\Omega'$ satisfy

$$\forall T \in \mathcal{T}_h \quad \text{diam}(T) \leq h \text{dist}(x'_T, \partial\Omega')^\alpha,$$

where $x'_T$ denotes the barycenter of $T$. Then there exists a constant $C$, which only depends on $\alpha$, such that we have for the solutions $\sigma^*$ and $\sigma^*_h$ of (3.1) resp. (5.1)

$$\left( \int_{\mathbb{R}^2} |(\nabla')^{-1/2}(\sigma^*_h - \sigma^*)|^2 \, dx' \right)^{1/2} \leq C |H'_\text{ext}| \left\{ \begin{array}{ll} h^{3/2} & \text{for } \alpha > \frac{2}{3} \\ \left(\ln \frac{1}{h} \right) h^{2(1-\alpha)} & \text{for } \alpha < \frac{2}{3} \end{array} \right..$$

(5.3)

**Proof of Theorem 2.**
We denote by $C$ a generic constant which only depends on $\alpha$. In view of (3.1) and (5.1), the discretization error in the energy norm is given by the approximation error. It thus suffices to show

$$\left( \int_{\mathbb{R}^2} |(\nabla')^{-1/2}(I\sigma^* - \sigma^*)|^2 \, dx' \right)^{1/2} \leq C |H'_\text{ext}| \left\{ \begin{array}{ll} h^{3/2} & \text{for } \alpha > \frac{2}{3} \\ \left(\ln \frac{1}{h} \right) h^{2(1-\alpha)} & \text{for } \alpha < \frac{2}{3} \end{array} \right.,$$

(5.4)

where the interpolation $I\sigma \in X_h$ is defined via

$$\forall T \in \mathcal{T}_h \quad \int_T I\sigma \, dx' = \int_T \sigma \, dx'.$$

In view of (3.3), i.e.

$$\left( \frac{1}{2} \int_{\mathbb{R}^2} |(\nabla')^{-1/2}(I\sigma^* - \sigma^*)|^2 \, dx' \right)^{1/2} = \sup \left\{ \frac{\int_{\Omega'} (I\sigma^* - \sigma^*) \, u \, dx'}{\left( \int_{\mathbb{R}^3} |\nabla u|^2 \, dx' \right)^{1/2}} \mid u : \mathbb{R}^3 \to \mathbb{R} \right\},$$

12
(5.4) amounts to show
\[
\sum_{T \in \mathcal{T}} \int_T (I \sigma^* - \sigma^*) (I u - u) \, dx' = \left| \int_{\Omega'} (I \sigma^* - \sigma^*) (I u - u) \, dx \right| \\
\leq C |H'_{ext}| \left( \int_{\mathbb{R}^3} |\nabla u|^2 \, dx \right)^{1/2} \left\{ \frac{h^2}{(\ln \frac{1}{h})^{1/2}} \right\} \tag{5.5}
\]
for an arbitrary test function \( u \).

We will need the following trace estimate for \( u \) on each triangle \( T \) and for exponents \( p \in \{2, 4\} \)
\[
\left( \int_T |u - Iu|^p \, dx' \right)^{1/p} \leq C \frac{\text{diam}(T)^{1/2}}{\text{area}(T)^{1/2 - 1/p}} \left( \int_{T \times \mathbb{R}} |\nabla u|^2 \, dx \right)^{1/2}. \tag{5.6}
\]
Indeed, let \( \hat{T} \) be the reference triangle on which we are sure to have the estimate, which is a mixture of a Sobolev estimate, a trace estimate (recall that \( H^1(\mathbb{R}^3) \) embeds into \( L^4(\mathbb{R}^2) \)) and a Poincaré estimate:
\[
\left( \int_T |u - Iu|^p \, dx' \right)^{1/p} \leq C \left( \int_{\hat{T} \times \mathbb{R}} |\hat{\nabla} u|^2 \, d\hat{x} \right)^{1/2} \tag{5.7}
\]
with a universal constant \( C \). Let \( \hat{x}' \mapsto A \hat{x}' + b \) be the affine map which maps \( \hat{T} \) onto \( T \). We change coordinates \( x = (x', x_3) \) according to
\[
x' = A \hat{x}', \quad \hat{x}' = A' \nabla', \quad x_3 = |A'| \hat{x}_3, \quad \hat{x}_3 = |A'| \partial_3. \tag{5.8}
\]
Then we have
\[
\left( \int_T |u - Iu|^p \, dx' \right)^{1/p} = \left( |\det A| \int_{\hat{T}} |u - Iu|^p \, d\hat{x}' \right)^{1/p} \leq C |\det A|^{1/p} \left( \int_{\hat{T} \times \mathbb{R}} |\hat{x}'|^2 + |\hat{x}_3'|^2 \, d\hat{x} \right)^{1/2} \leq C |\det A|^{1/p} \left( |A'|^2 \int_{T \times \mathbb{R}} |\nabla' u|^2 + |\partial_3 u|^2 \, dx \right)^{1/2} \leq C |\det A|^{1/p} \left( |A'| |\det A|^{-1} \int_{T \times \mathbb{R}} |\nabla' u|^2 + |\partial_3 u|^2 \, dx \right)^{1/2} = C |\det A|^{1/p - 1/2} |A|^{1/2} \left( \int_{T \times \mathbb{R}} |\nabla' u|^2 + |\partial_3 u|^2 \, dx \right)^{1/2}.
\]
Now (5.6) follows because of \(\frac{1}{2} |\det A| = \text{area}(T)\) and \(|A| \leq C \text{diam}(T)\).

We will distinguish between interior and boundary triangles.

- **Interior triangles** \(T\) are those for which
  \[
  \text{dist}(x'_T, \partial \Omega') \geq (2h)^{1-\alpha}.
  \]  
  Those triangles satisfy in particular
  \[
  \text{diam}(T) \leq h \text{dist}(x'_T, \partial \Omega') \leq \frac{1}{2} \text{dist}(x'_T, \partial \Omega'),
  \]  
  so that
  \[
  \forall x' \in T \quad \frac{1}{2} \text{dist}(x'_T, \partial \Omega') \leq \text{dist}(x', \partial \Omega') \leq 2 \text{dist}(x'_T, \partial \Omega').
  \]  
  These triangles furthermore have the property that
  \[
  \text{dist}(T, \partial \Omega') \geq \frac{1}{2} \text{dist}(x'_T, \partial \Omega') \geq h^{\frac{1}{1-\alpha}},
  \]  
  so that
  \[
  T \subset \{ x' \in \Omega' \mid \text{dist}(x', \partial \Omega') \geq h^{\frac{1}{1-\alpha}} \} =: \Omega'_h.
  \]  
  - **Boundary triangles** \(T\) are those for which (5.9) fails, i.e.
    \[
    \text{dist}(x'_T, \partial \Omega') < (2h)^{\frac{1}{1-\alpha}}.
    \]  
    Since
    \[
    \text{diam}(T) \leq h \text{dist}(x'_T, \partial \Omega') \leq (2h)^{\frac{1}{1-\alpha}},
    \]  
    it follows from (5.13) that
    \[
    \forall x' \in T \quad \text{dist}(x', \partial \Omega') \leq (4h)^{\frac{1}{1-\alpha}}.
    \]  
    We retain
    \[
    T \subset \{ x' \in \Omega' \mid \text{dist}(x', \partial \Omega') \leq (4h)^{\frac{1}{1-\alpha}} \} =: \partial \Omega'_h.
    \]  

We call \(T_{h,\text{int}}, T_{h,\text{bdry}}\) the partitioning of \(T_h\).

We start with an interior triangle \(T \in T_{h,\text{int}}\). By Cauchy–Schwarz inequality

\[
\left| \int_T (I \sigma^* - \sigma^*)(I u - u) \, dx' \right| \leq \left( \int_T |I \sigma^* - \sigma^*|^2 \, dx' \right)^{1/2} \left( \int_T |I u - u|^2 \, dx' \right)^{1/2}.
\]  

For the first factor, we use Poincaré’s estimate for a convex set \(T\)

\[
\left( \int_T |I \sigma^* - \sigma^*|^2 \, dx' \right)^{\frac{1}{2}} \leq C \left( \text{diam}(T)^2 \int_T |\nabla \sigma^*|^2 \, dx' \right)^{1/2}.
\]
and appeal to Theorem 1
\[
\left( \int_T |\nabla \sigma^*|^2 \, dx' \right)^{1/2} \leq C |H'_{ext}| \left( \int_T \text{dist}(x', \partial \Omega')^{-3} \text{dist}(x', P)^{-1} \, dx' \right)^{1/2},
\]
where \( P = \{(\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, -\frac{1}{2}), (-\frac{1}{2}, \frac{1}{2}), (-\frac{1}{2}, -\frac{1}{2})\} \) denotes the set of corners.

For the second factor in (5.15), we use (5.6) for \( p = 2 \):
\[
\left( \int_T |I u - u|^2 \, dx' \right)^{1/2} \leq C \left( \text{diam}(T) \int_{T \times \mathbb{R}} |\nabla u|^2 \, dx \right)^{1/2}.
\]
The combination yields
\[
\left| \int_T (I \sigma^* - \sigma^*)(I u - u) \, dx' \right| \leq C |H'_{ext}| \left( \text{diam}(T)^3 \int_T \text{dist}(x', \partial \Omega')^{-3} \text{dist}(x', P)^{-1} \, dx' \int_{T \times \mathbb{R}} |\nabla u|^2 \, dx \right)^{1/2}.
\]
(5.16)

We now observe
\[
\text{diam}(T)^3 \int_T \text{dist}(x', \partial \Omega')^{-3} \text{dist}(x', P)^{-1} \, dx' \overset{(5.2)}{\leq} h^3 \text{dist}(x_T, \partial \Omega)^3 \int_T \text{dist}(x', \partial \Omega')^{-3} \text{dist}(x', P)^{-1} \, dx' \overset{(5.11)}{\leq} C h^3 \int_T \text{dist}(x', \partial \Omega')^{3(\alpha - 1)} \text{dist}(x', P)^{-1} \, dx'.
\]
Thus (5.16) turns into
\[
\left| \int_T (I \sigma^* - \sigma^*)(I u - u) \, dx' \right| \leq C |H'_{ext}| h^{3/2} \left( \int_T \text{dist}(x', \partial \Omega')^{3(\alpha - 1)} \text{dist}(x', P)^{-1} \, dx' \int_{T \times \mathbb{R}} |\nabla u|^2 \, dx \right)^{1/2}. \quad (5.17)
\]

Cauchy–Schwarz yields in view of (5.12)
\[
\left| \sum_{T \in \mathcal{T}_{h, int}} \int_T (I \sigma^* - \sigma^*)(I u - u) \, dx' \right| \overset{(5.12)}{\leq} C |H'_{ext}| h^{3/2} \left( \int_{\Omega_h} \text{dist}(x', \partial \Omega)^3 \text{dist}(x', P)^{-1} \, dx' \int_{\mathbb{R}^3} |\nabla u|^2 \, dx \right)^{1/2}. \quad (5.18)
\]

Hence we are lead to discuss the scaling of the singular integral
\[
\int_{\Omega_h} \text{dist}(x', \partial \Omega)^3 \text{dist}(x', P)^{-1} \, dx'. \quad (5.19)
\]
in $h$. By symmetry, it is enough to consider one eighth of $\Omega_h'$, say $\Omega_h' \cap (\frac{-1}{2}, 0)^2 \cap \{x_1 > x_2\}$. After translation of the corner $(-\frac{1}{2}, -\frac{1}{2})$ into $(0, 0)$, we have $\text{dist}(x', \partial \Omega') = x_2$ and $\text{dist}(x', P) \geq x_1$ so that we need to consider
\[
\int_{\frac{h}{\sqrt{2}}}^{\frac{h}{\sqrt{2}}} x_2^{3(\alpha-1)} x_1^{-1} dx_2 dx_1 = \int_{\frac{h}{\sqrt{2}}}^{\frac{h}{\sqrt{2}}} x_2^{3(\alpha-1)} dx_2 dx_1.
\]
We notice
\[
\int_{\frac{h}{\sqrt{2}}}^{\frac{h}{\sqrt{2}}} x_2^{3(\alpha-1)} dx_2 \leq C \left\{ \begin{array}{ll}
x_1^{3 \alpha - 2} & \alpha > \frac{2}{3} \\
 h^{\frac{3 \alpha - 2}{\alpha - 1}} & \alpha < \frac{2}{3}
\end{array} \right\
\]
Hence we obtain for (5.19)
\[
\int_{\Omega_h'} \text{dist}(x', \partial \Omega')^{3(\alpha-1)} dx' \leq C \left\{ \begin{array}{ll}
\frac{1}{2} & \text{for } \alpha > \frac{2}{3} \\
 (\ln \frac{1}{h}) h^{\frac{3 \alpha - 2}{\alpha - 1}} & \text{for } \alpha < \frac{2}{3}
\end{array} \right\
\]
We use this in (5.18):
\[
\left| \sum_{T \in \mathcal{T}_{h, \text{int}}} \int_T (I \sigma^* - \sigma^*)(I u - u) \ dx' \right|
\leq C |H_{\text{ext}}'| \left( \int_{\mathbb{R}^3} |\nabla u|^2 \ dx \right)^{1/2} \left\{ \begin{array}{ll}
h^{3/2} & \text{for } \alpha > \frac{2}{3} \\
 (\ln^{1/2} \frac{1}{h}) h^{\frac{3 \alpha - 2}{\alpha - 1}} & \text{for } \alpha < \frac{2}{3}
\end{array} \right\}. \quad (5.20)
\]
We now turn to a boundary triangle $T \in \mathcal{T}_{h, \text{bdry}}$. We start with Hölder’s inequality
\[
\left| \int_T (I \sigma^* - \sigma^*)(I u - u) \ dx' \right|
\leq \left( \int_T |I \sigma^* - \sigma^*|^{4/3} \ dx' \right)^{3/4} \left( \int_T |I u - u|^4 \ dx' \right)^{1/4}
\leq 2 \left( \int_T |\sigma^*|^{4/3} \ dx' \right)^{3/4} \left( \int_T |I u - u|^4 \ dx' \right)^{1/4}.
\]
For the first factor, we appeal to Theorem 1. For the second factor, we evoke the trace estimate (5.6) with $p = 4$. This yields
\[
\left| \int_T (I \sigma^* - \sigma^*)(I u - u) \ dx' \right|
\leq C |H_{\text{ext}}'| \left( \int_T \text{dist}(x', \partial \Omega')^{-2/3} \text{dist}(x', P)^{-2/3} \ dx' \right)^{3/4}
\times \frac{\text{diam}(T)^{1/2}}{\text{area}(T)^{1/4}} \left( \int_{T \times \mathbb{R}} |\nabla u|^2 \ dx \right)^{1/2}.
\quad (5.21)
\]
We now need the following reverse Hölder estimate
\[
\int_T \text{dist}(x', \partial \Omega')^{-2/3} \text{dist}(x', P)^{-2/3} \, dx' \\
\times \left( \int_T \text{dist}(x', \partial \Omega') \, dx' \right)^{2/3} \left( \int_T \text{dist}(x', P) \, dx' \right)^{2/3} \\
\leq C \text{area}(T)^{7/3}.
\] (5.22)

By symmetry, it is enough to consider triangles \( T \) in \( \Omega' \cap (0, 1/2)^2 \cap \{ x_1 > x_2 \} \). After translation of the corner \((-1/2, -1/2)\) into \((0, 0)\), we have \( \text{dist}(x', \partial \Omega') = x_2 \) and \( x_1 \leq \text{dist}(x', P) \leq 2x_1 \) so that we need to show
\[
\left( \int_T x_2^{-2/3} x_1^{-2/3} \, dx' \right) \left( \int_T x_2 \, dx' \right)^{2/3} \left( \int_T x_1 \, dx' \right)^{2/3} \leq C \text{area}(T)^{7/3}.
\] (5.23)

The worst case is when \( T \) lies in the corner \((0, 0)\), i.e. when \( T \) is half of the rectangle \((0, \ell_1) \times (0, \ell_2)\). In this case (5.23) follows from
\[
\int_T x_2^{-2/3} x_1^{-2/3} \, dx' \leq \int_0^\ell_1 x_1^{-2/3} \, dx_1 \int_0^\ell_2 x_2^{-2/3} \, dx_2 \leq C \ell_1^{1/3} \ell_2^{1/3}, \\
\int_T x_2 \, dx' \leq \ell_1 \int_0^\ell_2 x_2 \, dx_2 \leq C \ell_1 \ell_2, \\
\int_T x_1 \, dx' \leq \ell_2 \int_0^\ell_1 x_1 \, dx_1 \leq C \ell_1^2 \ell_2, \\
\text{area}(T) = \frac{1}{2} \ell_1 \ell_2.
\]

We also need the following consequence of the standard Hölder inequality:
\[
\text{area}(T)^{3-\alpha} \left( \int_T \text{dist}(x', \partial \Omega') \, dx' \right)^{\alpha-1} \left( \int_T \text{dist}(x', P) \, dx' \right)^{-1} \\
\leq \int_T \text{dist}(x', \partial \Omega')^{\alpha-1} \text{dist}(x', P)^{-1} \, dx'.
\] (5.24)

Indeed (5.24) can be reformulated as
\[
\text{area}(T) \leq \left( \int_T \text{dist}(x', \partial \Omega')^{\alpha-1} \text{dist}(x', P)^{-1} \, dx' \right)^{\frac{1}{1-\alpha}} \\
\times \left( \int_T \text{dist}(x', \partial \Omega') \, dx' \right)^{\frac{1}{1-\alpha}} \left( \int_T \text{dist}(x', P) \, dx' \right)^{-\frac{1}{1-\alpha}},
\]
which follows from Hölder’s inequality by writing
\[
1 = \text{dist}(x', \partial \Omega')^{\frac{1}{1-\alpha}} \text{dist}(x', P)^{\frac{1}{1-\alpha}} \text{dist}(x', \partial \Omega')^{\frac{1}{1-\alpha}} \text{dist}(x', P)^{\frac{1}{1-\alpha}}.
\]
We now use (5.2) in form of
\[
\text{diam}(T) \leq C h \left( \text{area}(T)^{-1} \int_T \text{dist}(x', \partial \Omega') dx' \right)^\alpha
\]
and the reverse Hölder estimate (5.22) in (5.21). This yields
\[
\left| \int_T (I \sigma^* - \sigma^*)(I u - u) \, dx' \right|
\leq C |H'_{\text{ext}}| \left[ \text{area}(T)^{7/3} \left( \int_T \text{dist}(x', \partial \Omega') \, dx' \right)^{-2/3} \left( \int_T \text{dist}(x', P) \, dx' \right)^{-2/3} \right]^{3/4}
\times \text{area}(T)^{-1/4} \left[ h \left( \text{area}(T)^{-1} \int_T \text{dist}(x', \partial \Omega') \, dx' \right)^{\alpha/2} \left( \int_{T \times \mathbb{R}} |\nabla u|^2 \, dx \right)^{1/2} \right]^{1/2}
= C |H'_{\text{ext}}| h^{1/2} \left[ \text{area}(T)^{3/2-\alpha} \left( \int_T \text{dist}(x', \partial \Omega') \, dx' \right)^{\alpha-1} \right]^{1/2}
\times \left( \int_{T \times \mathbb{R}} |\nabla u|^2 \, dx \right)^{1/2}.
\]
We now apply to the above the Hölder inequality (5.24):
\[
\left| \int_T (I \sigma^* - \sigma^*)(I u - u) \, dx' \right|
\leq C |H'_{\text{ext}}| h^{1/2} \left( \int_T \text{dist}(x', \partial \Omega')^{\alpha-1} \text{dist}(x', P)^{-1} \, dx' \right)^{1/2}
\times \left( \int_{T \times \mathbb{R}} |\nabla u|^2 \, dx \right)^{1/2}.
\] (5.25)
Finally, we sum (5.25) over all \( T \in \mathcal{T}_{h,\text{bdry}} \). In view of (5.14), this yields by Cauchy–Schwarz
\[
\left| \sum_{T \in \mathcal{T}_{h,\text{bdry}}} \int_T (I \sigma^* - \sigma^*)(I u - u) \, dx' \right|
\leq C |H'_{\text{ext}}| h^{1/2} \left( \int_{\partial \Omega' \cap h} \text{dist}(x', \partial \Omega')^{\alpha-1} \text{dist}(x', P)^{-1} \, dx' \right)^{1/2}
\times \left( \int_{\mathbb{R}^3} |\nabla u|^2 \, dx \right)^{1/2}.
\] (5.26)
Hence we need to discuss
\[
\int_{\partial \Omega' \cap h} \text{dist}(x', \partial \Omega')^{\alpha-1} \text{dist}(x', P)^{-1} \, dx'.
\] (5.27)
Analogously to (5.19) we see that (5.27) behaves as

\[
\int_0^{1/2} \int_0^{\min\{h^{-\alpha}, x_1\}} x_2^\alpha x_1^{-1} dx_2 dx_1 \\
= \int_0^{1/2} \alpha^{-1} \min\{h^{-\alpha}, x_1^\alpha\} x_1^{-1} dx_1 \\
\leq C(\ln \frac{1}{h}) h^{\frac{\alpha}{\alpha}}.
\]

Therefore (5.26) turns into

\[
\left| \sum_{T \in \mathcal{T}_{h, \text{bdry}}} \int_T (I\sigma^* - \sigma^*)(I u - u) dx' \right| \\
\leq C |H'_{\text{ext}}| \left( \ln \frac{1}{h} h^{-\alpha} \right)^{1/2} \left( \int_{\mathbb{R}^3} |\nabla u|^2 \, dx \right)^{1/2}. \tag{5.28}
\]

Since \(\frac{1}{2(1-\alpha)} > \frac{3}{2}\) for \(\alpha > \frac{2}{3}\), (5.5) follows from combining (5.20) with (5.28).

\[
\square
\]

6 A–priori Refined Triangulation and Complexity Considerations

The numerical analysis from Section 5 suggests to consider triangulations \(\mathcal{T}_h\) which are refined at the boundary according to

\[
\text{diam}(T) \sim h \, \text{dist}(x_T', \partial\Omega')^\alpha, \tag{6.1}
\]

where

- \(\text{diam}(T)\) denotes the diameter of a triangle \(T\) and \(\text{area}(T)\) its area,
- \(\text{dist}(x_T', \partial\Omega')\) denotes the distance of the barycenter \(x_T'\) of \(T\) to the boundary \(\partial\Omega'\).

The \(H^2\)-matrix calculus suggests to consider triangulations which are regular in the sense of

\[
\text{area}(T) \sim \text{diam}(T)^2. \tag{6.2}
\]

Given \(h\), we use the following algorithm to generate \(\mathcal{T}_h\) with property (6.1) & (6.2):

- We start from a uniform triangulation of mesh size \(h\), see Figure 2.
- Then we mark all triangles which do not satisfy (6.1). The marked triangles are divided into four triangles of the same shape; this preserves (6.2). Neighboring triangles are divided so that the triangulation property is restored.
This procedure is repeated until all triangles meet (6.1).

Figure 1 shows an example of a locally refined triangulation.

We now work out how the number of triangles scales in $h$, depending on $\alpha$. We find that there is a cross-over at $\alpha = \frac{1}{2}$.

**Lemma 4.** For any constants $\Lambda \geq 1$ and $0 \leq \alpha_0 < 1$ there exists a constant $C = C(\Lambda, \alpha_0) < \infty$ with the following property: Let $0 \leq \alpha \leq \alpha_0$, $0 < h \leq \frac{1}{4}$ and triangulations $T_h$ of $\Omega' = (0, 1)^2$, which are refined towards the boundary in the sense of

$$\forall T \in T_h \quad \text{diam}(T) \leq h \, \text{dist}(x'_T, \partial \Omega')^\alpha \leq \Lambda \, \text{diam}(T)$$  \hspace{1cm} (6.3)

and regular in the sense of

$$\forall T \in T_h \quad \text{diam}(T)^2 \leq \Lambda \, \text{area}(T)$$  \hspace{1cm} (6.4)

be arbitrary. Then the number $\# T_h$ of triangles scales as

$$\frac{1}{C} \# T_h \leq \begin{cases} \frac{1}{1-2\alpha} h^{-2} & \text{for } h \leq e^{-\frac{1}{\sqrt{1-2\alpha}}} \text{ and } \alpha < \frac{1}{2} \\ (\ln \frac{1}{h^2}) h^{-2} & \text{for } h \geq e^{-\frac{1}{\sqrt{1-2\alpha}}} \text{ or } \alpha = \frac{1}{2} \\ \frac{1}{2\alpha} h^{-\frac{1}{1-\alpha}} & \text{for } h \leq e^{-\frac{1}{\sqrt{1-2\alpha}}} \text{ and } \alpha > \frac{1}{2} \end{cases} \leq C T_h.$$  \hspace{1cm} (6.5)

**Proof of Lemma 4.**

It is convenient to think in terms of the local number density of triangles

$$n(x') := \text{area}(T)^{-1} \quad \text{for} \quad x' \in T$$

since we have

$$\# T_h = \int_{\Omega'} n(x') \, dx'.$$  \hspace{1cm} (6.6)

Hence we need to estimate the number density $n$ by below and above.

Let $C < \infty$ denote a generic constant which only depends on $\Lambda$ and $\alpha_0$. We use the notation $A \lesssim B$ if there exists such a $C$ with $A \leq CB$. We write $A \sim B$ if $A \lesssim B$ and $B \lesssim A$.

In this notation our assumptions (6.3) and (6.4) read

$$\text{diam}(T) \sim h \, \text{dist}(x'_T, \partial \Omega')^\alpha, \quad \text{diam}(T)^2 \lesssim \text{area}(T),$$  \hspace{1cm} (6.7)

respectively. Since $\text{area}(T) \lesssim \text{diam}(T)^2$ comes for free, we have for any $x' \in T$:

$$n(x') = \text{area}(T)^{-1} \sim \text{diam}(T)^{-2} \sim h^{-2} \, \text{dist}(x'_T, \partial \Omega')^{-2\alpha}.$$  \hspace{1cm} (6.8)

We now argue that

$$\text{dist}(x'_T, \partial \Omega') \sim \max \left\{ \text{dist}(x', \partial \Omega'), h^\frac{1}{2} \right\}.$$  \hspace{1cm} (6.9)
To this purpose, we split (6.9) into the two statements
\[
|\text{dist}(x'_T, \partial\Omega') - \text{dist}(x', \partial\Omega')| \leq \frac{1}{2} \text{dist}(x'_T, \partial\Omega') + Ch^{\frac{1}{2\alpha}} \quad (6.10)
\]
and
\[
\text{dist}(x'_T, \partial\Omega') \gtrsim h^{\frac{1}{2\alpha}}. \quad (6.11)
\]
Indeed, (6.10) implies
\[
\text{dist}(x'_T, \partial\Omega') \lesssim \max \left\{ \text{dist}(x', \partial\Omega'), h^{\frac{1}{2\alpha}} \right\}, \quad (6.12)
\]
whereas the combination of (6.10) and (6.11) yields
\[
\text{dist}(x'_T, \partial\Omega') \overset{(6.10)}{\lesssim} \text{dist}(x'_T, \partial\Omega') + Ch^{\frac{1}{2\alpha}} \lesssim \text{dist}(x'_T, \partial\Omega'). \quad (6.13)
\]
Combining (6.13) once more with (6.11) we obtain the reverse of (6.12), i.e.
\[
\text{dist}(x'_T, \partial\Omega') \gtrsim \max \left\{ \text{dist}(x', \partial\Omega'), h^{\frac{1}{2\alpha}} \right\}.
\]
So indeed (6.9) follows from (6.10) and (6.11).

Let us now argue in favour of (6.10):
\[
|\text{dist}(x'_T, \partial\Omega') - \text{dist}(x', \partial\Omega')| \leq \text{diam}(T) \leq \Lambda h \text{dist}(x'_T, \partial\Omega')^\alpha = \Lambda(2\alpha)^\alpha h \left( \frac{1}{2\alpha} \text{dist}(x'_T, \partial\Omega') \right)^\alpha.
\]

We now appeal to Young’s inequality \(ab \leq (1 - \alpha)a^{\frac{1}{1-\alpha}} + \alpha b^\frac{1}{\alpha}\) and obtain
\[
|\text{dist}(x'_T, \partial\Omega') - \text{dist}(x', \partial\Omega')| \leq (1 - \alpha)(\Lambda(2\alpha)^\alpha h)^\frac{1}{1-\alpha} + \frac{1}{2} \text{dist}(x'_T, \partial\Omega') \leq Ch^{\frac{1}{1-\alpha}} + \frac{1}{2} \text{dist}(x'_T, \partial\Omega').
\]

We now establish (6.11). We notice that (6.7) implies that the angles of \(T\) are uniformly bounded away from zero so that
\[
\text{dist}(x'_T, \partial\Omega') \geq \text{dist}(x'_T, \partial T) \sim \text{diam}(T) \overset{(6.3)}{\gtrsim} h \text{dist}(x'_T, \partial\Omega')^\alpha.
\]
This yields (6.11).

The combination of (6.8) and (6.9) gives
\[
n(x') \sim h^{-2} \left( \max \left\{ \text{dist}(x', \partial\Omega'), h^{\frac{1}{2\alpha}} \right\} \right)^{-2\alpha} = h^{-2} \min \left\{ \text{dist}(x', \partial\Omega')^{-2\alpha}, h^{-\frac{2\alpha}{2\alpha}} \right\}.
\]
Hence according to (6.6)
\[ \# \mathcal{T}_h \sim h^{-2} \int_{\Omega'} \min \left\{ \text{dist}(x', \partial \Omega')^{-2\alpha}, h^{-\frac{2\alpha}{1-\alpha}} \right\} \, dx'. \] (6.14)

It remains to analyze the asymptotic behaviour of the integral in (6.14) for $h \downarrow 0$. By symmetry, it is enough to consider one eighth of $\Omega'$, namely $\Omega'_\frac{1}{8} := \{(x_1, x_2) \in \Omega' \mid x_2 < x_1 < \frac{1}{2}\}$. There, we have
\[ \text{dist}(x', \partial \Omega') = x_2 \quad \text{for} \quad x' \in \Omega'_\frac{1}{8} \]
so that for $h \leq \frac{1}{4}$ (and thus $h^{\frac{1}{1-\alpha}} \leq \frac{1}{4}$)
\[
\int_{\Omega'} \min \left\{ \text{dist}(x', \partial \Omega')^{-2\alpha}, h^{-\frac{2\alpha}{1-\alpha}} \right\} \, dx' \\
= 8 \int_0^{\frac{1}{4}} \int_0^{x_1} \min \left\{ x_2^{-2\alpha}, h^{-\frac{2\alpha}{1-\alpha}} \right\} \, dx_2 \, dx_1 \\
= 8 \left\{ \int_0^{\frac{1}{4}} x_1 h^{-\frac{2\alpha}{1-\alpha}} \, dx_1 + \int_{\frac{1}{4}}^{\frac{1}{2}} \left( h^{\frac{1}{1-\alpha}} h^{-\frac{2\alpha}{1-\alpha}} + \int_{\frac{1}{4}}^{x_1} x_2^{-2\alpha} \, dx_2 \right) \, dx_1 \right\}. \]

In the case of $\alpha \neq \frac{1}{2}$ we therefore have
\[
\int_{\Omega'} \min \left\{ \text{dist}(x', \partial \Omega')^{-2\alpha}, h^{-\frac{2\alpha}{1-\alpha}} \right\} \, dx' \\
= 8 \left\{ \frac{1}{2} \left( h^{\frac{1}{1-\alpha}} \right)^2 h^{-\frac{2\alpha}{1-\alpha}} + \left( \frac{1}{2} - h^{\frac{1}{1-\alpha}} \right) h^{\frac{1}{1-\alpha}} h^{-\frac{2\alpha}{1-\alpha}} \\
+ \frac{1}{1-2\alpha} \int_{\frac{1}{4}}^{\frac{1}{2}} \left( x_1^{-2\alpha} - \left( h^{\frac{1}{1-\alpha}} \right)^{1-2\alpha} \right) \, dx_1 \right\} \\
= 8 \left\{ \frac{1}{2} h^2 + \frac{1}{2} h^{\frac{1}{1-\alpha}} - h^2 + \frac{1}{1-2\alpha} \frac{1}{2(1-\alpha)} \left( \frac{1}{2} \right)^{2(1-\alpha)} - \left( h^{\frac{1}{1-\alpha}} \right)^{2(1-\alpha)} \\
- \frac{1}{1-2\alpha} \left( \frac{1}{2} - h^{\frac{1}{1-\alpha}} \right) \left( h^{\frac{1}{1-\alpha}} \right)^{1-2\alpha} \right\} \\
= 8 \left\{ \frac{\alpha}{2(1-\alpha)} h^2 + \frac{1}{1-2\alpha} \left( \frac{1}{1-\alpha} \left( \frac{1}{2} \right)^{3-2\alpha} - \alpha h^{\frac{1-2\alpha}{1-\alpha}} \right) \right\}. \] (6.15)

To have a closer look at $\frac{1}{1-2\alpha} \left( \frac{1}{1-\alpha} \left( \frac{1}{2} \right)^{3-2\alpha} - \alpha h^{\frac{1-2\alpha}{1-\alpha}} \right)$ for $\alpha \approx \frac{1}{2}$, we introduce
\[ \beta := 1 - 2\alpha \in (-1, 1) \quad \text{and} \quad g := h^{\frac{1}{1-\alpha}} \in (0, \frac{1}{4}), \]


so that for $\beta \ll 1$

$$
\frac{1}{1-2\alpha} \left( \frac{1}{1-\alpha} \left( \frac{1}{2} \right)^{3-2\alpha} - \alpha h^{1-2\alpha} \right) = \frac{1}{2\beta} \left( \frac{1}{(1+\beta)^2} - (1 - \beta)g^\beta \right) \\
\sim \frac{1}{\beta}(1 - g^\beta) \\
= (\ln \frac{1}{g}) \beta \ln \frac{1}{g} \left( 1 - e^{-\beta \ln \frac{1}{g}} \right)
$$

$$
\sim (\ln \frac{1}{g}) \left\{ \begin{array}{ll}
\frac{1}{|\beta|} & \text{for } \beta \ln \frac{1}{g} \gtrsim 1 \\
1 & \text{for } |\beta \ln \frac{1}{g}| \lesssim 1 \\
-\frac{1}{\beta \ln \frac{1}{g}} e^{-\beta \ln \frac{1}{g}} & \text{for } -\beta \ln \frac{1}{g} \gtrsim 1
\end{array} \right. \\
= \left\{ \begin{array}{ll}
\frac{1}{|\beta|} & \text{for } \ln \frac{1}{g} \gtrsim \frac{1}{|\beta|} \text{ and } \beta > 0 \\
\ln \frac{1}{g} & \text{for } \ln \frac{1}{g} \lesssim \frac{1}{|\beta|} \\
\frac{1}{|\beta|} g^{-|\beta|} & \text{for } \ln \frac{1}{g} \lesssim \frac{1}{|\beta|} \text{ and } \beta < 0
\end{array} \right. \\
(6.16)
$$

Formula (6.15) and estimate (6.16) translate into

$$
\int_{\Omega'} \min \left\{ \text{dist}(x', \partial \Omega')^{-2\alpha} - h^{-2\alpha} \right\} dx' \\
\sim \left\{ \begin{array}{ll}
\frac{1}{1-2\alpha} & \text{for } h \leq e^{-\frac{1}{|\beta|}} \text{ and } \alpha < \frac{1}{2} \\
(\ln \frac{1}{g}) & \text{for } h \geq e^{-\frac{1}{|\beta|}} \text{ or } \alpha = \frac{1}{2} \\
h^{-\alpha} & \text{for } h \leq e^{-\frac{1}{|\beta|}} \text{ and } \alpha > \frac{1}{2}
\end{array} \right. \\
= \left\{ \begin{array}{ll}
\frac{1}{1-2\alpha} & \text{for } h \leq e^{-\frac{1}{|\beta|}} \text{ and } \alpha < \frac{1}{2} \\
\frac{1}{2\alpha-1} h^{-2\alpha} & \text{for } h \leq e^{-\frac{1}{|\beta|}} \text{ and } \alpha > \frac{1}{2}
\end{array} \right.
$$

In view of (6.14), this yields the proposition for $\alpha \neq \frac{1}{2}$. The case $\alpha = \frac{1}{2}$ follows by continuity.

Table 1 and Table 2 show that the theoretical asymptotic result captures the actual experimental behavior.

In view of Theorem 2, which states that the discretization error $\epsilon$ is estimated as

$$
\epsilon \lesssim \left\{ \begin{array}{ll}
h^\frac{\alpha}{3} & \text{for } \alpha \geq \frac{2}{3} \\
h^\frac{1}{3(1-\alpha)} & \text{for } \alpha \leq \frac{2}{3}
\end{array} \right.
$$

(we neglect logarithmic terms) and of (6.5) we now ask the following question: For which degree of refinement is the $\epsilon - N$ relation most favorable? A glance at (6.5) and (6.17) reveals

$$
\epsilon \lesssim \left\{ \begin{array}{ll}
N^{-\frac{3(1-\alpha)}{2}} & \text{for } \frac{2}{3} \leq \alpha \leq \frac{2}{3} \\
N^{-\frac{1}{3(1-\alpha)}} & \text{for } \frac{1}{2} \leq \alpha \leq \frac{2}{3}
\end{array} \right.
$$

Hence the $\alpha$ in the range $[\frac{1}{2}, \frac{2}{3}]$ lead to the best $\epsilon - N$ relation. In the sequel, we shall focus on $\alpha = \frac{2}{3}$.
Table 1: Local refinement with $\alpha = 2/3$

<table>
<thead>
<tr>
<th>$h$</th>
<th>$N$</th>
<th>$Nh^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>5,824</td>
<td>372.74</td>
</tr>
<tr>
<td>0.3</td>
<td>15,224</td>
<td>411.05</td>
</tr>
<tr>
<td>0.2</td>
<td>50,464</td>
<td>403.71</td>
</tr>
<tr>
<td>0.1</td>
<td>416,456</td>
<td>416.46</td>
</tr>
</tbody>
</table>

Table 2: Local refinement with $\alpha = 1/2$

<table>
<thead>
<tr>
<th>$h$</th>
<th>$N$</th>
<th>$Nh^2 \ln^{-1} \frac{1}{h}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>3,136</td>
<td>38.97</td>
</tr>
<tr>
<td>0.1</td>
<td>15,440</td>
<td>33.53</td>
</tr>
<tr>
<td>0.05</td>
<td>72,216</td>
<td>30.13</td>
</tr>
<tr>
<td>0.025</td>
<td>328,576</td>
<td>27.84</td>
</tr>
</tbody>
</table>

7 Discretization

With this section, we start the discussion of algorithms to solve the Galerkin ansatz (5.1). Let $\chi_1, \chi_2, \ldots, \chi_N$ be the characteristic functions associated to the triangles $T_1, T_2, \ldots, T_N \in T_h$ (with $N = \#T_h$). The ansatz

$$\sigma_h = \sum_{i=1}^{N} z_i \chi_i$$

for the solution of (5.1) leads to the linear system of algebraic equations

$$\sum_{i=1}^{N} z_i \int_{T_i} \int_{T_j} \frac{1}{4\pi |x' - y'|} dy' dx' + \int_{T_j} H'_{ext} \cdot x' dx' = 0 .$$

These can be written in matrix–vector form as

$$Kz + b = 0,$$  \hspace{1cm} (7.1)

where $K$ is a positive definite and symmetric matrix with entries

$$K_{ij} := \int_{T_i} \int_{T_j} \frac{1}{4\pi |x' - y'|} dy' dx' ,$$  \hspace{1cm} (7.2)

$z = (z_j)_{j=1,\ldots,N}$ is the vector of the unknowns and $b$ is defined by

$$b_j := \int_{T_j} H'_{ext} \cdot x' dx'.$$

$K_{ij}$ represents the interaction energy of unit charges placed in $T_i$ and $T_j$. Due to the non–locality of the magnetostatic interaction, the matrix $K$ is fully populated.
We calculate the entries $K_{ij}$ by a semi-analytical approach. The interior integral in (7.2) can be computed analytically [1], whereas the exterior integral is evaluated by a high-order Gaussian quadrature rule [9].

Using direct methods such as Gaussian elimination or Cholesky factorization to solve (7.1) is prohibitive for large problems. Iterative solvers in general need $O(N^2)$ operations per iteration, so computational cost still becomes excessive for large $N$. Indeed, any algorithm based on an explicit representation of $K$ has at least $O(N^2)$—complexity in time and memory.

To solve (7.1) in sub-quadratic time, we combine the conjugate gradient method with two different methods of matrix compression:

- For a uniform triangulation (U) as in Figure 2, the interaction matrix $K$ bears a Toeplitz-type structure, i.e. the matrix entries depend only on the distance vector between the corresponding cells. This discrete convolution structure of course reflects the continuous convolution structure (3.2). A matrix-vector product in form of a discrete convolution can be efficiently computed by FFT with complexity $O(N \log N)$. By simply filling up the data by zeros up to the next dyadic size, no periodicity is required [2, 21].

- For a priori locally refined triangulations (R) (see Figure 1) we implemented a more flexible, yet (near) optimal-complexity algorithm to carry out the convolution, based on the concept of $H$-matrices (hierarchical matrices) [14]. $H$-matrices are natural approximations of fully populated stiffness matrices as they appear in the finite element-based Galerkin discretization of non-local integral operators: Depending on the smoothness properties of the kernel, submatrices of the stiffness matrix are replaced by suitable low rank approximations. We implemented a new class of hierarchical matrices, the so-called $H^2$-matrices [15], see Section 8.

### 7.1 FFT-Based Algorithm

We consider a triangulation which comes from a cartesian grid and is divided into lower and upper triangles, see Figure 2. For a suitable choice of indices the matrix $K$ and the vector $z$ have a special block structure

$$
K = \begin{pmatrix} K^{(l,l)} & K^{(l,u)} \\ K^{(u,l)} & K^{(u,u)} \end{pmatrix}, \quad z = \begin{pmatrix} z^{(l)} \\ z^{(u)} \end{pmatrix},
$$

where $z^{(l)}$ represents the data on lower triangles and $z^{(u)}$ those on upper triangles. $K^{(l,l)}$ is the interaction matrix for lower triangles, $K^{(l,u)}$ is the interaction matrix for one fixed lower triangle with all the upper ones, and so on.

Since the given triangulation is translation invariant, each of the $K$-blocks bears a Toeplitz-type structure, i.e. the matrix elements depend only on the distance
vector between the corresponding cells, thus instead of storing $N^2/4$ matrix elements we only have to store a vector of length $N/2$. The mathematical form of the matrix–vector multiplication $K^{(l,j)}z^{(l)}$ is that of a discrete convolution, reflecting the continuous convolution structure $(3.2)$.

Direct computation of the discrete convolution of two $n$–long sequences requires $O(n^2)$ operations. Computational cost is reduced by taking advantage of the discrete convolution theorem: The discrete Fourier transform (DFT) of the convolution of two discrete periodic functions equals the componentwise product of the DFTs of the two functions. By performing the DFT with the help of Fast Fourier Transform (FFT) algorithms, the over–all complexity of the discrete convolution is reduced to $O(n \log n)$.

Due to the non–periodicity of our data we have to extend the vector $z$ by zeros. This technique of zero–padding avoids any end effects as described in [21, Chapter 12] but increases the size of the input data for the FFT in 2d by a factor of four.

Once the DFTs of the $K$–blocks are given, the matrix–vector product $Kz$ is computed by applying a 2d–FFT routine four times (two forward and two backward transforms). Since the input data is purely real we can take advantage of a special real–to–complex variant of the FFT improving speed and memory usage roughly by
a factor of two.

If $n$ is a power of 2, the classical Cooley–Tukey FFT algorithm for $n$ complex data requires asymptotically $\frac{n}{2} \log_2 n$ complex multiplications, whereas the real-to-complex FFT of $n$ real data results in asymptotically $n \log_2 n$ real multiplications. So if $N = 2^\nu$ is the total number of triangles (i.e. we have $N/2$ lower and $N/2$ upper triangles), for each of the four FFTs the data-size is $n = 4 \cdot \frac{N}{2}$ (the factor 4 is due to zero-padding), which results in a total number of approximately $8N \log_2 N$ real multiplications in the computation of $Kz$. The number of multiplications in frequency space is of order $N$ and thus asymptotically negligible.

In practice, so-called split-radix FFTs that rely on a combination of elementary 2-point and 4-point transforms are more efficient than pure radix-2 algorithms like the Cooley–Tukey.

8 $\mathcal{H}$–Matrices

The notion of $\mathcal{H}$–matrices (hierarchical matrices) was introduced in [14]. These matrices are data-sparse: they can be described by relatively few data. $\mathcal{H}$–matrices are natural approximations of fully populated stiffness matrices as they appear in finite element–based Galerkin discretization of non-local integral operators [18]. The use of $\mathcal{H}$–matrices reduces storage requirements and the complexity of matrix-vector multiplication to almost linear complexity $\mathcal{O}(N \log N)$.

The reliability of $\mathcal{H}$–matrices for the approximation of integral operators is based on the smoothness properties of the kernel. Typically, one should think of the kernel as a singular convolution kernel as in (3.2). Hence the kernel becomes singular only at the diagonal of the stiffness matrix $K$. Away from the diagonal one replaces submatrices of $K$ by suitable low rank approximations. This is motivated by the fact that off–diagonal submatrices are numerically of low rank, i.e. the columns are nearly linearly dependent. The approximate matrix is then stored in a hierarchical data structure of size $\mathcal{O}(N \log N)$, thus allowing an evaluation of matrix-vector products in $\mathcal{O}(N \log N)$ time. This is a common strategy, particularly in boundary element methods and the $n$-body problem, leading to algorithms like the panel clustering technique [19] and the fast multipole method [13]. The basic idea can be summarized...
by the following recipe: The near-field component of the long range interaction is evaluated by direct computation whereas the far-field component is approximated using a hierarchical clustering of distant elements.

The complexity of the matrix-vector multiplication is further reduced to optimal complexity $O(N)$ by a new class of hierarchical matrices, the so-called $H^2$–matrices, introduced in [17]. Let us focus on interpolation–based low rank approximations: Polynomial interpolation in one argument of the kernel $k(.,.)$ leads to $H$–matrix approximations, interpolation in both arguments leads to so–called uniform $H$–matrices. The polynomial interpolation allows for a hierarchical change from the finite element basis to the cluster basis and vice versa; this is the $H^2$–matrix approximation with optimal complexity in the matrix–vector multiplication. For a comparison of the different types of hierarchical matrices, see [16].

$H^2$–matrices have already been applied in computational electromagnetism to deal with non–sparse matrices arising from Galerkin boundary element discretizations [3]. An application of $H^2$–matrices to 3–d stray field computation in micromagnetics is given in [5], where the large–body limit derived in [6] by $\Gamma$–convergence is considered.

The $H^2$–matrix method consists of a setup phase, i.e. the computation of the compressed matrix representation (which has to be done only once for a given triangulation), and an algorithm for fast matrix–vector multiplication.

We suppress all primes for the two–dimensional variables in this section.

### 8.1 Setup Phase

The setup phase starts with a rearrangement of the data structure. Let $I$ denote the index set of the finite element basis. In our case, $I$ just enumerates the triangles and has cardinality $N$.

- $I$ is decomposed along a binary tree introducing clusters $\tau \subset I$, see Figure 3. The subset $T_\tau \subset \Omega$ is the union of the supports of the basis functions $\chi_i$, $i \in \tau$. In our case $T_\tau$ is just the union of the triangles $T_i$, $i \in \tau$, see Figure

![Figure 3: Cluster tree for $I$](image-url)
The bounding box $B_\tau \subset \mathbb{R}^2$ is the smallest rectangle with axes in $x_1$ and $x_2$ direction which contains $T_\tau$, see Figure 4. In order to generate the tree, we use a cardinality–based algorithm:

- $I$ is the root of the binary tree.
- For each cluster $\tau$, we split the corresponding boundary box $B_\tau$ along its longer side. We split it such that each side contains the same number of barycenters of triangles $i \in \tau$. This defines the two son clusters.
- We stop when a cluster contains $\leq p^2$ triangles. This introduces a parameter $p$ which we choose to be $p = 3$.

Notice that by construction on each level of the binary tree the clusters have comparable cardinality $\#\tau_1 \sim \#\tau_2$ and the bounding boxes are not too anisotropic in the sense of $\text{area}(B_\tau) \sim \text{diam}(B_\tau)^2$.

- A pair of clusters $\tau_1 \times \tau_2$ is called an admissible block if a geometric condition is satisfied that bounds the diameter of the clusters by their distance, i.e.

$$\max\{\text{diam}(B_{\tau_1}), \text{diam}(B_{\tau_2})\} \leq 2\eta \text{ dist}(B_{\tau_1}, B_{\tau_2}) .$$

This introduces a parameter $\eta$. It has turned out in our experiments that a good value for the parameter $\eta$ to choose here is $\eta = 0.9$. This defines a block (= pair of clusters) tree, which is a decomposition of $I \times I$, by the following algorithm:

- $I \times I$ is the root of the block tree.
- Recursively, each block $\tau_1 \times \tau_2$ is decomposed into the four sub–blocks (formed by the pairs of son clusters of $\tau_1, \tau_2$) until one of the following two termination criteria holds:
  1. $\tau_1 \times \tau_2$ is admissible,
  2. $\tau_1$ or $\tau_2$ is a leaf.

The leaves of this block tree induce a partitioning of $I \times I$, see Figure 5.
8.2 Low Rank Approximation of the Kernel

The block partitioning of $I \times I$ induces a block partitioning of the matrix $K$. The idea is to approximate the restriction of $K$ on each admissible block $\tau_1 \times \tau_2$ (the shaded ones in Figure 5) by a matrix of fixed rank $p^2$. This is done by (tensor product) Chebyshev interpolation of the kernel $k(x, y)$, $(x, y) \in \Omega \times \Omega$, with polynomials of degree $(p-1)$ in both arguments. More precisely, for every cluster $\tau$, let $(x^\tau_i)$, denote the $p^2$ interpolation points in the bounding box $B_\tau$. They are given by the zeros of the Chebyshev polynomials. Let $p^\tau_{i\kappa}$ be the corresponding Lagrange polynomials. For an admissible cluster $\tau_1 \times \tau_2$, the kernel function $k(x, y)$ is replaced by

$$
\tilde{k}(x, y) = \sum_{i,\kappa=1}^{p^2} k(x^\tau_1_i, y^\tau_2_\kappa) p^\tau_{i\tau_1}(x) p^\tau_{\kappa\tau_2}(y)
$$

(8.1)

for $(x, y) \in B_{\tau_1} \times B_{\tau_2}$. On the blocks which are not admissible (the non–shaded area in Figure 5), $K$ remains unchanged. Notice that this in particular holds for the entries near the singular diagonal. Hence the approximation distinguishes between the far–field and near–field components of the long range interaction described by the kernel function $k$.

8.3 Matrix–Vector Multiplication

Consider the operator $\tilde{K}$ defined by the kernel $\tilde{k}$ in (8.1). Its matrix representation $(\tilde{K}_{ij})_{(i,j) \in I \times I}$ w.r.t. the canonical finite element basis is given by

$$
\tilde{K}_{ij} = \sum_{i,\kappa} \int_{T_i} p^\tau_{i\tau_1}(x) dx \int_{T_j} p^\tau_{\kappa\tau_2}(y) dy ,
$$

(8.2)

where $\tau_1 \times \tau_2$ is the unique block in the partitioning of $I \times I$ which contains $(i, j)$. Remember that $K_{ij}$ is only replaced by $\tilde{K}_{ij}$ if $\tau_1 \times \tau_2$ is an admissible block. We
write this as
\[ K_{ij} = \sum_{\iota,\kappa} V_{i\iota}^{\tau_1} S_{i\kappa}^{\tau_1 \times \tau_2} V_{j\kappa}^{\tau_2} = [V^{\tau_1} S^{\tau_1 \times \tau_2} (V^{\tau_2})^T]_{ij}, \] (8.3)

where
\[ S_{i\kappa}^{\tau_1 \times \tau_2} := k(x_{i\iota}^{\tau_1}, y_{\kappa}^{\tau_2}), \quad V_{i\iota}^{\tau_1} := \int_{T_{i\iota}} p_{i\iota}^{\tau_1} (x) dx \]

with the implicit understanding that \( i \in \tau_1 \). Notice that the action of \( V^{\tau_1} \) can be interpreted as a transformation from the finite element basis (as enumerated by \( i \in I \)) to the cluster basis (as parametrized by all the clusters of the binary tree). We observe that each cluster corresponds to \( p^2 \) data. Hence the matrix–vector multiplication can be interpreted as a forward transformation \( V^T \), a matrix–vector multiplication with \( S \), and a backward transformation \( V \).

The additional idea of \( H^2 \)-matrices is to organize the transformation \( V \) in a hierarchical way. It is based on the insight that the Lagrange polynomials \( (p_i^\tau(x))_i \) for a cluster \( \tau \) can be expressed as a linear combination of \( (p_\kappa^\tau(x))_\kappa \) for any other cluster \( \tau_1 \), in particular one of the two son clusters of \( \tau \):
\[ p_i^\tau(x) = \sum_\kappa p_\kappa^\tau(x)p_i^{\tau_1}(x_{\kappa}^{\tau_1}). \]

This implies the hierarchical relationship
\[ V_{i\iota}^{\tau} = \sum_\kappa V_{i\kappa}^{\tau_1} B_{i\kappa}^{\tau_1,\tau} = [V^{\tau_1} B^{\tau_1,\tau}]_{i\iota}, \] (8.4)

provided \( i \in \tau_1 \subset \tau \) where
\[ B_{i\kappa}^{\tau_1,\tau} := p_i^{\tau_1}(x_{\kappa}^{\tau_1}). \]

Hence the forward transformation requires an initial transformation from the finite element basis to the leaves of the cluster tree. Then one uses (8.4) to get recursively down the binary tree to the root.

### 8.4 Complexity of the Matrix–Vector Multiplication

Within the above cardinality–based algorithm, the cluster tree (see Figure 3) is nearly balanced, which means

- it has approximately \( N/p^2 \) leaves and
- it has approximately \( N/p^2 \) internal nodes.

In our case, the generated block tree is well–behaved in the following sense:

- The number of blocks in the partitioning of \( I \times I \) (see Figure 5) is approximately proportional to the number of leaves \( N/p^2 \), see Table 3 for a uniform triangulation and Tables 4 and 5 for a locally refined triangulation as described in Section 6. This property can be theoretically derived for any triangulation which is regular in the sense of 6.2, see [12, Lemma 4.5]. The prefactor depends on the type of the triangulation and, of course, on the parameter \( \eta \).
From these data we conclude

- Each involved $V$– and $B$–matrix has $p^2 \cdot p^2 = p^4$ entries. The $V$–matrices are applied on the $\approx N/p^2$ leaves. Each time one moves from two son clusters to their father cluster, two $B$–matrices are involved. There are $\approx N/p^2$ internal nodes. Hence the forward transformation requires $\approx (p^4 + 2p^4) \cdot (N/p^2) = 3Np^2$ multiplications.

- The backward transformation likewise requires $\approx 3Np^2$ multiplications.

- Each $S$–matrix has $p^4$ entries. There is an $S$–matrix for each of the $O(N/p^2)$ blocks in the partitioning of $I \times I$. Hence the matrix multiplication with $S$ takes $O(Np^2)$ multiplications.

<table>
<thead>
<tr>
<th>Table 3: Uniform triangulation</th>
<th>$N$</th>
<th>#blocks</th>
<th>#blocks/(Np^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8,192</td>
<td>24,808</td>
<td>0.336</td>
<td></td>
</tr>
<tr>
<td>32,768</td>
<td>104,530</td>
<td>0.354</td>
<td></td>
</tr>
<tr>
<td>131,072</td>
<td>430,378</td>
<td>0.365</td>
<td></td>
</tr>
<tr>
<td>524,288</td>
<td>1,744,468</td>
<td>0.370</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4: Local refinement with $\alpha = 2/3$</th>
<th>$h$</th>
<th>$N$</th>
<th>#blocks</th>
<th>#blocks/(Np^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>5,824</td>
<td>38,038</td>
<td>0.726</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>15,224</td>
<td>86,482</td>
<td>0.631</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>50,464</td>
<td>383,446</td>
<td>0.844</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>416,456</td>
<td>3,215,884</td>
<td>0.858</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5: Local refinement with $\alpha = 1/2$</th>
<th>$h$</th>
<th>$N$</th>
<th>#blocks</th>
<th>#blocks/(Np^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>3,136</td>
<td>14,260</td>
<td>0.5052</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>15,440</td>
<td>71,956</td>
<td>0.5178</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>72,216</td>
<td>266,554</td>
<td>0.4101</td>
<td></td>
</tr>
<tr>
<td>0.025</td>
<td>328,576</td>
<td>1,149,866</td>
<td>0.3888</td>
<td></td>
</tr>
</tbody>
</table>

9 Numerical Results

We compare the performance of both methods for a square $\Omega' = (0,1)^2$ and external field $H_{ext}' = (0.2,0.2)$. Since we do not know the exact solution $\sigma^*$ of (7.1), we derive a reliable approximation $E(\sigma^*)$ to the exact energy of $\sigma^*$ in the following way: We
compute the charge density $\sigma_h$ by the FFT–based conjugate gradient method with relative residuum $10^{-12}$ for several values of $h$. Linear extrapolation to $h = 0$ then yields $E(\sigma^*)$, since $E(\sigma_h)$ depends almost linearly on $h$ (see Theorem 2 with $\alpha = 0$).

Then we start the FFT–based and the $H^2$–based conjugate gradient method on triangulations (U) and (R) (see Section 5), respectively, and terminate the iteration when a prescribed error $\epsilon$ relative to $\sigma^*$ is attained. Here $\epsilon$ denotes the relative error between the approximate solution $\sigma_h$ and $\sigma^*$ in the homogeneous part of the continuous $H^{-1/2}(\mathbb{R}^2)$–norm:

$$\epsilon := \frac{\left(\int_{\mathbb{R}^2} |(\nabla')^{-1/2}(\sigma^* - \sigma_h)|^2 dx'\right)^{1/2}}{\left(\int_{\mathbb{R}^2} |(\nabla')^{-1/2}\sigma^*|^2 dx'\right)^{1/2}}.$$

We can express $\epsilon$ in terms of the energies since $E$ is the sum of a quadratic form and a linear functional

$$E(\sigma_h) = B(\sigma_h, \sigma_h) + L(\sigma_h),$$

where $B(\sigma_h, \sigma_h)$ is one half of the squared homogeneous part of the $H^{-1/2}(\mathbb{R}^2)$–norm of $\sigma_h$. Therefore we have

$$E(\sigma_h) - E(\sigma^*) = B(\sigma_h + \sigma^*, \sigma_h - \sigma^*) + L(\sigma_h - \sigma^*)$$

$$= B(\sigma_h - \sigma^*, \sigma_h - \sigma^*) + 2B(\sigma^*, \sigma_h - \sigma^*) + L(\sigma_h - \sigma^*)$$

$$= B(\sigma_h - \sigma^*, \sigma_h - \sigma^*),$$

since $2B(\sigma^*, \sigma_h - \sigma^*) + L(\sigma_h - \sigma^*) = 0$ is just the weak formulation of the Euler–Lagrange equation tested with $(\sigma_h - \sigma^*)$. So the relative error $\epsilon$ can be rewritten as

$$\epsilon = \sqrt{\frac{E(\sigma_h) - E(\sigma^*)}{E(\sigma^*)}}. \quad (9.1)$$

In particular, the notion of error is independent of the chosen grid. Notice that the conjugate gradient algorithm monotonically decreases the energy norm distance (9.1) to the solution $\sigma^*$, i.e. we have $E(\sigma^*) \leq E(\sigma_h^{(k+1)}) \leq E(\sigma_h^{(k)})$ for each iteration step $k$, see [21, Chapter 10.6].

Table 6 and Table 7 show the CPU times in seconds for the prescribed relative error $\epsilon$. $N$ is the number of degrees of freedom, given in Table 6 for a triangulation (U) and in Table 7 for a triangulation (R) (with the refinement rate $\alpha = 2/3$ a priori predicted by the numerical analysis, see the discussion in Section 5). Cells of the table without entry indicate that the prescribed relative error is not attainable on the given triangulation. Since we want to test the performance of FFT–based versus $H^2$–based matrix–vector products in the conjugate gradient iterations, we do not consider the CPU times for the setup phases and the computation of $\epsilon$. In the case of triangulation (R) the conjugate gradient method is preconditioned by the diagonal
matrix \( D = \text{diag}(K_{11}, \ldots, K_{NN}) \). This simple preconditioning has no effect on the performance of the conjugate gradient method in the case of triangulation (U). Our numerical experiments reveal a cross-over in computation time at a relative error of the order of 5%.

<table>
<thead>
<tr>
<th>( \epsilon )</th>
<th>( N=8,192 )</th>
<th>( N=32,768 )</th>
<th>( N=131,072 )</th>
<th>( N=524,288 )</th>
<th>( N=2,097,152 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.09</td>
<td>0.41</td>
<td>1.95</td>
<td>8.22</td>
<td>35.89</td>
</tr>
<tr>
<td>0.07</td>
<td>—</td>
<td>0.70</td>
<td>2.61</td>
<td>11.05</td>
<td>49.21</td>
</tr>
<tr>
<td>0.05</td>
<td>—</td>
<td>—</td>
<td>4.23</td>
<td>15.17</td>
<td>65.02</td>
</tr>
<tr>
<td>0.03</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>112.76</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \epsilon )</th>
<th>( N=5,824 )</th>
<th>( N=15,224 )</th>
<th>( N=50,464 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.18</td>
<td>0.72</td>
<td>1.72</td>
</tr>
<tr>
<td>0.07</td>
<td>0.24</td>
<td>0.72</td>
<td>1.72</td>
</tr>
<tr>
<td>0.05</td>
<td>0.86</td>
<td>1.34</td>
<td>2.79</td>
</tr>
<tr>
<td>0.03</td>
<td>—</td>
<td>—</td>
<td>5.13</td>
</tr>
</tbody>
</table>

A further remark on the computation of \( \epsilon \) must be added in the case of the \( H^2 \)-matrix method. Since in this method the evaluation of the energy \( E \) in (9.1) is based on the approximate matrix \( \tilde{K} \), we calculate only an approximation \( \tilde{\epsilon} \) to \( \epsilon \). To test the deviation numerically, we compute the exact energy norm distance \( \epsilon \) of solutions \( \sigma_h \) to \( \sigma^* \) with the help of the full matrix \( K \) in the case of a prescribed relative error of 5%. Table 8 shows that the deviation is of higher order.

<table>
<thead>
<tr>
<th>( \epsilon )</th>
<th>( \tilde{\epsilon} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0454</td>
<td>0.0500</td>
</tr>
<tr>
<td>0.0418</td>
<td>0.0486</td>
</tr>
<tr>
<td>0.0372</td>
<td>0.0423</td>
</tr>
</tbody>
</table>

Figure 6 shows the graph of a solution \( \sigma_h \) on a triangulation (R) with \( N = 5824 \) and prescribed relative error of 5%. For the convenience of the reader the piecewise constant function \( \sigma_h \) has been linearly interpolated.

In Figure 7 we compare the performance of a single matrix–vector multiplication \( Kz \) for the two matrix compression methods on triangulations (U). As non–trivial vector \( z \) we take the solution of the FFT–based conjugate gradient method computed above (see Table 6). Here the FFT–based method is exact up to rounding errors, whereas
Figure 6: Graph of a solution $\sigma_h$ (linearly interpolated) on triangulation (R).

the $H^2$–matrix method computes an approximate solution $\tilde{K}z$. Figure 8 shows the relative error
\[
\epsilon = \sqrt{\frac{z^T K z - z^T \tilde{K} z}{z^T K z}}
\]

between the corresponding exact and approximate quadratic form evaluated on $z$. This is the value of interest in our applications since $z^T K z$ is the discretized version of the stray field energy.

As expected, the FFT–based method beats the $H^2$–matrix method since for triangulations (U) the Toeplitz–type structure is a much stronger property than the block structure of the $H^2$–matrix induced by the (mesh–independent) self–similarity in the decay of the kernel function. Moreover, we take advantage of a highly optimized FFT implementation [11] achieving practically optimal complexity over the range accessible in our numerical experiments. All numerical experiments were done on a Pentium III processor running at 700 MHz.
Figure 7: CPU–time required for matrix–vector multiplication.

Figure 8: Relative error of $\mathcal{H}^2$ matrix–vector product.

10 Acknowledgements

We thank Ralf Hiptmair for introducing us to the $\mathcal{H}^2$–method and the Raviart–Thomas element. He also provided us with software. JD thanks Jörg Ostrowski and Lars Grasedyck for their advice on the implementation of hierarchical matrices. We thank Angela Kunoth for many helpful suggestions. JD is supported by the Deutsche Forschungsgemeinschaft through the SFB 611 at the university of Bonn.

References


Verzeichnis der erschienenen Preprints ab No. 125

<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
<th>Journal/Conference</th>
</tr>
</thead>
<tbody>
<tr>
<td>141.</td>
<td>Albeverio, Sergio; Liang, Song</td>
<td>A Remark on Different Lattice Approximations and Continuum Limits for $\phi_4^3$-Fields</td>
<td></td>
</tr>
<tr>
<td>142.</td>
<td>Albeverio, Sergio; Nizhnik, Leonid</td>
<td>A Schrödinger Operator with Point Interactions on the Sobolev Spaces</td>
<td>Letters in Mathematical Physics</td>
</tr>
<tr>
<td>143.</td>
<td>Labesse, Jean-Pierre; Müller, Werner</td>
<td>Weak Weyl’s Law for Congruence Subgroups</td>
<td></td>
</tr>
<tr>
<td>144.</td>
<td>Arndt, Marcel; Griebel, Michael</td>
<td>Higher Order Gradient Continuum Description of Atomistic Models for Crystalline Solids</td>
<td>Proceedings der Konferenz ECCOMAS 2004, Jyväskylä, Finnland</td>
</tr>
<tr>
<td>145.</td>
<td>Albeverio, Sergio; Herzberg, Frederik S.</td>
<td>On an Internal Random Walk Representation of Measurable Lévy Processes and their Stochastic Integrals</td>
<td></td>
</tr>
<tr>
<td>146.</td>
<td>Otto, Felix; Tzavaras, Athanasios E.</td>
<td>Continuity of Velocity Gradients in Suspensions of Rod-Like Molecules</td>
<td></td>
</tr>
<tr>
<td>147.</td>
<td>Menon, Govind; Otto, Felix</td>
<td>Dynamic Scaling in Miscible Viscous Fingering</td>
<td></td>
</tr>
<tr>
<td>148.</td>
<td>Arndt, Marcel; Griebel, Michael</td>
<td>Derivation of Higher Order Gradient Continuum Models from Atomistic Models for Crystalline Solids</td>
<td>Multiscale Modeling and Simulation</td>
</tr>
<tr>
<td>149.</td>
<td>Griebel, Michael; Preusser, Tobias; Rumpf, Martin; Schweitzer, Marc Alexander; Telea, Alexandru</td>
<td>Flow Field Clustering via Algebraic Multigrid</td>
<td>Visualization 2004</td>
</tr>
<tr>
<td>150.</td>
<td>Roessler, Thomas</td>
<td>Discretizing the Porous Medium Equation Based on its Gradient Flow Structure – A Consistency Paradox</td>
<td></td>
</tr>
<tr>
<td>151.</td>
<td>Griebel, Michael; Scherer, Karl; Schweitzer, Marc Alexander</td>
<td>Robust Norm Equivalencies and Optimal Preconditioners for Diffusion Problems</td>
<td>Mathematics of...</td>
</tr>
</tbody>
</table>
155. Griebel, Michael; Jager, Lukas; Voigt, Axel: Computing Diffusion Coefficients of Intrinsic Point Defects by Atomistic Simulations

156. Engel, Martin; Griebel, Michael: Flow Simulation on Moving Boundary-Fitted Grids and Application to Fluid-Structure Interaction Problems

157. Croce, Roberto; Griebel, Michael; Schweitzer, Marc Alexander: A Parallel Level-Set Approach for Two-Phase Flow Problems with Surface Tension in Three Space Dimensions; eingereicht bei: Journal of Computational Physics

158. Bass, Richard F.; Kassmann, Moritz: Hölder Continuity of Harmonic Functions with Respect to Operators of Variable Order


160. Giacomelli, Lorenzo; Knüpfer, Hans; Otto, Felix: Maximal Regularity for a Degenerate Operator for Fourth Order

161. Drwenski, Jörg; Otto, Felix: $H^2$-Matrix Method vs. FFT in Thin-Film Stray-Field Computations