NESTED NEWTON STRATEGIES FOR ENERGY-CORRECTED FINITE ELEMENT METHODS

U. RÜDE∗, C. WALUGA†, AND B. WOHLMUTH‡

Abstract. Energy-corrected finite element methods provide an attractive technique for dealing with elliptic problems in domains with re-entrant corners. Optimal convergence rates in weighted $L^2$-norms can be fully recovered by a local modification of the stiffness matrix at the re-entrant corner, and no pollution effect occurs. Although the existence of optimal correction factors is established, it remains open how to determine these factors in practice. First, we show that asymptotically a unique correction parameter exists and that it can be formally obtained as the limit of level dependent correction parameters which are defined as roots of an energy defect function. Second, we propose three nested Newton-type algorithms using only one Newton step per refinement level and show local or even global convergence to this asymptotic correction parameter.

Key words. corner singularities, energy-corrected finite element methods, optimal convergence rates, pollution effect, re-entrant corners

AMS subject classifications. 65N30, 65N12, 65N15

DOI. 10.1137/130935392

1. Introduction. Elliptic partial differential equations with singular solution components are of special interest in many applications, e.g., in fracture mechanics or in heterogeneous porous media flow. In the presence of singular solution components, standard finite element methods on uniform meshes show poor results compared to the best approximation in the $L^2$-norm. This effect is commonly referred to as pollution and is related to the fact that the solution of the dual problem has, in general, no $H^2$-regularity.

In this work, we consider the Laplace problem in a two-dimensional domain $\Omega$ with a re-entrant corner as a prototype for a systematic approach to pollution effects. We refer the reader to [21] for a numerical study of more general settings, including second order finite elements, eigenvalue problems, and heterogeneous coefficients. It is well known (see, e.g., [17, 18, 22, 26]) that in the presence of corners with interior angle $\pi < \theta < 2\pi$ the solution will, in general, have singular components of type $r^{k\pi/\theta} \sin(k\pi \phi/\theta)$, $k \in \mathbb{N}$, even when the data are smooth. Here, $r$ denotes the distance to the singular corner and $\phi$ the angle. Thus, the singular components of the solution are not smooth, i.e., $s := r^{\pi/\theta} \sin(\pi \phi/\theta) \notin H^\alpha(\Omega)$, $\alpha > 1 + \pi/\theta$. As a consequence, only reduced convergence rates are obtained by standard finite element methods on a sequence of quasi-uniform meshes; i.e., the estimates $\|\nabla (u - R_h u)\|_0 = O(h^{\pi/\theta})$ and $\|u - R_h u\|_0 = O(h^{2\pi/\theta})$ for the Ritz projector $R_h$ are, in general, sharp with respect to powers of the mesh-size $h$. Although the convergence order in the $H^1$-norm is the same as the order of the best approximation, this does not hold for the $L^2$-norm, where a gap of $1 - \pi/\theta$ in the convergence order can be observed [6, 7, 30]. For a

*Submitted to the journal’s Methods and Algorithms for Scientific Computing section September 3, 2013; accepted for publication (in revised form) April 4, 2014; published electronically July 1, 2014.

http://www.siam.org/journals/sisc/36-4/93539.html

†Department of Computer Science 10, University Erlangen-Nuremberg, Cauerstraße 6, D–91058 Erlangen, Germany (ulrich.rude@informatik.uni-erlangen.de).

‡Institute for Numerical Mathematics, Technische Universität München, Boltzmannstrasse 3, D–85748 Garching b. München, Germany (waluga@ma.tum.de, wohlmuth@ma.tum.de).
weighted $L^2$-norm with weight $r^{1-\pi/\theta}$, the gap is even bigger. The nodal interpolant then shows $O(h^2)$ convergence while such a weight does not recover any additional convergence rate for the Ritz projector, i.e., typically a gap of $2(1-\pi/\theta)$ can be seen.

Standard techniques to improve the convergence are to deal with weighted Sobolev space norms in combination with graded meshes [1, 2, 10, 12, 13, 15] and/or enrichment of the finite element space [3, 6, 9, 11, 14, 20, 25, 30]. Alternative techniques are first order systems least squares methods that provide the flexibility to select appropriate weights in the norms; see, e.g., [4, 5, 8, 16, 23]. Most strategies require more than a local modification and aim to improve the finite element approximation at the singularity. However, the quantity of interest is often not a global standard norm but a functional or weighted norm which excludes or relaxes the influence of the neighborhood of the re-entrant corner. Examples are the stress intensity factor that can be evaluated as a line integral at a fixed distance from the re-entrant corner, the eigenvalues, and the flux at some given interface not including the re-entrant corner. To obtain improved error reduction rates for such quantities, an accurate representation of the solution at the re-entrant corner is not required.

This observation is our motivation to focus on energy correction schemes [21, 28, 29] that do not enrich the finite element spaces associated with a sequence of uniformly refined meshes. The basic idea was originally introduced for finite difference schemes in [33]. In the recent contribution [19], a mathematically rigorous analysis for finite element methods is presented, and it was shown that a careful modification of the original Galerkin method can drastically improve the convergence. Namely, for conforming low-order finite element spaces, if $\pi < \theta < 3\pi/2$, then $O(h^2)$ convergence on a sequence of uniformly refined meshes can be observed in a suitable weighted $L^2$-norm, assuming that a sufficiently accurate correction parameter is known. If additionally the initial mesh restricted to elements touching the re-entrant corner is symmetric, the same argument also holds for $3\pi/2 \leq \theta < 2\pi$.

The rest of this paper is organized as follows: First, in section 2, we sketch the energy corrected finite element method and motivate the need for an efficient algorithm to determine its parameter. In section 3, we introduce a level-dependent nonlinear energy defect function and recall that its unique root defines a suitable level-dependent parameter. Based on the properties of the energy defect, we establish global convergence of a Newton algorithm and give reliable stopping criteria. Section 4 is devoted to the proof that the roots of these functions converge, and we provide convergence rates. Additionally, we show that the limit value defines the unique level-independent parameter in energy-corrected finite element methods. We discuss, in section 5, several nested Newton-type strategies to approximate the roots and show that the asymptotic correction parameter can be approximated with almost no extra cost. Numerical results that illustrate the convergence of the different strategies can be found in section 6. Here we also investigate the dependency of the asymptotic correction factor on the angle of the re-entrant corner and the number of attached elements. Additionally, we apply the approach to a domain with several re-entrant corners and more general boundary conditions.

2. Energy-corrected finite element method. In this section, we sketch the idea of energy-corrected finite element methods. For simplicity we restrict ourselves to a simple model problem and to a bounded polygonal domain $\Omega \subset \mathbb{R}^2$ with one re-entrant corner. We consider the numerical solution of the Poisson problem

$$-\Delta u = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial \Omega.$$ (2.1)
The standard bilinear form associated with (2.1) is given by $a(v, w) := \int_{\Omega} \nabla v \cdot \nabla w \, dx$, $v, w \in H^1(\Omega)$, and $(\cdot, \cdot)$ denotes the usual $L^2$-scalar product.

2.1. Definition of the energy-corrected finite element method. To define the energy correction, we introduce

$$a_l(v, w) := \int_{\omega_l} \nabla v \cdot \nabla w \, dx,$$

where $\omega_l \subset B_{k_l h_l}$ is a union of elements in $T_l$. The sequence $T_l$ forms a nested hierarchy of uniformly refined simplicial meshes with mesh-size $h_l$, and $B_{k_l h_l}$ is a ball with radius $k_l h_l$ and center at the re-entrant corner. Asymptotically, it is essential that $k_l$ be fixed and bounded independently of the level $l$, since this guarantees that the number of elements in $\omega_l$ is bounded independently of $l$; see subsection 3.3 for a theoretical discussion. We note that Algorithm 1 of subsection 3.5 gives us for all our numerical examples $k_l = 1$. For given $\gamma \in [0, 0.5]$, we define the parameter-dependent bilinear form by

$$a_{ec}(v, w) := a(v, w) - \gamma a_l(v, w)$$

and note that it depends on the mesh-dependent subdomain $\omega_l$ and also on the scalar parameter $\gamma$. Obviously, we have $0.5 a(v, v) \leq a_{ec}(v, v) \leq a(v, v)$ for $v \in H^1(\Omega)$. The assumption $\gamma \in [0, 0.5]$ guarantees that $a_{ec}(\cdot, \cdot)$ is uniformly coercive on $H^1_0(\Omega)$.

A modified finite element formulation of (2.1) then reads as follows: Find $u_l(\gamma) \in V_l \cap H^1_0(\Omega)$ such that

$$a_{ec}(u_l(\gamma), v) = (f, v), \quad v \in V_l \cap H^1_0(\Omega).$$

Here, $V_l$ stands for the standard conforming piecewise linear finite element space associated with $T_l$. The modification (2.3) does not change the structure of the standard finite element stiffness matrix and changes only a small number of its coefficients. Hence, it is cheap and easy to implement into existing codes, provided that $\gamma_l$ and $\omega_l$ are given. Moreover, fast high performance solvers may profit from using data structures for uniformly refined grids that avoid the logistic overhead of unstructured and adaptive mesh techniques.

Remark 2.1. We note that for $\gamma = 0$ the standard finite element solution is recovered. Recalling that the Poisson equation models the normal displacement of a homogeneous membrane, the effect of the modification with $\gamma \in (0, 0.5]$ can be regarded as a softening of the material in $\omega_l$.

We emphasize that the quality of $u_l(\gamma)$ is determined by the choice of $\gamma$ and $\omega_l$. The choice of $\omega_l$ is motivated by the fact that the modification should change the original stiffness matrix as little as possible and should not deteriorate the convergence order. In [19] it has been shown that such $\omega_l$ and a level-dependent parameter $\gamma$ exist such that no pollution occurs; i.e., there is no gap in the convergence order between the interpolation and energy-corrected finite element approximation. Moreover, second order convergence in a suitably weighted $L^2$-norm can be recovered. A suitable correction parameter $\gamma$ can be defined by $\gamma := \gamma_l$, where $\gamma_l$ is the root of a nonlinear energy defect function that will be introduced in section 3. In all our numerical examples of section 6 and also those discussed in [19, 21], it has been sufficient to choose the correction domain $\omega_l$ as the union of elements adjacent to the re-entrant corner. We now call the modified finite element approach (2.3) the energy-corrected finite element method if $\omega_l$ and the possibly level-dependent parameter $\gamma$ are selected such that optimal order convergence rates can be observed on uniformly refined meshes.
2.2. Numerical example. We start with an illustrating example that shows the performance of an energy-corrected finite element method. Moreover from this example it will become apparent how important the proper selection of the parameter $\gamma$ is. To demonstrate the accuracy of the energy-correction method and to motivate the algorithms proposed in this paper, we consider here a triangulation of a polygonal domain $\Omega$ with a single re-entrant corner at $x_c = (0,0)$ and the interior angle $\theta = 3\pi/2$ (L-shape). We set nonhomogeneous Dirichlet boundary conditions given in polar coordinates as $u = s := r^{2/3}\sin(2\phi/3)$ on $\partial \Omega$ and zero forcing $f = 0$ such that the exact solution $u = s \in H^{1+\pi/\theta-\varepsilon}(\Omega)$ for any $\varepsilon > 0$. Besides the standard norms, we employ the weighted $L^2$-norm ($L^2_\rho$) and the weighted $H^1$-seminorm ($H^1_\rho$) defined by

\[
\|u - u_h\|_{0,\rho} := \|\rho(u - u_h)\|_0, \quad |u - u_h|_{1,\rho} := \|\rho \nabla (u - u_h)\|_0,
\]

with the radial weighting function defined as $\rho := r^{1-\pi/\theta}$. This weight is illustrated in Figure 1 along with the initial triangulation and the solution.

![Initial mesh, weighting function, and solution for the L-shape example.](image)

We next conduct a convergence study for different values of the correction parameter $\gamma$ on a series of uniformly refined meshes. All finite element formulations show almost the same quantitative results in the $H^1$-norms, but there is a significant difference in the performance for the $L^2$-norms. Table 1 presents the errors for the standard finite element method when no correction is used ($\gamma = 0$). As is well known, this shows the pollution effect, i.e., it results in suboptimal convergence rates, in both weighted and standard $L^2$-norms.

<table>
<thead>
<tr>
<th>$l$</th>
<th>$L^2$ error</th>
<th>Rate</th>
<th>$L^2_\rho$ error</th>
<th>Rate</th>
<th>$H^1$ error</th>
<th>Rate</th>
<th>$H^1_\rho$ error</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.704e-2</td>
<td>-</td>
<td>2.362e-2</td>
<td>-</td>
<td>2.506e-1</td>
<td>-</td>
<td>1.625e-1</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>1.434e-2</td>
<td>1.37</td>
<td>8.248e-3</td>
<td>1.52</td>
<td>1.622e-1</td>
<td>0.63</td>
<td>8.677e-2</td>
<td>0.91</td>
</tr>
<tr>
<td>2</td>
<td>5.545e-3</td>
<td>1.37</td>
<td>3.002e-3</td>
<td>1.46</td>
<td>1.039e-1</td>
<td>0.64</td>
<td>4.595e-2</td>
<td>0.92</td>
</tr>
<tr>
<td>3</td>
<td>2.150e-3</td>
<td>1.37</td>
<td>1.129e-3</td>
<td>1.41</td>
<td>6.620e-2</td>
<td>0.65</td>
<td>2.414e-2</td>
<td>0.93</td>
</tr>
<tr>
<td>4</td>
<td>8.370e-4</td>
<td>1.36</td>
<td>4.336e-4</td>
<td>1.38</td>
<td>4.199e-2</td>
<td>0.66</td>
<td>1.259e-2</td>
<td>0.94</td>
</tr>
<tr>
<td>5</td>
<td>3.273e-4</td>
<td>1.35</td>
<td>1.688e-4</td>
<td>1.36</td>
<td>2.657e-2</td>
<td>0.66</td>
<td>6.535e-3</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Next, we guess a correction parameter $\gamma = 0.1$ for which we observe a significant improvement of the solution accuracy in the $L^2$-norms as shown in Table 2. However, here the asymptotic behavior remains suboptimal. This indicates that the guessed correction parameter may be considered good enough for the first two mesh levels $l = 1, 2$ but that it is not sufficiently accurate for higher levels. Extending the results
of [19], we will show in section 3.4 that on each mesh level there exists a suitable interval of energy correction parameters. Asymptotically there additionally exists a unique parameter $\gamma_\infty$ which works for all levels. We will design efficient algorithms to approximate this limit value accurately enough, which in this example is given by $\gamma_\infty = 0.11917674 \ldots$. Using this asymptotically correct parameter for computing the results in Table 3 yields the optimal convergence rates predicted by the theory for both weighted and standard $L^2$-norms.

### Table 2

<table>
<thead>
<tr>
<th>$I$</th>
<th>$L^2$ error</th>
<th>Rate</th>
<th>$L^2_\rho$ error</th>
<th>Rate</th>
<th>$H^1$ error</th>
<th>Rate</th>
<th>$H^1_\rho$ error</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.016e-2</td>
<td></td>
<td>1.784e-2</td>
<td></td>
<td>2.541e-1</td>
<td></td>
<td>1.645e-1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>9.722e-3</td>
<td>1.63</td>
<td>4.664e-3</td>
<td>1.94</td>
<td>1.652e-1</td>
<td>0.62</td>
<td>8.801e-2</td>
<td>0.90</td>
</tr>
<tr>
<td>2</td>
<td>3.140e-3</td>
<td>1.63</td>
<td>1.257e-3</td>
<td>1.89</td>
<td>1.060e-1</td>
<td>0.64</td>
<td>4.653e-2</td>
<td>0.92</td>
</tr>
<tr>
<td>3</td>
<td>1.019e-3</td>
<td>1.62</td>
<td>3.564e-4</td>
<td>1.82</td>
<td>6.754e-2</td>
<td>0.65</td>
<td>2.439e-2</td>
<td>0.93</td>
</tr>
<tr>
<td>4</td>
<td>3.388e-4</td>
<td>1.61</td>
<td>1.086e-4</td>
<td>1.71</td>
<td>4.284e-2</td>
<td>0.66</td>
<td>1.270e-2</td>
<td>0.94</td>
</tr>
<tr>
<td>5</td>
<td>1.108e-4</td>
<td>1.59</td>
<td>3.597e-5</td>
<td>1.59</td>
<td>2.711e-2</td>
<td>0.66</td>
<td>6.577e-3</td>
<td>0.95</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>$I$</th>
<th>$L^2$ error</th>
<th>Rate</th>
<th>$L^2_\rho$ error</th>
<th>Rate</th>
<th>$H^1$ error</th>
<th>Rate</th>
<th>$H^1_\rho$ error</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.911e-2</td>
<td></td>
<td>1.700e-2</td>
<td></td>
<td>2.557e-1</td>
<td></td>
<td>1.659e-1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>9.147e-3</td>
<td>1.67</td>
<td>4.229e-3</td>
<td>2.01</td>
<td>1.666e-1</td>
<td>0.62</td>
<td>8.887e-2</td>
<td>0.90</td>
</tr>
<tr>
<td>2</td>
<td>2.882e-3</td>
<td>1.67</td>
<td>1.055e-3</td>
<td>2.00</td>
<td>1.069e-1</td>
<td>0.64</td>
<td>4.696e-2</td>
<td>0.92</td>
</tr>
<tr>
<td>3</td>
<td>9.084e-4</td>
<td>1.67</td>
<td>2.646e-4</td>
<td>2.00</td>
<td>6.816e-2</td>
<td>0.65</td>
<td>2.459e-2</td>
<td>0.93</td>
</tr>
<tr>
<td>4</td>
<td>2.862e-4</td>
<td>1.67</td>
<td>6.619e-5</td>
<td>2.00</td>
<td>4.323e-2</td>
<td>0.66</td>
<td>1.279e-2</td>
<td>0.94</td>
</tr>
<tr>
<td>5</td>
<td>9.021e-5</td>
<td>1.67</td>
<td>1.656e-5</td>
<td>2.00</td>
<td>2.735e-2</td>
<td>0.66</td>
<td>6.626e-3</td>
<td>0.95</td>
</tr>
</tbody>
</table>

However, the question of how to compute the parameter $\gamma_\infty$ or how to find a suitable level-dependent parameter interval remains open. Unfortunately, there is so far no analytical formula known to determine these quantities, which depend on the angle and on the initial mesh at the re-entrant corner but not on the global mesh and not on the solution. The main contribution of this paper is to develop and analyze Newton-type algorithms for the approximation of $\gamma_\infty$. Moreover we show that the computed approximations on each level define an energy-corrected finite element method. The three algorithms proposed in section 5 each require one step of a Newton iteration per refinement level of the mesh to determine such a correction parameter. They differ in whether the analytic exact energy of the dominating singular function must be known and in how many finite element systems must be solved per step.

### 3. The energy defect function.

For the ease of presentation, we assume again that $\Omega \subset \mathbb{R}^2$ has one re-entrant corner $x_c = (0,0)$, that a part of the positive x-axis starting at the origin is in $\partial \Omega$, and that all $x \in \overline{\Omega}$ can be represented in polar coordinates as $x = (r \cos \phi, r \sin \phi)$ with $\phi \in [0, \theta]$. Let $s := r^{\pi/\theta} \sin(\pi \phi/\theta)$; then $s$ is a solution of the Dirichlet boundary value problem: Find $u$ such that

$$\Delta u = 0 \quad \text{in} \; \Omega, \quad u = s \quad \text{on} \; \partial \Omega. \quad (3.1)$$

Moreover $a(s,v) = 0$ for all $v \in H^1_0(\Omega)$.

In terms of the bilinear form $a_{ec}(\cdot, \cdot)$ given in (2.2), we define a finite element approximation $s_I(\gamma) \in V_I$ such that the inhomogeneous Dirichlet boundary conditions
s_l(\gamma)(p_l) = s(p_l) are satisfied for all vertices of \( T_l \) being on \( \partial \Omega \) and

\begin{equation}
  a_{ec}(s_l(\gamma), v_l) = 0, \quad v_l \in V_l \cap H^1_0(\Omega).
\end{equation}

We recall that for \( \gamma \in [0, 0.5) \), the bilinear form \( a_{ec} (\cdot, \cdot) \) is uniformly elliptic, and thus a unique solution exists. Moreover, we have

\begin{equation}
  \| \nabla (s - s_l(\gamma)) \|_{0, \Omega \setminus B} \leq C h_l, \quad ch_l^{2\pi/\theta} \geq \| \nabla (s - s_l(\gamma)) \|_0^2 \geq 2C_0 h_l^{2\pi/\theta},
\end{equation}

where \( B \) denotes a ball with a fixed and positive diameter and center at the re-entrant corner. The first bound follows from the regularity of \( s \) on \( \Omega \setminus B \), the global \( L^2 \)-estimate \( \| s - s_l(\gamma) \|_0 = O(h_l) \), and Wahlbin-type arguments for the finite element error on subdomains \([32]\). The second equivalence follows from the best approximation properties and the properties of the modified bilinear form \( a_{ec}(\cdot, \cdot) \). We refer the reader to \([19]\) for details.

Now we define on each level \( l \) the energy defect function \( g_l(\gamma) \), for \( \gamma \in [0, 0.5) \), as

\begin{equation}
  g_l(\gamma) := a(s, s) - a_{ec}(s_l(\gamma), s_l(\gamma)).
\end{equation}

The main difference compared to \([19]\) is that we define the energy defect function in terms of the singular function \( s \) and not with respect to a cut-off of \( s \) having homogeneous Dirichlet boundary values on \( \partial \Omega \). Observing that \( s_l(\gamma) \) restricted to \( \partial \Omega \) does not depend on \( \gamma \), we get that \( s'_l(\gamma) \in V_l \cap H^1_0(\Omega) \), where the prime \( ' \) stands for the derivative with respect to \( \gamma \). A straightforward computation then shows that

\begin{equation}
  g'_l(\gamma) = -2a_{ec}(s_l(\gamma), s'_l(\gamma)) + a(s_l(\gamma), s_l(\gamma)) = a(s_l(\gamma), s_l(\gamma)) \geq 0,
\end{equation}

and thus \( g_l(\cdot) \) defined by (3.4) is, in contrast to the definition given in \([19]\), a monotonically increasing function on \([0, 0.5) \). Moreover, it is easy to see that \( s'_l(\gamma) \) satisfies the variational problem

\begin{equation}
  a_{ec}(s'_l(\gamma), v_l) = a(s_l(\gamma), v_l), \quad v_l \in V_l \cap H^1_0(\Omega).
\end{equation}

### 3.1. Algebraic representation of the energy defect function.

For simplicity of notation, we suppress in the algebraic representation the level index \( l \) and indicate the \( \gamma \) dependence by an upper index. We decompose the nontrivial degrees of freedom of the algebraic representation of \( s_l(\gamma) \) into two blocks. Then \( s_l(\gamma) \in V_l \) can be identified with the vector \( s^\gamma \) having the two block components \( s^\gamma_l \) and \( s^\gamma_R \). The values of \( s^\gamma_l \) are the nodal values of \( s_l(\gamma) \) at the inner vertices, and the values of \( s^\gamma_R \) are the nodal values of \( s_l(\gamma) \) at a subset of boundary vertices. More precisely, let \( \theta_p \) be the angle of the polar coordinates of the boundary vertex \( p \); then we include \( p \) in the subset of boundary nodes if and only if \( \theta_p \in (0, \theta) \); see Figure 2.

We assemble the matrices with respect to the reduced space \( \tilde{V}_l := \{ v_l \in V_l; v_l(x) = 0, x \in \partial \Omega_R \} \) with \( \partial \Omega_R := \{ x \in \partial \Omega, x = (r_x, 0) \text{ or } x = (r_x \cos \theta, r_x \sin \theta) \} \), and \( n_l := \dim(\tilde{V}_l) \). By \( A \) we denote the standard stiffness matrix on level \( l \) associated with Neumann boundary conditions, by \( A_D \) the stiffness matrix associated with Dirichlet conditions, and by \( \tilde{B} \) the matrix associated with \( a_{ec}(\cdot, \cdot) \). All three stiffness matrices \( A, A_{D}, \tilde{B} \in \mathbb{R}^{n_l \times n_l} \) have a \( 2 \times 2 \) block structure associated with the degrees of freedom in the two blocks \( I \) and \( R \). In this algebraic notation, the coefficient vector \( s^\gamma \) of the modified finite element solution \( s_l(\gamma) \) satisfies

\begin{equation}
  (A_D - \gamma \tilde{B}) s^\gamma =: \left( \begin{array}{cc}
    A_H & A_{HR} \\
    0 & \text{Id}
  \end{array} \right) \gamma \left( \begin{array}{c}
    B \\
    0
  \end{array} \right) \left( \begin{array}{c}
    s^\gamma_l \\
    s^\gamma_R
  \end{array} \right) = \left( \begin{array}{c}
    0 \\
    s^\gamma_R
  \end{array} \right),
\end{equation}

where \( \text{Id} \) is the identity matrix.
Fig. 2. The entries in the block component I are marked with filled circles and the entries in the block component R are marked with empty squares (N = 4 for k = 1).

where \( s_R \) stands for the vector obtained by nodal interpolation of the singular function \( s \) on part of the boundary. Having (3.7), \( g_l(\gamma) \) defined by (3.4) and \( g'_l(\gamma) \) defined in (3.5) can be expressed equivalently by

\[
\begin{align*}
g_l(\gamma) &= a(s, s) - s_R^\top (A_{II} - A_{IR}(A_{II} - \gamma B)^{-1} A_{IR}) s_R, \\
g'_l(\gamma) &= s_R^\top A_{IR}(A_{II} - \gamma B)^{-1} B(A_{II} - \gamma B)^{-1} A_{IR} s_R,
\end{align*}
\]

with \( A_{RI} := A_{IR}^\top \). We note that the matrix \( B \) is symmetric and positive semidefinite and of low rank \( N \), where \( N \) is the number of interior vertices contained in \( \Omega_l \); see also Figure 2.

The matrix \( A_{II} \) is positive definite, and moreover a straightforward computation shows that

\[
x^\top B x \leq qx^\top A_{II} x,
\]

with \( q = 1 \) for \( k_0 > 1 \) and \( q < 1 \) for \( k_0 = 1 \).

### 3.2. Relation to singular enrichment.

As an alternative to the previous rank-\( N \) modification of the stiffness matrix, one could consider a rank-1 modification having the same local effect. This can be accomplished by first enriching \( \tilde{V}_l \) with one basis function \( \phi_E \) compactly supported in \( \Omega \). More precisely, we require that the support of \( \phi_E \) be bounded away from the outer boundary, \( \text{supp} \phi_E \cap T = \emptyset \) if \( \partial T \cap (\partial \Omega \setminus \partial \Omega_R) \neq \emptyset \). Second, we consider the original variational problem on the enriched space and apply static condensation to obtain

\[
(\begin{pmatrix} A_{II} & A_{IR} \\ 0 & \text{Id} \end{pmatrix} - \gamma \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix}) \begin{pmatrix} s_I^\top \\ s_R^\top \end{pmatrix} = \begin{pmatrix} 0 \\ s_R \end{pmatrix},
\]

where the submatrix \( S \) has the form \( S := A_{IE} A_{EE}^{-1} A_{EI} \) and \( \gamma = 1 \). Here \( A_{EE} := a(\phi_E, \phi_E) \) is a scalar, \( A_{IE} = A_{EI}^\top \), and \( A_{EI} \) is a row vector whose entries are obtained by evaluating \( a(\cdot, \phi_E) \) with the basis functions from the I-block. Now we set \( \gamma \in [0, 1] \) as parameter. For \( \gamma = 0 \) we obtain the standard formulation and for \( \gamma = 1 \) the enriched form as described above. For \( \gamma \in (0, 1) \) we get by construction a rank-1 modification acting as a local softening of the material. Moreover if \( \phi_E \) is supported in \( \omega_l \), then \( S \) has the same sparsity pattern as \( B \) in (3.7), i.e., it has at most \( N^2 \) non-zero entries. If \( \phi_E \) is a cut-off of the singular function being supported in a subdomain independently of \( h \) and \( \gamma = 1 \), we obtain the condensed form of a singular enrichment. In this case, however, the number of nonzero entries in \( S \) grows as the mesh-size tends to zero.

**Lemma 3.1.** If the modification in its algebraic form is given by (3.10) with \( S = z z^\top \), then the energy defect function has a unique root if \( \beta := z^\top y - \alpha s_R^\top A_{RI} y^\top A_{IR} s_R \neq 0 \).
0. Here $\alpha^{-1} := a(s, s) - a(s_l(0), s_l(0))$ and $y := A_{II}^{-1}z$. Moreover the unique root is then analytically given by $\gamma_l = \frac{\beta}{\omega_l}$.

Proof. The proof follows by an application of the Sherman–Morrison formula in (3.8a) and the fact that $\gamma_l$ is characterized by $g_l(\gamma_l) = 0$. We recall that the Sherman–Morrison formula as a special case of the Sherman–Morrison–Woodbury formula reads as

$$(M + wv^\top)^{-1} = M^{-1} - \frac{M^{-1}wv^\top M^{-1}}{1 + v^\top M^{-1}w}$$

for quite general matrices $M$ and vectors $w, v$. Setting $M := A_{II}$ and $w := -\gamma z$ and $v := z$, we get that $\gamma_l$ is defined by

$$a(s, s) - s_R^\top (A_{RR} - A_{RI} (A_{II}^{-1} - \frac{A_{II}^{-1}(-\gamma z) z^\top A_{II}^{-1}}{1 + z^\top A_{II}^{-1}(-\gamma z)}) A_{IR}) s_R = 0.$$ 

Observing that $a(s_l(0), s_l(0)) = s_R^\top (A_{RR} - A_{RI} A_{II}^{-1} A_{IR}) s_R$, a straightforward computation gives $\gamma_l = \beta^{-1}$.  

Based on Lemma 3.1 the computation of $\gamma_l$ would cost the solution of two unmodified discrete boundary value problems. However, it does not provide the well-posedness and makes no statement about the convergence of such a $\gamma_l$. Thus in the rest of this paper, we do not discuss this option any further but focus on the modification defined by (2.2) and provide efficient algorithms to approximate suitable correction parameters.

Remark 3.2. Equation (3.8a) shows that the nonlinearity in $g_l(\cdot)$ stems from the term $(A_{II} - \gamma B)^{-1}$. Recalling that $B$ is a low rank matrix, one can use a low rank representation of $(\textrm{Id} - \gamma A_{II}^{-1/2} B A_{II}^{-1/2})^{-1}$ using the Sherman–Morrison–Woodbury formula to rewrite $g_l(\cdot)$ as a rational function in $\gamma$, i.e., $g_l(\gamma) = P_l(\gamma)/Q_l(\gamma)$, where $P_l$ and $Q_l$ are polynomials of degree at most $N$. The coefficients of these polynomials can be computed numerically by solving $N$ times a discrete boundary value problem.

In [19] it has been shown that optimal convergence order can be observed if the parameter $\gamma$ on each level $l$ is selected such that $|g_l(\gamma)| \leq C h_l^2$ with $C$ fixed and moderate. In the rest of this section, we first provide existence results and then propose a Newton algorithm including a reliable stopping criteria for the selection of $\omega_l$ and $\gamma$ on each level.

3.3. Existence of $\omega_l$. The properties of $g_l(\cdot)$ depend on the choice of $\omega_l$. To determine a suitable $\omega_l$, we follow along the lines of [19, 29, 28, 33]. Setting $\tilde{\omega}_l^1 := \{T \in T_l, x_c \in \partial T\}$ and then recursively enlarging the neighborhood of $x_c$ by

$$\tilde{\omega}_l^{k+1} := \{T \in T_l, \partial T \cap \tilde{\omega}_l^k \neq \emptyset\},$$

we define

$$\gamma_l^k := \frac{|a(I_l s, I_l s) - a(s, s)|}{\int_{\tilde{\omega}_l^k} \nabla I_l s \cdot \nabla I_l s \, dx}$$

where $I_l$ is the standard nodal interpolation operator. Then it is obvious that $\gamma_l^{k+1} < \gamma_l^k$. Considering the numerator in more detail, we obtain by integration by parts and
from the regularity of $s$ the upper bound
\[ |a(I ls, ls) - a(s, s)| \leq a(Is - s, Is - s) + 2|a(Is - s, s)| \]
\[ \leq \|\nabla(s - Is)\|^2_0 + 2\int_{\partial\Omega} \left| \frac{\partial s}{\partial n}(s - Is) \right| \, d\tau \]
\[ \leq C(h^2_1 + h^2_1) \leq C(h^2_1 + h^2_1), \]
where in the last step we used that $s = Is = 0$ on $\partial\Omega \subset \partial\Omega$. The denominator can be bounded from below in terms of the triangle inequality by
\[ \|\nabla Is\|_{0;\omega^h_1} \geq \|\nabla s\|_{0;\omega^h_1} - \|\nabla(s - Is)\|_{0;\omega^h_1} \]
\[ \geq \|\nabla s\|_{0;\omega^h_1} - \|\nabla(s - Is)\|_0 \geq \left(C_2 k^\pi/\theta - C_1\right)h^\pi/\theta. \]
From now on we fix $\omega_l := \omega^k_0$ with $k_0 \in \mathbb{N}$ large enough such that $(C_2 k_0^\pi/\theta - C_1)^2 > 2C^2_1$. This choice guarantees $\tilde{\gamma}^k_0 < 0.5$.

3.4. Existence of an interval for $\gamma$. We proceed in two steps. In Lemma 3.3, we show that $g_l(\cdot)$ has a unique root in $(0, 0.5)$. In Lemma 3.4, we establish lower and upper bounds for $g_l(\cdot)$. These two preliminary results allow us to specify a closed interval $J_l$ such that for $\gamma \in J_l$, we obtain an energy-corrected finite element method.

Lemma 3.3. There exists a coarse level $l_0 \in \mathbb{N}$ such that for $l \geq l_0$ there is a unique $\gamma_l \in (0, 0.5)$ with $g_l(\gamma_l) = 0$.

Proof. The proof is similar to the proof of [19, Lemma 5.2]; see also [28, Lemma 3]. In a first step, we show existence and in a second step uniqueness. Since $g_l(\cdot)$ is continuous, we start with the evaluation of $g_l(\cdot)$ at $\gamma = 0$ and at $\gamma = 0.5$ and deduce $g_l(0)g_l(0.5) \leq 0$. For $\gamma = 0$, we find in terms of (3.3) and the fact that $\Delta s = 0$
\[ g_l(0) = a(s, s) - a(s(0), s(0)) = -a(s - s(0), s - s(0)) + 2a(s, s - s(0)) \]
\[ \leq -2C_0 h^2_1 + 2\int_{\partial\Omega} \left| \frac{\partial s}{\partial n}(s - Is) \right| \, d\tau \leq -2C_0 h^2_1 + C_3 h^2_1 \]
\[ \leq h^2_1 (C_3 h^2_{1-\pi/\theta} - C_0). \]
Thus for $l_0$ large enough, we have $C_3 h^2_{1-\pi/\theta} \leq C_0$ and
\[ g_l(0) \leq -C_0 h^2_1, \quad l \geq l_0. \]
Noting that $s_l(\gamma) - Is \in V_l \cap H^1_0(\Omega)$ and thus $a_{cc}(s_l(\gamma), s_l(\gamma) - Is) = 0$, we get for $\gamma = 0.5$
\[ g_l(0.5) = a(s, s) - a(Is, Is) + 0.5 a_l(Is, Is) + \|Is - Is(0.5)\|^2_{H^1_{0.5}} \]
\[ \geq (0.5 - \tilde{\gamma}^k_0 a_l(Is, Is) + \|Is - Is(0.5)\|^2_{H^1_{0.5}} > 0, \]
where $\|u_l\|^2_{H^1_{0.5}} := a(u_l, u_l) - a u_l(u_l, v_l)$. Due to the continuity of $g_l(\cdot)$, there exists a root $\gamma_l \in (0, 0.5)$ for $l$ large enough.

To guarantee uniqueness, it is sufficient to sharpen the monotonicity of (3.5) and show $g'_l(\gamma_l) \neq 0$. Assuming that $a_l(s_l(\gamma_l), s_l(\gamma_l)) = 0$, then $s_l(\gamma_l)$ restricted to $\omega_l$ is equal to zero, and thus $s_l(\gamma_l) = s_l(0)$. Now $g_l(\gamma_l) = 0$, and $g_l(\gamma_l)$ is $g_l(0)$ yields a contradiction to $g_l(0) < 0$, and thus $g_l(\cdot)$ has one unique root.

Lemma 3.4. There exist two constants $0 < \alpha \leq \beta < \infty$ and a level $l_0$ such that for $l \geq l_0$ and $\gamma \in (0, 0.5)$
Algorithm 1. Determine $\omega_l$ and calculate $\gamma$ on level $l \geq l_0$.

$k_0 := 1$

while $g_l(0.5) \leq 0$ do

$k_0 \leftarrow k_0 + 1$

end while

$\sigma := 0.01$

$k := 0; \gamma^0 := 0.5$

while $g_l(\gamma^k) > \sigma h_l^2$ do

$\gamma^{k+1} := \gamma^k - \frac{a(s,s) - a_s(s_l(\gamma^k),s_l(\gamma^k))}{a_l(s_l(\gamma^k),s_l(\gamma^k))}$

$k \leftarrow k + 1$

end while

Each Newton step requires the solution of one finite element system (3.2) and some extra $O(1)$ cost evaluations since $s_l(\gamma) - s_l(0.5) \in H^1_0(\Omega)$. The parameter $\sigma$ can be configured for the problem at hand. Too small values may require more Newton
steps, while for larger values, the approximation of \( \gamma_l \) may be too poor. In all our tests, setting \( \sigma = 0.01 \) worked very well.

**Theorem 3.5.** Algorithm 1 terminates and defines an energy-corrected finite element method.

**Proof.** The theoretical results of section 3.3 guarantee the existence of a finite level-independent \( k_0 \) such that \( g_l(0.5) > 0 \), and thus the first while statement terminates.

To show that the second while statement terminates, we first establish that the sequence \( g_l(\gamma^k) \) is strictly decreasing. Since \( g_l''(\gamma) \geq 0 \) (see (3.13)) and \( g_l(0.5) > 0 \), all our \( \gamma^k \) satisfy \( 0 < \gamma_l \leq \gamma^k \leq 0.5 \) and thus \( g_l(\gamma^k) \geq 0 \). A straightforward computation shows in terms of the equivalence (3.12) that

\[
g_l(\gamma^k) \leq \left( 1 - \frac{\alpha}{\beta} \right) g_l(\gamma^{k-1}) \leq \left( 1 - \frac{\alpha}{\beta} \right)^k g_l(0.5).
\]

Since \( g_l(0.5) \leq C h_l^{2\pi/\theta} \), the number of required Newton steps is at most \( O(l) \).

**Remark 3.6.** For all our numerical results \( k_0 \) is equal to one. If \( N + 1 \) is the number of elements in \( \omega_l \), then only \( 3N - 2 \) entries of the stiffness matrix have to be modified. Typically \( N = 2, 3, 4, 5, 6, 7 \), hence less than 20 entries have to be modified; see also Figure 2, where \( N = 4 \).

**Remark 3.7.** A nested Newton variant improves the performance significantly. In particular, if \( g_l(\gamma^0) = O(h_l^2) \), then the number of required Newton steps will be level-independent. This is the case, e.g., for \( \gamma^0 := \gamma_{l-1} \).

4. Convergence of the level-dependent correction factor. The results of the previous sections indicate that on each level we have to solve a nonlinear problem to determine an accurate enough approximation for \( \gamma_l \). This can be done by Algorithm 1, where each step requires the solution of a finite element problem. Thus for higher levels it is quite expensive and makes the scheme unattractive for practical applications. We are then interested in designing nested Newton schemes which can be easily embedded in a full multigrid method and which require only one Newton step per refinement level. To guarantee global convergence for such a scheme we have to provide theoretical results for the energy defect function.

4.1. Properties of the energy defect function. In this subsection, we consider in more detail the properties of \( g_l(\cdot) \) given by (3.4) as a function of \( \gamma \). Recall that, different from [19], we do not work with homogeneous Dirichlet boundary conditions but with a boundary value problem with a homogeneous right-hand side. As we will see, in this case \( g_l(\cdot) \) is strictly increasing.

**Lemma 4.1.** For the constants \( 0 < \alpha \leq \beta < \infty \) and \( l_0 \) of Lemma 3.4 we have for \( l \geq l_0 \) and \( \gamma, \tilde{\gamma} \in [0, 0.5] \)

\[
\begin{align*}
0 \leq g_l''(\gamma) & \leq \frac{2}{1 - \gamma} g_l'(\gamma), \\
g_l''(\tilde{\gamma}) & \leq \frac{4\beta}{\alpha} g_l'(\gamma).
\end{align*}
\]

**Proof.** The proof of (4.1a) is based on the equality in (3.13) and the upper bound (3.9). Denoting the algebraic representation of \( s_l(\gamma) \) with \( ds_l \), then \( ds_l = 0 \). Moreover we have \( g_l''(\gamma) = 2(s_l)\top B d s_l \), and (3.5) yields \( g_l(\gamma) = (s_l)\top B s_l \). In terms
of (3.9), we then get
\[(1 - \gamma)(ds^\gamma_i)^T Bds^\gamma_i \leq (ds^\gamma_i)^T (A_\Omega - \gamma B)ds^\gamma_i = a_{cc}(s^\gamma_i, s^\gamma_i)\]
\[= a_i(s^\gamma_i, s^\gamma_i) = (s^\gamma_i)^T Bds^\gamma_i \leq ((s^\gamma_i)^T Bs^\gamma_i)^{\frac{1}{2}} \left( ((ds^\gamma_i)^T Bds^\gamma_i) \right)^{\frac{1}{2}}\]
and therefore \((1 - \gamma)\left((ds^\gamma_i)^T Bds^\gamma_i\right)^{\frac{1}{2}} \leq ((s^\gamma_i)^T Bs^\gamma_i)^{\frac{1}{2}}\), i.e.,
\[g_i''(\gamma) \leq \frac{2}{1 - \gamma} ((s^\gamma_i)^T Bs^\gamma_i)^{\frac{1}{2}} \left( ((s^\gamma_i)^T Bs^\gamma_i) \right)^{\frac{1}{2}} = \frac{2}{1 - \gamma} \left((s^\gamma_i)^T Bs^\gamma_i\right)^{\frac{1}{2}} = \frac{2}{1 - \gamma} g_i'(\gamma).\]

Finally, the proof of (4.1b) follows directly from (3.12) and (4.1a).

4.2. Convergence of \(\gamma_l\). The choice \(\gamma = \gamma_l\) results in a fairly expensive algorithm if the \(\gamma_l\) are not precomputed. The main theoretical result of this section is to establish convergence of \(\gamma_l\). This observation then allows us to formulate a nested one-step Newton algorithm for the approximation of \(\gamma_l\). To link \(\gamma_l\) with \(\gamma_{l+1}\), we first introduce an auxiliary quantity \(\gamma_i^l\), \(i = 1, 2\), and second relate \(\gamma_i^l\) to \(\gamma_i\).

Let \(\Omega_i \subset \Omega\) be defined by \(\Omega_i := \bar{\omega}_i^1\), \(i = 1, 2, 3\); see Figure 3 for cases \(i = 1\) and \(i = 2\). Associated with \(\Omega_i\), we define discrete solutions \(s_i^l(\gamma) \in V_l\) and \(\tilde{s}_i^l(\gamma) \in V_l\) such that both are given by interpolation on \(\Omega_i\), i.e., \(s_i^l(\gamma)\) \((p_i) = s_i^l(\gamma)\) \((p_i) = s_i^l(\gamma)\) for all vertices of \(\tilde{T}_l\) being in \(\Omega_i\) \(\Omega_i\) and by variational equality on \(\Omega_i\), \(\Omega_i \setminus \Omega_3\), respectively, i.e.,
\[
\begin{align*}
(4.2a) & \quad a(s_i^l(\gamma), v_l) - \gamma a(s_i^l(\gamma), v_l) = 0, \quad v_l \in V_l, \quad v_l|\Omega_i \in H^1_0(\Omega_i), \\
(4.2b) & \quad a(\tilde{s}_i^l(\gamma), v_l) = 0, \quad v_l \in V, \quad \text{supp} \, v_l \subset \bar{\Omega}_i \setminus \Omega_3.
\end{align*}
\]

To obtain a well-defined \(\tilde{s}_i^l(\gamma)\) on \(\Omega\), we set it equal to \(s_i^l(\gamma)\) on \(\Omega_3\).

![Fig. 3. The construction of \(\Omega_i\), \(i = 1, 2\).](image)

Similarly to (3.2), (3.4), and \(\gamma_l\), we define now \(\gamma_i^l\), \(i = 1, 2\), as the unique solution of
\[
(4.3) \quad g_i^l(\gamma) := a|_{\chi}(s, s) - a|_{\chi}(s_i^l(\gamma), s_i^l(\gamma)) + \gamma a_i(s_i^l(\gamma), s_i^l(\gamma)) = 0,
\]
where the discrete solution \(s_i^l(\gamma) \in V_l\) on \(\Omega_i\) is defined by (4.2a). Here we use the notation \(a|_{\chi}(\cdot, \cdot)\) for the bilinear form \(a(\cdot, \cdot)\) with the integral restricted to the subdomain \(\chi \subset \Omega\). We point out that the results of section 3 also apply for the energy defect functions \(g_i^l(\gamma)\), \(i = 1, 2\).

**Lemma 4.2.** Let \(\gamma_i^1\) and \(\gamma_i^2\) be the unique correction parameters on level \(l\) and \(l + 1\) such that (4.3) holds for \(i = 1\) and \(i = 2\), respectively. Then we have
\[\gamma_i^1 = \gamma_i^2.\]
Proof. Without loss of generality, we have assumed that \( x_\gamma \) is the origin of the coordinate system, and thus the linear mapping \( F(x) = 2x \) maps \( \Omega_2 \) onto \( \Omega_1 \). Moreover, for \( x \in \partial \Omega_2 \), we find that \( F(x) \in \partial \Omega_1 \) and \( s \circ F = 2^{\pi/\theta} s \). By construction of \( \Omega_1 \) and \( \Omega_2 \), the boundary nodes on level \( l + 1 \) of domain \( \Omega_2 \) are mapped by \( F \) onto the boundary nodes on level \( l \) of domain \( \Omega_1 \), and \( \omega_{l+1} \) is mapped onto \( \omega_l \). The definition (4.2a) of \( s_l^\gamma(\gamma) \) now yields

\[
s^{2}_{l+1}(\gamma)(x) = 2^{-\pi/\theta} s^{1}_l(\gamma)(2x), \quad x \in \overline{\Omega}_2,
\]

and thus in terms of (4.3), we get \( g^{1}_l(\gamma) = 2^{2\pi/\theta} g^{2}_{l+1}(\gamma) \).

**Lemma 4.3.** Let \( \gamma_l \) and \( \gamma^i_l \), \( i = 1, 2 \), be the correction parameters on level \( l \) such that \( g_l(\gamma_l) = 0 \) and (4.3) hold, respectively. Then we have

\[
|\gamma^i_l - \gamma_l| \leq C h^{2(1-\pi/\theta)}_l.
\]

**Proof.** In a first step, we provide an upper bound for \( \gamma_l - \gamma^i_l \). Using the fact that \( s^i_l(\gamma_l) - s_\gamma(\gamma_l) \in H^1(\Omega) \), we find in terms of (3.2) that

\[
a(s^i_l(\gamma_l), s^i_l(\gamma_l)) - a(s_\gamma(\gamma_l), s_\gamma(\gamma_l)) \geq a(s_1(\gamma_l), s_1(\gamma_l)) - a(s_1(\gamma_l), s_1(\gamma_l)).
\]

Now the definition of \( \gamma^i_l \) yields

\[
\tilde{g}^l_l := a(s_1(\gamma_l), s_1(\gamma_l)) + a(s^i_l(\gamma_l), s^i_l(\gamma_l)) \leq g_l(\gamma_l) = 0.
\]

Decomposing \( \Omega \) into \( \Omega_1 \) and \( \Omega_1^i := \Omega \setminus \overline{\Omega}_2 \), we can rewrite \( \tilde{g}^l_l \) as

\[
\tilde{g}^l_l = a|\Omega_1(s, s) - a|\Omega_1(s^i_l(s, s)^i_l(s_s^i(\gamma_l), s^i_l(\gamma_l))) + a(s^i_l(\gamma_l), s^i_l(\gamma_l)) + \gamma_l a(s_\gamma(\gamma_l), s_\gamma(\gamma_l))
\]

\[
|\Omega_1(s, s) - a|\Omega_1(s^i_l(s, s)^i_l(s_s^i(\gamma_l), s^i_l(\gamma_l))) + a(s^i_l(\gamma_l), s^i_l(\gamma_l)) + \gamma_l a(s_\gamma(\gamma_l), s_\gamma(\gamma_l))
\]

\[
= a|\Omega_1(s, s) - a|\Omega_1(s^i_l(s, s)^i_l(s_s^i(\gamma_l), s^i_l(\gamma_l))) + (\gamma_l - \gamma_l) a(s_\gamma(\gamma_l), s_\gamma(\gamma_l))
\]

\[
= a|\Omega_1(s, s) - a|\Omega_1(s^i_l(s, s)^i_l(s_s^i(\gamma_l), s^i_l(\gamma_l))) + (\gamma_l - \gamma_l) a(s_\gamma(\gamma_l), s_\gamma(\gamma_l))
\]

To obtain an upper bound for \( \gamma_l - \gamma^i_l \), we recall that \( (g^i_l(s^i_l(\gamma_l), s^i_l(\gamma_l))) = a(s^i_l(\gamma_l), s^i_l(\gamma_l)) \) and thus get in terms of Lemma 3.4 that \( a(s^i_l(\gamma_l), s^i_l(\gamma_l)) \geq c h^{2\pi/\theta}_l > 0 \). On the other hand, noting that \( \text{dist}(x, x_c) = \mathcal{O}(1) \) for all \( x \in \Omega_1^i \), we find \( |a|\Omega_1(s, s) - a|\Omega_1(s^i_l(s, s))| \leq C h^2_l \). These two observations yield

\[
(4.4)
\]

\[
\gamma_l - \gamma^i_l \leq C h^{2(1-\pi/\theta)}_l.
\]

The proof of the lower bound for \( \gamma_l - \gamma^i_l \) requires the use of \( s^i_l(\gamma_l) \) and applies similar arguments. We note that \( s^i_l(\gamma_l) - s^i_l(\gamma_l) \) restricted to \( \Omega_i \) is in \( H^1(\Omega_i) \), \( i = 1, 2 \). Then the definition (4.2a) yields that

\[
a|\Omega_1(s^i_l(\gamma_l), s^i_l(\gamma_l)) - a|\Omega_1(s^i_l(\gamma_l), s^i_l(\gamma_l)) + \gamma_l a(s_\gamma(\gamma_l), s_\gamma(\gamma_l)) \geq a(s^i_l(\gamma_l), s^i_l(\gamma_l)) - \gamma_l a(s^i_l(\gamma_l), s^i_l(\gamma_l)).
\]

Now, using the definition of \( \gamma^i_l \) and the fact that \( g_l(\gamma_l) = 0 \), we find in terms of \( s^i_l(\gamma_l) = s_\gamma(\gamma_l) \) on \( \Omega_3 \) and \( s^i_l(\gamma_l) = I_s \) on \( \Omega_1^i \) that

\[
0 = g^{1}_l(\gamma_l) \geq a|\Omega_1(s, s) - a|\Omega_1(s^i_l(\gamma_l), s^i_l(\gamma_l)) + \gamma_l a(s_\gamma(\gamma_l), s_\gamma(\gamma_l))
\]

\[
= a(s, s) - a(s_\gamma(\gamma_l), s_\gamma(\gamma_l)) + \gamma_l a(s_\gamma(\gamma_l), s_\gamma(\gamma_l)) + a|\Omega_1(I_s, I_s) - a|\Omega_1(s, s)
\]

\[
+ a(s_\gamma(\gamma_l), s_\gamma(\gamma_l)) + \gamma_l a(s_\gamma(\gamma_l), s_\gamma(\gamma_l)) + (\gamma_l - \gamma_\gamma) a(s_\gamma(\gamma_l), s_\gamma(\gamma_l))
\]

\[
= a|\Omega_1(I_s, I_s) - a|\Omega_1(s, s) - \|\nabla(s^i_l(\gamma_l) - s_\gamma(\gamma_l))\|^2 + (\gamma_l - \gamma_\gamma) a(s_\gamma(\gamma_l), s_\gamma(\gamma_l)).
\]
Applying the definition (4.2b), we get \[ \| \nabla (\tilde{s}_l^i(\gamma_l) - s_l(\gamma_l)) \|_0^2 = \| \nabla (\tilde{s}_l^i(\gamma_l) - s_l(\gamma_l)) \|_{0;\Omega_l^0}^2. \]
In a next step, we can further decompose \( \Omega_l^3 \) into \( \Omega_l^1 \) and \( \Omega_l \setminus \Omega_l^3 \). On \( \Omega_l \setminus \Omega_l^3 \), both discrete functions \( \tilde{s}_l^i(\gamma_l) \) and \( s_l(\gamma_l) \) are discrete harmonic. Moreover, \( \tilde{s}_l^i(\gamma_l) - s_l(\gamma_l) \) is equal to zero on \( \partial \Omega_3 \). As a consequence, the \( H^1 \)-seminorm on \( \Omega_l \setminus \Omega_3 \) can be bounded by the \( H^{1/2} \)-seminorm on \( \partial \Omega_3 \cap \Omega_l \), which can be bounded by the \( H^1 \)-seminorm on \( \Omega_l^0 \); see, e.g., [31]. Using \( \tilde{s}_l^i(\gamma_l) = Is \) on \( \Omega_l^0 \), we have due to (3.3)
\[ \| \nabla (\tilde{s}_l^i(\gamma_l) - s_l(\gamma_l)) \|_0 \leq C \| \nabla (\tilde{s}_l^i(\gamma_l) - s_l(\gamma_l)) \|_{0;\Omega_l^0} \]
\[ \leq C (\| \nabla (Is - s) \|_{0;\Omega_l^0} + \| \nabla (s - s_l(\gamma_l)) \|_{0;\Omega_l^0}) \leq Ch. \]

Finally, we recall that \( a_l(s_l(\gamma_l), s_l(\gamma_l)) \geq ch^{2\pi/\theta} > 0 \), which implies
\[ \gamma_l^i - \gamma_l \leq Ch^{2(1-\pi/\theta)}. \]
Having the upper bound (4.4) and the lower bound, the distance between \( \gamma_l^i \) and \( \gamma_l \) can be bounded. \( \square \)

**Theorem 4.4.** Let \( \omega_l \) be such that the root \( \gamma_l \in [0, 0.5] \). Then we have for \( \pi < \theta < 2\pi \)
\[ |\gamma_l - \gamma_{l+1}| \leq Ch^{2(1-\pi/\theta)}. \]
Moreover \( \gamma_l \) converges to \( \gamma_\infty \in (0, 0.5] \) with \( |\gamma_l - \gamma_\infty| \leq Ch^{2(1-\pi/\theta)} \) and
\[ |g_l(\gamma_\infty)| \leq Ch^2. \]

Proof. Combining the results of Lemmas 4.2 and 4.3, we get
\[ |\gamma_l - \gamma_{l+1}| = |\gamma_l - \gamma_l^i + \gamma_l^i - \gamma_{l+1}| \leq C \left( h_l^{2(1-\pi/\theta)} + h_{l+1}^{2(1-\pi/\theta)} \right) \leq Ch^{2(1-\pi/\theta)}. \]

Using the triangle inequality and the boundedness of a geometric series, it can be easily seen that \( \gamma_l \) defines a Cauchy sequence and converges with the given rate. To prove the bound for \( g_l(\gamma_\infty) \), we use the fact that there exists a \( \xi_l \in [0, 0.5] \) such that
\[ |g_l(\gamma_\infty)| = |g_l(\gamma_l) + (\gamma_\infty - \gamma_l)g_l'(\xi_l)| \leq Ch_l^{2(1-\pi/\theta)}h_l^{2\pi/\theta}, \]
which follows by Taylor expansion and Lemma 3.4. \( \square \)

**5. Nested Newton algorithms.** Although Algorithm 1 converges globally, it does not exploit the convergence of \( \gamma_l \). In this section, we present several nested Newton strategies for the approximative computation of \( \gamma_\infty \). We assume that we start with a coarse initial mesh and use uniform refinement.

**5.1. A nested Newton iteration using the exact energy: Algorithm 2.**
Our first algorithm is a simple Newton strategy based on the observation that \( \gamma_l \) is the root of \( g_l(\cdot) \). We select \( \gamma_0^a \in [0, 0.5] \) and for \( l = 0, 1, \ldots \) compute

(Algorithm 2) \[ \gamma_{l+1}^a := \min(0.5, max(0, \gamma_{l+1}^a)), \quad \gamma_{l+1}^a := \gamma_l^a - \frac{g_l(\gamma_l^a)}{g_l'(\gamma_l^a)}. \]
Recall that the index \( l \) stands for the refinement level, and on each level only one Newton step is carried out. There is no need for reiteration, and there is almost no extra computational cost, provided that we know \( a(s, s) \).

**Theorem 5.1.** Let us assume that \( \omega_l = \omega_{l+1}^k \) with \( k_0 \) fixed such that \( \gamma_\infty < 0.5 \) if \( k_0 > 1 \) or \( \gamma_\infty \leq 0.5 \) if \( k_0 = 1 \). Then there exists a constant \( C_\alpha < \infty \) such that
\[ |\gamma_{l+1}^a - \gamma_\infty| \leq C(q, \gamma_\infty)|\gamma_l^a - \gamma_\infty|^2 + C_\alpha h_l^{2(1-\pi/\theta)}, \]

\[ |\gamma_{l+1}^a - \gamma_\infty| \leq C(q, \gamma_\infty)|\gamma_l^a - \gamma_\infty|^2 + C_\alpha h_l^{2(1-\pi/\theta)}, \]
with \( C(g, \gamma_\infty) := \frac{q}{1 - q \gamma_\infty} \) and \( q \) as in (3.9). Thus we have convergence for all \( \gamma_\infty^0 \in [0, 0.5] \).

**Proof.** Let \( C_1 := A_{11} - \gamma_1 B \) and \( C_2 := A_{11} - \gamma_2 B \); then we have for \( \gamma_1, \gamma_2 \in [0, 0.5] \) the relation

\[
C_1^{-1} = C_2^{-1} + (\gamma_1 - \gamma_2)C_2^{-1} BC_2^{-1} (\text{Id} - (\gamma_1 - \gamma_2)B) - 1 B C_2^{-1}
\]

\[
= C_2^{-1} + (\gamma_1 - \gamma_2)C_2^{-1} BC_2^{-1}
\]

\[
+ (\gamma_1 - \gamma_2)^2 C_2^{-1} BC_2^{-1} \left( \sum_{i=0}^{\infty} (\gamma_1 - \gamma_2)^i (C_2^{-\frac{1}{2}} BC_2^{-\frac{1}{2}})^i \right) C_2^{-\frac{1}{2}} BC_2^{-1}
\]

\[
= C_2^{-1} + (\gamma_1 - \gamma_2)C_2^{-1} BC_2^{-1}
\]

\[
+ (\gamma_1 - \gamma_2)^2 C_2^{-1} BC_2^{-1} \left( \text{Id} - (\gamma_1 - \gamma_2)(C_2^{-\frac{1}{2}} BC_2^{-\frac{1}{2}})^{-1} C_2^{-\frac{1}{2}} BC_2^{-1} \right)
\]

\[
= C_2^{-1} + (\gamma_1 - \gamma_2)^2 C_2^{-1} BC_2^{-1} + (\gamma_1 - \gamma_2)^2 C_2^{-1} B (C_2 - (\gamma_1 - \gamma_2)B)^{-1} BC_2^{-1}
\]

\[
= C_2^{-1} + (\gamma_1 - \gamma_2)^2 C_2^{-1} BC_2^{-1} + (\gamma_1 - \gamma_2)^2 C_2^{-1} BC_2^{-1}.
\]

In terms of this elementary equality and by means of (3.8a) and (3.8b), we obtain

\[
g(\gamma_1) = g(\gamma_2) + (\gamma_1 - \gamma_2)g'(\gamma_2) + (\gamma_1 - \gamma_2)^2 g''(\gamma_2) s_R A_{11} C_2^{-1} BC_2^{-1} A_{HR} s_R,
\]

\[
g(\gamma_1) \leq g(\gamma_2) + (\gamma_1 - \gamma_2)g'(\gamma_2) + (\gamma_1 - \gamma_2)^2 g''(\gamma_2) s_R A_{11} C_2^{-1} BC_2^{-1} A_{HR} s_R.
\]

In contrast to a standard Taylor expansion, the quadratic term in \( (\gamma_2 - \gamma_1)^2 \) is weighted by the first derivative and not by a second one. Setting \( \gamma_1 = \gamma_\infty \), \( \gamma_2 = \gamma_l^a \) and using (3.8b), we find

\[
\frac{g(\gamma_\infty) - g(\gamma^a_l)}{g'(\gamma^a_l)} = (\gamma_\infty - \gamma_l^a) + (\gamma_\infty - \gamma^a_l)^2 s_R A_{11} C_2^{-1} BC_2^{-1} A_{HR} s_R.
\]

The definition of \( \gamma^a_{\text{trial}} \) and (5.3) result in

\[
|\gamma^a_l - \gamma_\infty| \leq |\gamma^a_{\text{trial}} - \gamma_\infty| = \left| \gamma^a_l - \gamma_\infty - \frac{g(\gamma^a_l) - g(\gamma_\infty)}{g'(\gamma^a_l)} \right|
\]

\[
\leq (\gamma^a_l - \gamma_\infty)^2 \| B_\infty^2 (A_{11} - \gamma_\infty B)^{-1} B_\infty^2 \| + \frac{g(\gamma_\infty)}{g'(\gamma^a_l)}.
\]

Due to (3.9) the first term on the right can be bounded, and Theorem 4.4 and Lemma 3.4 yield a bound for the second term. Altogether we get

\[
|\gamma^a_{l+1} - \gamma_\infty| \leq \frac{q}{1 - q \gamma_\infty} (\gamma^a_l - \gamma_\infty)^2 + C h_l^{-2\pi/\theta}
\]

\[
\leq C(q, \gamma_\infty) \max(\gamma_\infty, 0.5 - \gamma_\infty) |\gamma^a_l - \gamma_\infty| + C h_l^{-2\pi/\theta}.
\]

Under the assumptions on \( \gamma_\infty \) and \( h_l \), we have that the first term on the right is a contraction, and global convergence is obtained. \( \square \)

**5.2. A nested Newton iteration on two levels:** **Algorithm 3.** The main disadvantage of Algorithm 2 is that the determination of \( \gamma^a_{l+1} \) requires the exact evaluation of \( g_l(\gamma^a_l) \), which depends on \( a(s, s) \) and \( s_l(\gamma^a_l) \). The unknown energy \( a(s, s) \) can possibly be evaluated analytically or up to order \( h_l^2 \) accurate by quadrature formulas for \( \int_{\partial\Omega} s \cdot \partial_{\Gamma} s \, ds \) on the edges of \( \partial\Omega \). Our main interest is the formulation of
an algorithm which does not require this evaluation. A first step into this direction is to propose an algorithm which does not require the explicit knowledge of \( a(s, s) \).

To this end we define an alternative characterization of the correction parameter by requiring approximately that the energy defect function on two consecutive levels coincide: \( g_l(\gamma) = g_{l-1}(\gamma) \). We then set \( \gamma_l^b = \gamma_{l-1}^b = 0 \) and define \( \gamma_{l+1}^b := \min(0.5, \max(0, \gamma_{\text{trial}}^b), l = 1, 2, \ldots) \), with

\[
(\text{Algorithm 3}) \quad \gamma_{\text{trial}}^b := \frac{a(s_l(\gamma_l^b), s_l(\gamma_l^b)) - a(s_{l-1}(\gamma_{l-1}^b), s_{l-1}(\gamma_{l-1}^b))}{a_l(s_l(\gamma_{l-1}^b), s_l(\gamma_l^b)) - a_{l-1}(s_{l-1}(\gamma_{l-1}^b), s_{l-1}(\gamma_{l-1}^b))}
\]

Algorithm 3 is motivated by the observation that \( \gamma_{\text{trial}}^b \) can be equivalently written as

\[
(5.4) \quad \gamma_{\text{trial}}^b = \gamma_l^b - \frac{g_l(\gamma_l^b) - g_{l-1}(\gamma_{l-1}^b)}{g_l(\gamma_l^b) - g_{l-1}(\gamma_{l-1}^b)}.
\]

This can be interpreted as one Newton step with start iterate \( \gamma_l^b \) applied for solving \( g_l(\gamma) = g_{l-1}(\gamma) \).

Before we consider the convergence of the sequence \( \gamma_l^b \) given by Algorithm 3, we develop a relation between \( g_l(\gamma) \) and \( g_{l+1}(\gamma) \).

**Lemma 5.2.** The energy defect function on level \( l - 1 \) is related to the energy defect function on level \( l \) by

\[
(5.5a) \quad g_{l-1}(\gamma) = 2^{2\pi/\theta} g_l(\gamma) + \mathcal{O}(h_l^2),
\]

\[
(5.5b) \quad g_{l-1}'(\gamma) = 2^{2\pi/\theta} g_l'(\gamma) + \mathcal{O}(h_l^{1+\pi/\theta}).
\]

**Proof.** We first use the results and notation of section 4 to relate \( g_l(\gamma) \) to \( g_l^b(\gamma) \) and \( g_{l-1}(\gamma) \) to \( g_{l-1}^b(\gamma) \) via

\[
\begin{align*}
g_l(\gamma) &= g_l^b(\gamma) + (a|\Omega|^2(s, s) - a|\Omega|^2(I_l^2, I_l^2)) + \|s_l(\gamma) - s_l^2(\gamma)\|^2_{h_l^\gamma}, \\
g_{l-1}(\gamma) &= g_{l-1}^b(\gamma) + (a|\Omega|^2(s, s) - a|\Omega|^2(I_l^2, I_l^2)) + \|s_{l-1}(\gamma) - s_{l-1}^2(\gamma)\|^2_{h_{l-1}^\gamma},
\end{align*}
\]

where we recall that \( \|v\|^2_{h_l^\gamma} := a(v, v) - a\omega(v, v) \). Using the same arguments as in the proof of Lemma 4.3, we find that the second and third terms on the right are of order \( h_l^2 \). Taking into account the equality \( g_{l-1}^b(\gamma) = 2^{2\pi/\theta} g_l^b(\gamma) \) yields (5.5a).

To obtain a similar relation for the derivatives, we have to consider \( g_l'(\gamma) \), \( g_{l-1}'(\gamma) \) and \( (g_l^b)'(\gamma), (g_{l-1}^b)'(\gamma) \) in more detail. Starting with the trivial equalities

\[
\begin{align*}
g_l'(\gamma) - (g_l^b)'(\gamma) &= a_l(s_l(\gamma), s_l(\gamma)) - a_l(s_l^2(\gamma), s_l^2(\gamma)), \\
g_{l-1}'(\gamma) - (g_{l-1}^b)'(\gamma) &= a_{l-1}(s_{l-1}(\gamma), s_{l-1}(\gamma)) - a_{l-1}(s_{l-1}^2(\gamma), s_{l-1}^2(\gamma)),
\end{align*}
\]

we have to show that the differences on the right are of order \( h_l^{1+\pi/\theta} \). Without loss of generality, we restrict ourselves to the term \( D_l := |a_l(s_l(\gamma), s_l(\gamma)) - a_l(s_l^2(\gamma), s_l^2(\gamma))| \).

Using the definitions of \( s_l(\gamma) \) and \( s_l^2(\gamma) \), we get

\[
D_l \leq a_l(s_l(\gamma) - s_l^2(\gamma), s_l(\gamma) - s_l^2(\gamma)) + 2|a(s_l(\gamma), s_l(\gamma) - s_l^2(\gamma))| \\
\leq Ch_l^2 + 2|a_l(s_l(\gamma), s_l(\gamma) - s_l^2(\gamma))| \leq C \left( h_l^2 + h_l^{1+\pi/\theta} \right) \leq Ch_l^{1+\pi/\theta},
\]

and thus (5.5b) holds.

**Theorem 5.3.** Let us assume that \( \omega_l = \omega_{k_0}^l \) with \( k_0 \) fixed such that \( \gamma_{\infty} < 0.5 \) if \( k_0 > 1 \) or \( \gamma_{\infty} \leq 0.5 \) if \( k_0 = 1 \). Then there exists a constant \( C_\theta < \infty \) such that

\[
|\gamma_{l+1}^b - \gamma_{\infty}| \leq C(q, \gamma_{\infty})|\gamma_{l}^b - \gamma_{\infty}|^2 + C_\theta h_l^{2(1-\pi/\theta)},
\]

with \( C(q, \gamma_{\infty}) \) as in Theorem 5.1. Thus we have convergence for all \( \gamma_{0}^b \in [0, 0.5] \).
Proof. The proof is based on the equivalent representation of $\gamma^b_{\text{trial}}$ by (5.4). We follow along the lines of the proof of Theorem 5.1, use (5.5a) and (5.5b), and get for $h_l$ small enough

$$|\gamma^b_{\text{trial}} - \gamma^\infty| = |\gamma^b_l - \gamma^\infty - \frac{g_l(\gamma^b_l) - g_{l-1}(\gamma^b_l)}{g'_l(\gamma^b_l) - g'_{l-1}(\gamma^b_l)}|$$

$$= |\gamma^b_l - \gamma^\infty + \frac{(2^{2\pi/\theta} - 1)g_l(\gamma^b_l) + \mathcal{O}(h^2_l)}{(2^{2\pi/\theta} - 1)g'_l(\gamma^b_l) + \mathcal{O}(h_1^{1+\pi/\theta})}|$$

$$\leq \left|\gamma^b_l - \gamma^\infty + \frac{g_l(\gamma^b_l)}{g'_l(\gamma^b_l)}\right| + C \left(h^2_l h_1^{1-\pi/\theta} + g_l(\gamma^b_l) h_1^{1+\pi/\theta} h_l^{-2\pi/\theta}\right)$$

$$\leq C(q, \gamma^\infty)(\gamma^b_l - \gamma^\infty)^2 + C h_1^{1-\pi/\theta} \left(h^2_l h_1^{1-\pi/\theta} + h^2 + |\gamma^b_l - \gamma^\infty|h_1^{2\pi/\theta}\right)$$

$$\leq C(q, \gamma^\infty)(\gamma^b_l - \gamma^\infty)^2 + C h_1^{2(1-\pi/\theta)}.$$ 

We note that in the last step, we have used $1 - \pi/\theta \leq 2\pi/\theta$ and that $|\gamma^b_l - \gamma^\infty|$ is bounded. Let (5.6) hold; then the first term on the right is by construction a contraction, and thus global convergence can be observed.  

5.3. An inexact nested Newton iteration: Algorithm 4. Although Algorithm 3 does not require the value $a(s, s)$, we have to solve the finite element equation on both levels, $l$ and $l - 1$, with the given value $\gamma^b_l$. Thus we propose a further simplification of the algorithm where we reuse the results of the previous computations by replacing $\gamma^b_l$ in the finite element approximation on level $l - 1$ by $\gamma^b_{l-1}$. We set $\gamma^b_0 = \gamma^c_0 = 0$ and define $\gamma^c_{l+1} := \min(0.5, \max(0, \gamma^c_{l+1}, \gamma^c_{l+1}))$, $l = 1, 2, \ldots$, with

(Algorithm 4) \[
\gamma^c_{\text{trial}} := \frac{a_l(s_l(\gamma^c_l), s_l(\gamma^c_l)) - a_l(s_{l-1}(\gamma^c_{l-1}), s_{l-1}(\gamma^c_{l-1}))}{a_l(s_l(\gamma^c_l), s_l(\gamma^c_l)) - a_l(s_{l-1}(\gamma^c_{l-1}), s_{l-1}(\gamma^c_{l-1}))}. 
\]

Theorem 5.4. Under the assumptions of Theorem 5.3, there exist a $\tau > 0$ small enough and a level $l_0$ such that for $|\gamma^c_{l_0} - \gamma^\infty|, |\gamma^c_{l_0+1} - \gamma^\infty| \leq \tau$, we have $|\gamma^c_l - \gamma^\infty| \leq \tau, l \geq l_0$, and moreover

\[
|\gamma^c_{l+1} - \gamma^\infty| \leq C_{\text{c}} \left(h^2_1(1-\pi/\theta) + |\gamma^c_l - \gamma^\infty|^2 + |\gamma^c_{l-1} - \gamma^\infty|^2\right), \quad l \geq l_0, 
\]

with $C_{\text{c}} < \infty$, and thus local convergence is guaranteed.

Proof. The proof is technical but essentially follows the arguments of the proofs of Theorems 5.1 and 5.3. Thus we do not work out all details on the constants but only sketch the main aspects. We start by reformulating the definition of Algorithm 4. Using (3.4) and (3.5), we can rewrite $\gamma^c_{\text{trial}}$ in terms of $g_{l-1}(\gamma^c_{l-1}), g_l(\gamma^c_l)$ and $g'_{l-1}(\gamma^c_{l-1}), g'_l(\gamma^c_l)$, obtaining

\[
\gamma^c_{\text{trial}} = \frac{g_{l-1}(\gamma^c_{l-1}) - g_{l-1}(\gamma^c_{l-1}) - g_l(\gamma^c_l) + g_l(\gamma^c_l)}{g'_l(\gamma^c_l) - g'_{l-1}(\gamma^c_{l-1})} \gamma^c_l \gamma^c_{l-1} - \gamma^c_{l-1} \gamma^c_{l-1} = \gamma^c_l + \frac{g_{l-1}(\gamma^c_{l-1}) - g_l(\gamma^c_l) + (\gamma^c_l - \gamma^c_{l-1})g'_{l-1}(\gamma^c_{l-1})}{g'_l(\gamma^c_l) - g'_{l-1}(\gamma^c_{l-1})}.
\]

Using the notation of the proof of Theorem 5.1 on level $l - 1$, and setting $\gamma_1 = \gamma^c_1$, $\gamma_2 = \gamma^c_{l-1}$ in (5.2), we can reformulate the numerator as

\[
\gamma^c_{\text{trial}} = \frac{g_l(\gamma^c_l) - g_{l-1}(\gamma^c_l) + (\gamma^c_l - \gamma^c_{l-1})^2 s_{l-1} A_{\text{HR}} C_{2-1} B_{1-1} B_{1-1} A_{\text{HR}} s_{l-1}}{g'_l(\gamma^c_l) - g'_{l-1}(\gamma^c_{l-1})}.
\]
Taylor expansion of $g_l'(\cdot)$ around $\gamma^c_{-1}$ with a suitable $\xi_l \in [0, 0.5 + \tau]$ yields in terms of (5.5b) that
\[
g_l'(\gamma^c_l) = g_l'(\gamma^c_{-1}) + (\gamma^c_l - \gamma^c_{-1})g''_l(\xi_l)
\]
\[
= 2^{-2\pi/\theta}g_{l-1}(\gamma^c_{-1}) + O(h_{l}^{1+\pi/\theta}) + (\gamma^c_l - \gamma^c_{-1})g''_l(\xi_l).
\]
Now due to (3.12) and (4.1b), we get for a suitable $\sigma < \infty$ independent of $l$
\[
|g_l'(\gamma^c_l) - g_{l-1}(\gamma^c_{-1})| = |(2^{-2\pi/\theta} - 1)g_{l-1}(\gamma^c_{-1}) + O(h_{l}^{1+\pi/\theta}) + (\gamma^c_l - \gamma^c_{-1})g''_l(\xi_l)|
\]
\[
\geq (1 - 2^{-2\pi/\theta} - \sigma h_{l}^{1-\pi/\theta} - \sigma |\gamma^c_l - \gamma^c_{-1}|)g_{l-1}(\gamma^c_{-1}).
\]
For $l \geq l_0$, $l_0$ large enough, and for $|\gamma^c_l - \gamma^c_{-1}| \leq 2\tau$, $\tau$ small enough, we find $|g_l'(\gamma^c_l) - g_{l-1}(\gamma^c_{-1})| \geq \sigma g_{l-1}(\gamma^c_{-1})$ with $\sigma$ depending on $l_0$, $\tau$, and $\theta$ but not on $l \geq l_0$. We recall that for $\gamma^c_l \leq 3/5$, we have $B^{1/2}C_1^{-1}B^{1/2} \leq B^{1/2}(A_{II} - 3/5 B)^{-1}B^{1/2} \leq A_{II}^{-1}(A_{II} - 3/5 A_{II})^{-1}A_{II}^{1/2} = 5/2 I_d$. These preliminary considerations yield by means of the algebraic representation (3.8b)
\[
\frac{s_R^T A_{RI} C_2^{-1} B C_2^{-1} A_{RI}}{|g_l'(\gamma^c_l) - g_{l-1}(\gamma^c_{-1})|} \leq \frac{1}{\sigma} \frac{s_R^T A_{RI} C_2^{-1} B C_2^{-1} A_{RI}}{s_{l-1}(\gamma^c_{-1})}
\]
\[
= \frac{1}{\sigma} \frac{s_R^T A_{RI} C_2^{-1} B C_2^{-1} A_{RI}}{s_{l-1}(\gamma^c_{-1})} \leq \frac{5}{2\sigma}.
\]
Using the triangle inequality, we obtain the upper bound
\[
\frac{\gamma^c_c - \gamma^c_{\infty}}{\gamma^c_l - \gamma^c_{-1}} \leq \frac{g_l'(\gamma^c_l) - g_{l-1}(\gamma^c_{-1})}{g_l'(\gamma^c_l) - g_{l-1}(\gamma^c_{-1})} + \frac{5}{2\sigma} (\gamma^c_l - \gamma^c_{-1})^2.
\]
To further bound the first term on the right, we follow along the lines of the proof of Theorem 5.1 and use $g_l'(\gamma^c_l) - g_{l-1}(\gamma^c_{-1}) = (2^{2\pi/\theta} - 1)g_l'(\gamma^c_l) + O(h_{l}^{1+\pi/\theta})$ as well as $g_l'(\gamma^c_l) - g_{l-1}(\gamma^c_{-1}) = (2^{2\pi/\theta} - 1)g_l'(\gamma^c_l) + O(h_{l}^{1+\pi/\theta}) + g_{l-1}(\gamma^c_l) - g_{l-1}(\gamma^c_{-1})$. For $|\gamma^c_l - \gamma^c_{-1}| \leq 2\tau$ and $\tau$ small enough, we obtain by means of the properties of the energy defect function
\[
\frac{\gamma^c_l - \gamma^c_{\infty}}{g_l'(\gamma^c_l) - g_{l-1}(\gamma^c_{-1})} \leq \gamma^c_l - \gamma^c_{\infty} - \frac{(2^{2\pi/\theta} - 1)g_l'(\gamma^c_l) + O(h_{l}^{1+\pi/\theta})}{(2^{2\pi/\theta} - 1)g_l'(\gamma^c_l) + O(h_{l}^{1+\pi/\theta}) + (\gamma^c_l - \gamma^c_{-1})g''_l(\xi_l)}
\]
\[
\leq \gamma^c_l - \gamma^c_{\infty} - \frac{(2^{2\pi/\theta} - 1)g_l'(\gamma^c_l) + O(h_{l}^{1+\pi/\theta}) + (\gamma^c_l - \gamma^c_{-1})g''_l(\xi_l)}{O(h_{l}^{1+\pi/\theta}) + (\gamma^c_l - \gamma^c_{-1})g''_l(\xi_l) + Ch_{l}^{2(1-\pi/\theta)}}
\]
\[
\leq \gamma^c_l - \gamma^c_{\infty} - \frac{g_l'(\gamma^c_l)}{g_l'(\gamma^c_l)} + C \left( \left| g_l'(\gamma^c_l)(h_{l}^{1+\pi/\theta} + (\gamma^c_l - \gamma^c_{-1}) + h_{l}^{2(1-\pi/\theta)} \right) \right)
\]
\[
\leq C(\gamma^c_l - \gamma^c_{-1})^2 + C \left( \left| g_l'(\gamma^c_l)(h_{l}^{1+\pi/\theta} + (\gamma^c_l - \gamma^c_{-1}) + h_{l}^{2(1-\pi/\theta)} \right) \right).
\]
Finally, (5.8) yields
\[
|\gamma_l+1 - \gamma_{\infty}| \leq C_{\infty} \left( h_{l}^{2(1-\pi/\theta)} + (\gamma^c_l - \gamma^c_{\infty})^2 + (\gamma^c_{-1} - \gamma^c_{\infty})^2 \right).
\]
NESTED NEWTON FOR ENERGY-CORRECTED FINITE ELEMENTS

provided that the assumptions are satisfied, $\tau$ is small enough, and $l_0$ is large enough. Moreover if $l_0$ is large enough, we obtain $|\gamma_{l+1}^c - \gamma_\infty| \leq \tau$, and thus (5.7) follows by induction.

**Remark 5.5.** In contrast to Theorems 5.1 and 5.3, Theorem 5.4 does not guarantee global but only local convergence. Although all three upper bounds (5.1), (5.6), and (5.7) have the same structure, there is one characteristic difference. In (5.1) and (5.6) the constants in front of the quadratic error terms can be more precisely specified, and thus convergence for all admissible start iterates is given. Starting with Algorithm 3 on coarse levels and then switching to Algorithm 4 guarantees global convergence.

**Remark 5.6.** Although Algorithms 3 and 4 can be applied without the explicit knowledge of $a(s, s)$, we still require $s$ to set the boundary conditions of our auxiliary problems. However, in any problem setup where the singular component $s$ causes the dominating error contribution, the same algorithms can be used.

**6. Numerical results.** Unless mentioned otherwise, we consider the problem (3.1) where we know the exact solution that is given by the singular function $u = s$. Starting from coarse meshes $T_0$, we generate a sequence of meshes $T_l$ by uniform midpoint refinements. For our numerical tests, we often consider the symmetric (circular) L-shape geometry and the slit domain with interior angles $\theta = 3\pi/2$ and $\theta = 2\pi$, respectively; cf. Figure 4.

![Fig. 4. Meshes $T_0$ and $T_1$ for $n = 6$ elements attached to the re-entrant corner and interior angles $\theta = 3\pi/2$ and $\theta = 2\pi$, respectively.](image)

All implementations are based on the Python interface of the DOLFIN (v. 1.2.0) finite element environment [24].

**6.1. A comparison of the nested Newton algorithms.** In the following numerical experiments, we compute the errors in weighted $L^2$-norms defined in (2.4). For the standard finite element method without energy correction, we expect a sub-optimal asymptotic convergence rate of $2\pi/\theta$, i.e., $4/3$ for the L-shape and 1 for the slit domain. This behavior can indeed be observed in Tables 4 and 5 for the series of finite element solutions obtained without energy correction; see also [19].

Next, we compare the different variants of the energy corrected method, as proposed in sections 3 and 5, namely, the exact Newton on each level (Algorithm 1), the nested iteration in terms of the exact energy (Algorithm 2), the approach on two levels (Algorithm 3), and the inexact method (Algorithm 4). The errors for subsequently refined meshes of the L-shaped and slit domains are displayed in Tables 4 and 5, respectively. All four algorithms successfully recover the optimal asymptotic convergence rates. Furthermore, in terms of the absolute error, the three nested algorithms reach almost the same or even slightly better results than when using Algorithm 1 to approximate the root $\gamma_l$. The underlined values indicate that with energy correction approximately two levels of refinement can be saved in case of the L-shaped domain, as compared with the unmodified finite element solution. In the case of the slit domain, the error on level 3 with correction is already smaller than the error on level 6 for the uncorrected solution.
Table 4
$L^2_\rho$ errors obtained with uncorrected vs. energy corrected methods for $n = 6$ and $\theta = 3\pi/2$.

<table>
<thead>
<tr>
<th>$l$</th>
<th>Uncorrected</th>
<th>Algorithm 1</th>
<th>Algorithm 2</th>
<th>Algorithm 3</th>
<th>Algorithm 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.362e-2</td>
<td>1.774e-2</td>
<td>1.696e-2</td>
<td>1.706e-2</td>
<td>1.706e-2</td>
</tr>
<tr>
<td>1</td>
<td>8.248e-3</td>
<td>4.458e-3</td>
<td>4.272e-3</td>
<td>4.283e-3</td>
<td>4.105e-3</td>
</tr>
<tr>
<td>2</td>
<td>3.002e-3</td>
<td>1.115e-3</td>
<td>1.068e-3</td>
<td>1.070e-3</td>
<td>1.071e-3</td>
</tr>
<tr>
<td>3</td>
<td>1.129e-3</td>
<td>2.788e-4</td>
<td>2.672e-4</td>
<td>2.678e-4</td>
<td>2.676e-4</td>
</tr>
<tr>
<td>5</td>
<td>1.688e-4</td>
<td>1.743e-5</td>
<td>1.673e-5</td>
<td>1.676e-5</td>
<td>1.676e-5</td>
</tr>
<tr>
<td>6</td>
<td>6.621e-5</td>
<td>4.361e-6</td>
<td>4.186e-6</td>
<td>4.195e-6</td>
<td>4.195e-6</td>
</tr>
</tbody>
</table>

Table 5
$L^2_\rho$ errors obtained with uncorrected vs. energy corrected methods for $n = 6$ and $\theta = 2\pi$.

<table>
<thead>
<tr>
<th>$l$</th>
<th>Uncorrected</th>
<th>Algorithm 1</th>
<th>Algorithm 2</th>
<th>Algorithm 3</th>
<th>Algorithm 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4.245e-2</td>
<td>2.587e-2</td>
<td>2.456e-2</td>
<td>2.472e-2</td>
<td>2.472e-2</td>
</tr>
<tr>
<td>1</td>
<td>1.874e-2</td>
<td>6.397e-3</td>
<td>6.251e-3</td>
<td>6.234e-3</td>
<td>8.819e-3</td>
</tr>
<tr>
<td>2</td>
<td>8.890e-3</td>
<td>1.587e-3</td>
<td>1.561e-3</td>
<td>1.557e-3</td>
<td>1.605e-3</td>
</tr>
<tr>
<td>3</td>
<td>4.347e-3</td>
<td>3.945e-4</td>
<td>3.892e-4</td>
<td>3.887e-4</td>
<td>4.351e-4</td>
</tr>
<tr>
<td>5</td>
<td>1.071e-3</td>
<td>2.444e-5</td>
<td>2.417e-5</td>
<td>2.415e-05</td>
<td>2.415e-5</td>
</tr>
<tr>
<td>6</td>
<td>5.343e-4</td>
<td>6.088e-6</td>
<td>6.024e-6</td>
<td>6.019e-6</td>
<td>6.019e-6</td>
</tr>
</tbody>
</table>

In Figure 5 we plot the evolution of the correction parameters $\gamma_l$ with the refinement levels $l$ for the different algorithms. We point out that all three nested Newton algorithms proposed in section 5 converge with the same order but quantitatively faster to $\gamma_\infty$ than to the mesh-dependent root $\gamma_l$ of the energy defect function. Algorithms 2 and 3 produce almost the same curves, whereas the cheapest method, Algorithm 4, shows oscillations in the preasymptotic range; cf. also Remark 5.5.
6.2. The influence of the domain and the coarse mesh on $\gamma_\infty$. In this subsection, we consider numerically the influence of the interior angle $\theta$ at the re-entrant corner and of the number of elements touching the re-entrant corners on $\gamma_\infty$. All tests are set up with $k_0 = 1$, and thus $\omega_l$ is the union of the $n$ elements $T \in T_l$ such that $x_c$ is a vertex of $T$. The asymptotic parameter $\gamma_\infty^n = \gamma_\infty(\theta)$ is regarded as a function on $\theta \in [\pi, 2\pi]$. Here isosceles triangles with angle $\theta/n$ are used; the effects of the element shapes and asymmetry near the corner are discussed in [19].

To obtain an approximation of $\gamma_\infty$ that is sufficiently accurate for many practical computations and to save an explicit computation, we propose a nonlinear fit which is constructed as follows: For each $n \in \{3, \ldots, 12\}$ and 60 samples of $\theta$ in $[\pi, 2\pi]$, we apply Algorithm 3 to compute parameters $\gamma_l$ on a series of 7 meshes, which are successively refined according to the Bulirsch sequence (i.e., 1, $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{6}$, ... ) in order to reduce the cost per evaluation; cf. Figure 6.

Given this data, and assuming an asymptotic expansion for $\gamma(h) = \gamma_\infty + c_1 h^{2-2\pi/\theta} + o(h^{2-2\pi/\theta})$, we use a Richardson extrapolation of the form

$$\gamma_\infty(\theta, n) \approx \gamma_l + (\gamma_{l+1} - \gamma_l)/(1 - (\frac{h_{l+1}}{h_l})^{2-2\pi/\theta})$$

on the last level to eliminate the dominating error term in the asymptotic expansion and obtain improved values for the correction parameters. Having a closer look at the numerical results obtained by this procedure, we find furthermore that $\gamma_\infty(\theta, n)$ can be accurately approximated by a fit in the form of

$$\gamma_{kl}(\theta, n) = \sigma_{1,n}(\exp(-2(\theta - \pi)) - 1) + \sigma_{2,n}(\theta - \pi).$$

In Table 6, we list the coefficients $\sigma_{1,n}$ and $\sigma_{2,n}$ obtained for Dirichlet- and Neumann-type singularities for typical numbers of elements attached to the singularity.

<table>
<thead>
<tr>
<th>n</th>
<th>$\sigma_{1,n}$</th>
<th>$\sigma_{2,n}$</th>
<th>$\sigma_{1,n}$</th>
<th>$\sigma_{2,n}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.0998183980437</td>
<td>0.1896155427030</td>
<td>0.1033975735530</td>
<td>0.1917937440240</td>
</tr>
<tr>
<td>4</td>
<td>0.0555624193920</td>
<td>0.1280415576990</td>
<td>0.0559892102051</td>
<td>0.128230329260</td>
</tr>
<tr>
<td>5</td>
<td>0.0415019850858</td>
<td>0.1072128902000</td>
<td>0.042530924472</td>
<td>0.1076059179200</td>
</tr>
<tr>
<td>6</td>
<td>0.0363481781425</td>
<td>0.0979881012145</td>
<td>0.0369403570671</td>
<td>0.0981918095083</td>
</tr>
<tr>
<td>7</td>
<td>0.0328888599638</td>
<td>0.092597177024</td>
<td>0.0319701432127</td>
<td>0.0922167915794</td>
</tr>
<tr>
<td>8</td>
<td>0.031092655216</td>
<td>0.089440510842</td>
<td>0.0303097852573</td>
<td>0.0890798605368</td>
</tr>
<tr>
<td>9</td>
<td>0.0304135897967</td>
<td>0.0874557266743</td>
<td>0.0296697187321</td>
<td>0.0871982940151</td>
</tr>
<tr>
<td>10</td>
<td>0.0289942470411</td>
<td>0.0857622163158</td>
<td>0.0290198736762</td>
<td>0.0858256530158</td>
</tr>
<tr>
<td>11</td>
<td>0.0279010673090</td>
<td>0.0844853266090</td>
<td>0.0273173027417</td>
<td>0.0842839957719</td>
</tr>
<tr>
<td>12</td>
<td>0.0279439846719</td>
<td>0.0838604929991</td>
<td>0.0266036813107</td>
<td>0.0834159415925</td>
</tr>
</tbody>
</table>

Fig. 6. Series of meshes refined according to the Bulirsch sequence for $n = 6$ and $\theta = 7\pi/4$. 

In Table 6, we list the coefficients $\sigma_{1,n}$ and $\sigma_{2,n}$ obtained from the nonlinear fit.
Numerical tests indicate that the straightforward implementation of this fit of the correction parameters in a finite element code already significantly improves the solution in the presence of corner singularities; cf., e.g., [21] for an application to eigenvalue problems in nonconvex domains. Moreover, in convergence studies with randomly chosen values of \( n \) and \( \theta \), we have always observed a substantially better solution compared to finite elements without energy correction. In addition, the fit also provides a good initial guess for the previously discussed nested Newton algorithms.

Figure 7 illustrates the influence of \( n \) and \( \theta \) on \( \gamma_\infty(\theta, n) \). As the right plot indicates, the correction parameter assumes a maximal value at \( \gamma_\infty(2\pi, 3) = 1/2 \), and in the numerical algorithms this value is obtained sharply. Whether this is mere chance or not needs to be further investigated. Moreover we see that for \( n \to \infty \), \( \gamma_\infty(\theta, n) \) converges and that for \( \theta \in [3\pi/2, 2\pi] \), we obtain almost linear dependence for all \( n \).

6.3. An example with idealized cracks. Finally, we consider the domain \( \overline{\Omega} = [-2, 2] \times [-1, 1] \) that has seven idealized cracks, i.e., re-entrant corners with \( \theta = 2\pi \), as depicted in Figure 8 (left). The problem is defined as

\[
-\Delta u = 0 \quad \text{in } \Omega,
\]

\[
u = \frac{1}{4}\cos(\pi x_2) + 1 \quad \text{on } \Gamma_1 := \{-2\} \times (-1, 1),
\]

\[
\frac{1}{4}\cos(\pi x_2) \quad \text{on } \Gamma_2 := \{2\} \times (-1, 1),
\]

\[
\nabla u \cdot \mathbf{n} = 0 \quad \text{on } \partial \Omega \setminus (\Gamma_1 \cup \Gamma_2).
\]

Such problems may arise, for example, in heat conduction studies in materials with cracks.

Given a triangulation of the domain, our implementation automatically detects all corners and extracts the patch of elements attached to the corner. This time we use the cheaper method, Algorithm 4, to compute a series of correction parameters on meshes, which are again successively refined according to the Bulirsch sequence. Using the Richardson extrapolation (6.1), we obtain sufficiently accurate values for the correction parameters \( \gamma_\infty = 0.28033007 \ldots \) at the respective corners.

To measure the errors in this example, we use a weighted \( L^2 \)-norm

\[
\|u - u_h\|_{0, \rho} = \|\rho(u - u_h)\|_0.
\]
with the weighting function \( \rho \) defined as \( \rho := \min_i \{ 2 |x - x_i|^{1/2} \} \), and where \( x_i \) denote the positions of the crack-tips, respectively. Figure 8 (left) illustrates the geometry and the weighting function. In addition to the weighted \( L_2 \) error, we tabulate the discretization error of the flux at the right boundary in the norm

\[
\| h^{1/2} \nabla (\Pi_l u - u_l) \cdot n \|_{L_2(\Gamma_2)},
\]

where \( \Pi_l \) denotes the \( L_2 \) projection into the discrete function space \( V_l \) in which we compute \( u_l \). Since we have no analytic solution \( u \) for this problem available, it is approximated by the corrected method with mesh level 7. The convergence of the uncorrected vs. corrected finite element method is then examined on a series of uniformly refined meshes. The results are shown in Tables 7 and 8. Additionally, the solution on level 3 is depicted in Figure 8 (right).

### Table 7
**Numerical results for standard finite elements in the domain with cracks.**

<table>
<thead>
<tr>
<th>( t )</th>
<th>( L^2 ) error</th>
<th>Rate</th>
<th>( L_2^\rho ) error</th>
<th>Rate</th>
<th>Flux error</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.176e-2</td>
<td>-</td>
<td>1.933e-2</td>
<td>-</td>
<td>1.463e-2</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>1.044e-2</td>
<td>1.06</td>
<td>9.013e-3</td>
<td>1.10</td>
<td>4.385e-3</td>
<td>1.74</td>
</tr>
<tr>
<td>2</td>
<td>5.112e-3</td>
<td>1.03</td>
<td>4.379e-3</td>
<td>1.04</td>
<td>1.414e-3</td>
<td>1.63</td>
</tr>
<tr>
<td>3</td>
<td>2.525e-3</td>
<td>1.02</td>
<td>2.163e-3</td>
<td>1.02</td>
<td>4.786e-4</td>
<td>1.56</td>
</tr>
<tr>
<td>4</td>
<td>1.254e-3</td>
<td>1.01</td>
<td>1.076e-3</td>
<td>1.01</td>
<td>1.660e-4</td>
<td>1.53</td>
</tr>
<tr>
<td>5</td>
<td>6.246e-4</td>
<td>1.01</td>
<td>5.373e-4</td>
<td>1.00</td>
<td>5.814e-5</td>
<td>1.51</td>
</tr>
</tbody>
</table>

### Table 8
**Numerical results for the energy corrected method in the domain with cracks.**

<table>
<thead>
<tr>
<th>( t )</th>
<th>( L^2 ) error</th>
<th>Rate</th>
<th>( L_2^\rho ) error</th>
<th>Rate</th>
<th>Flux error</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.307e-2</td>
<td>-</td>
<td>1.078e-2</td>
<td>-</td>
<td>5.782e-3</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>4.032e-3</td>
<td>1.70</td>
<td>2.701e-3</td>
<td>2.00</td>
<td>1.253e-3</td>
<td>2.21</td>
</tr>
<tr>
<td>2</td>
<td>1.312e-3</td>
<td>1.62</td>
<td>6.784e-4</td>
<td>1.99</td>
<td>2.465e-4</td>
<td>2.35</td>
</tr>
<tr>
<td>3</td>
<td>4.418e-4</td>
<td>1.57</td>
<td>1.700e-4</td>
<td>2.00</td>
<td>4.598e-5</td>
<td>2.42</td>
</tr>
<tr>
<td>4</td>
<td>1.517e-4</td>
<td>1.54</td>
<td>4.267e-5</td>
<td>1.99</td>
<td>8.192e-6</td>
<td>2.49</td>
</tr>
<tr>
<td>5</td>
<td>5.260e-5</td>
<td>1.53</td>
<td>1.019e-5</td>
<td>2.06</td>
<td>1.194e-6</td>
<td>2.78</td>
</tr>
</tbody>
</table>

We again observe a beneficial effect of the energy correction, since the standard element method exhibits reduced convergence, while optimal orders of convergence are recovered by the correction. Additionally, the numerical accuracy of the flux values can be significantly improved. Here, the discretization error in the uncorrected
solution is already of the theoretically optimal order $O(h^{3/2})$. Surprisingly, the error for the energy-corrected approach yields a superconvergence effect, showing a rate of roughly $O(h^{5/2})$ that cannot be predicted by the refined analysis in [27]. We conducted an additional set of tests with an uncorrected reference solution to ensure that this superconvergence is not due to the comparison with a solution obtained by the energy corrected method. However, the differences between both ways of evaluating errors were negligible for both cases.

7. Conclusions and outlook. In this paper we have developed efficient algorithms and heuristics to determine the correction parameter that is required in the energy correction technique for handling corner singularities in elliptic partial differential equations. The algorithms are based on Newton’s method and can be embedded efficiently in a multilevel context, where just a few or even only a single iteration is required on each refinement level. Furthermore, we have demonstrated numerically that the coefficients can be tabulated and can be approximated by a heuristic function. Additionally, the theoretically shown asymptotic behavior of the correction parameter can be exploited to find improved values by performing Richardson extrapolation on these values. In all these cases, optimal convergence in the $L^2$-norm is recovered away from the singularity.

Future work will deal with extensions of the energy correction technique, such as second order finite elements, the computation of eigenvalues, and strongly heterogeneous materials; see [21] for preliminary results. Finally, important directions for future research include a full integration in a multigrid solver framework and the practical application to mechanical problems.

REFERENCES