



# **Least-Squares Finite-Elemente-Methoden für die Stokes-Gleichungen**

Bachelorarbeit  
Humboldt-Universität zu Berlin  
Mathematisch-Naturwissenschaftliche Fakultät II  
Institut für Mathematik

Eingereicht von Philipp Bringmann

Betreut durch  
Prof. Dr. rer. nat. Carsten Carstensen

**Korrigierte Version**  
Berlin, den 26. November 2013

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Theoretische Grundlagen</b>	<b>4</b>
2.1	Vorbereitung . . . . .	4
2.2	Least-Squares-Formulierung . . . . .	5
2.3	Wohlgestelltheit . . . . .	7
<b>3</b>	<b>Least-Squares Finite-Elemente-Methoden</b>	<b>10</b>
3.1	Diskretisierung . . . . .	10
3.2	LSFEM mit konformer Wahl für die Geschwindigkeit . . . . .	14
3.3	LSFEM mit nichtkonformer Wahl für die Geschwindigkeit . . . . .	15
3.4	Über den Zusammenhang von LS-NCFEM, PSFEM und NCFEM . . . . .	16
3.5	Fehlerschätzer . . . . .	18
3.5.1	Least-Squares-Schätzer . . . . .	18
3.5.2	Drei weitere Fehlerschätzer . . . . .	18
<b>4</b>	<b>Implementierung</b>	<b>20</b>
4.1	Berechnung der Steifigkeitsmatrix . . . . .	20
4.1.1	Block A . . . . .	20
4.1.2	Block B . . . . .	22
4.1.3	Block C . . . . .	24
4.1.4	Block D . . . . .	25
4.1.5	Block F . . . . .	25
4.2	Berechnung der Fehlerschätzer . . . . .	26
4.2.1	Least-Squares-Schätzer . . . . .	26
4.2.2	Berechnung der eta-Schätzer . . . . .	28
4.3	Realisierung . . . . .	29
4.3.1	Lösungsmethoden . . . . .	30
4.3.2	Schätzmethoden . . . . .	32
4.3.3	Exakter Fehler . . . . .	34
<b>5</b>	<b>Numerische Beispiele</b>	<b>35</b>
5.1	„Colliding flow“-Beispiel . . . . .	36
5.2	L-Gebiet-Beispiel . . . . .	36
5.3	Schlitzgebiet-Beispiel . . . . .	42
5.4	Poisson-Problem-Beispiel . . . . .	42
5.5	„backward facing step“-Beispiel . . . . .	47
<b>6</b>	<b>Diskussion der numerischen Ergebnisse</b>	<b>53</b>
6.1	Adaptivität . . . . .	53
6.2	„Locking“-artiges vorasymptotisches Verhalten . . . . .	53

6.3	Eigenschaften der Fehlerschätzer . . . . .	53
6.4	supercloseness . . . . .	54
6.5	Fazit und Ausblick . . . . .	54
<b>Literatur</b>		<b>56</b>
<b>Anhang</b>		<b>58</b>

### **Zusammenfassung**

Der Gegenstand dieser Arbeit ist die Least-Squares Finite-Elemente-Methode zur Lösung der Stokes-Gleichungen für den Fluss von zähen, inkompressiblen Flüssigkeiten. Die Ordnung des Systems partieller Differentialgleichungen reduzieren wir dabei durch Umformulierung unter Verwendung der Pseudospannung. Die Implementierung zweier Lösungsmethoden, eine mit einem konformen Finiten Element für die Geschwindigkeit und eine mit einem nichtkonformen, wird beschrieben und ausgeführt, wobei sich herausstellt, dass die nichtkonforme Least-Squares Methode mit Pseudospannungsmethode bis auf einen Datenfehler übereinstimmt. Die erstellten Programme werden an fünf klassischen Beispielen getestet, wobei sich qualitativ unterschiedliche Ergebnisse beobachten lassen.

# 1 Einleitung

Wie müssen Turbinen im Wasserkraftwerk oder Windräder konstruiert werden, um möglichst viel elektrische Energie zu produzieren? Wie kann erreicht werden, dass hinter einem Flugzeug so wenig störende Luftverwirbelungen wie möglich entstehen? Die Untersuchung von Strömungen ist vor allem in der Technik von immenser Bedeutung. Luft und Wasser zählen zu den newtonschen Gasen und Flüssigkeiten, für die grundsätzlich die gleichen Gesetzmäßigkeiten gelten.

Zur Beschreibung des Strömungsverhaltens dieser Stoffe haben die Mathematiker Claude Navier und George Stokes im 19. Jahrhundert unabhängig voneinander ein System partieller Differentialgleichungen entwickelt, die Navier-Stokes-Gleichungen. Deren Spezialfall für inkompressible Flüssigkeiten mit Viskosität 1 sind die *Stokes-Gleichungen*, mit denen ich mich im Rahmen dieser Arbeit beschäftigen werde.

Auf einem beschränkten Lipschitz-Gebiet  $\Omega \subseteq \mathbb{R}^2$  mit äußerem Normaleneinheitsvektor  $\nu : \partial\Omega \rightarrow \mathbb{R}^2$  ist dabei eine Volumenkraftdichte  $f : \Omega \rightarrow \mathbb{R}^2$  gegeben, d.h. die angreifende Kraft je Volumeneinheit. Für das Geschwindigkeitsfeld  $u : \Omega \rightarrow \mathbb{R}^2$  und den Druck  $p : \Omega \rightarrow \mathbb{R}$  gilt dann auf  $\Omega$  der Zusammenhang

$$-\Delta u + \operatorname{D} p = f. \quad (1.1)$$

Zur eindeutigen Lösbarkeit sind Randdaten notwendig. Wir gehen in dieser Arbeit ausschließlich von reinen Dirichlet-Randdaten  $g : \partial\Omega \rightarrow \mathbb{R}^2$  aus. Die Inkompressibilität der Flüssigkeit hat zur Folge, dass bei der Geschwindigkeit keine Quellen oder Senken auftreten können, sie muss also auf ganz  $\Omega$  divergenzfrei sein,

$$\operatorname{div} u = 0. \quad (1.2)$$

Der Divergenzsatz von Gauß liefert damit folgende notwendige Bedingung an die Randdaten,

$$0 = \int_{\partial\Omega} g \cdot \nu \, ds.$$

Außerdem ist bekannt, dass der Druck nur bis auf eine additive Konstante eindeutig ist. Daher wird er normiert auf

$$\int_{\Omega} p \, dx = 0. \quad (1.3)$$

Für die numerische Lösung der Stokes-Gleichungen sind bereits konforme [Bra03, Taylor-Hood- und MINI-Element in §7], nicht-konforme [Gal11, Bra03, Car09b] und gemischte [CGS12] Finite-Elemente-Methoden vielfältig untersucht worden. Bei diesen Betrachtungen werden die Gleichungen jedoch als Sattelpunktsproblem umformuliert. Dabei müssen die Lösungsräume die LBB-Bedingung (auch inf-sup-Bedingung) erfüllen, um die Existenz einer eindeutigen Lösung zu garantieren.

Der Nachweis dieser Bedingung ist mithin sehr umständlich und nicht immer möglich [Car09b, siehe 4.3.2]. Zur Vermeidung dieser Probleme gibt es einen alternativen Ansatz, bei dem ein gesondertes Minimierungsproblem ohne Nebenbedingung gestellt wird. Dazu wird ein Funktional definiert, das die Normen der Residuen des Problems enthält (üblicherweise die  $L^2$ -Norm). Zur Lösung des Problems kann dann direkt die Galerkin-Methode angewendet werden. Dieser Ansatz nennt sich *Least-Squares-Methode*.

Ad hoc erhalten wir bei diesem Vorgehen einen a posteriori Fehlerschätzer. Sofern das homogene Least-Squares-Funktional äquivalent zur Norm des untersuchten Funktionenraumes ist, ist es für die Abschätzung der Fehler geeignet, da es bereits die Normen der Residuen enthält (vgl. hierzu [BG09, Abschnitt 12.14]).

Ein Nachteil dieser Methode ist die höhere Regularitätsanforderung an die Lösung (vgl. hierzu ebenfalls [BG09, Abschnitt 2.2.2]). Um dies zu umgehen, kann das Problem in ein System erster Ordnung umformuliert werden. Dazu sind Hilfsvariablen nötig, die wiederum zu einem erhöhten Speicherbedarf führen. Ob sich die Least-Squares-Methode dennoch für eine gute numerische Approximation der Lösung eignet, werden wir in dieser Arbeit untersuchen.

Außerdem interessieren uns die Effizienz und Zuverlässigkeit dreier weiterer Fehlerschätzer, die analog zu Termen für das Poisson-Modellproblem aus [CP12] definiert sind. Die Arbeit von Carstensen und Park inspiriert uns darüber hinaus zur Betrachtung der sogenannten „supercloseness“-Eigenschaft, die den Zusammenhang der Lösungen der Least-Squares-Methode mit den Lösungen der Crouzeix-Raviart nichtkonformen sowie der Raviart-Thomas gemischten Finite-Elemente-Methode beschreiben. Dazu vergleichen wir die Größenordnung der beiden Beiträge zum Least-Squares-Funktional, was in den theoretischen Überlegungen in [CP12] die Grundlage für die „supercloseness“-Eigenschaften ist.

Unser Ziel ist es, numerische Beispielrechnungen durchzuführen, die uns Hinweise liefern können, ob diese Eigenschaften gelten. Für unsere Rechnungen benutzen wir das Matlab-Softwarepaket [Car09a]. Wir implementieren zwei Least-Squares-Methoden, wobei wir ein konformes bzw. ein nichtkonformes Element für die Geschwindigkeit wählen. Die Netze werden entweder uniform oder adaptiv verfeinert.

Die Arbeit beginnt in Kapitel section 2 nach einer kurzen Vorbereitung mit einigen theoretischen Überlegungen zur Least-Squares-Methode, wie unter anderem der Umformulierung in ein System erster Ordnung und dem Nachweis der Wohlgestelltheit des Problems. Im darauffolgenden Kapitel section 3 werden das Problem diskretisiert und die genannten Methoden eingeführt. Es folgen Betrachtungen zum Zusammenhang mit anderen bekannten Methoden für das Stokes-Problem und zu den Fehlerschätzern. Die Berechnung der Steifigkeitsmatrix und der Fehlerschätzer geschieht in Kapitel section 4. Zudem wird die Realisierung dieser Berechnungen in den Matlab-Funktionen dieser Arbeit beschrieben. In Kapitel section 5 sind die Ergebnisse der Beispielrechnungen dargestellt. Die Schlussfolgerungen aus diesen

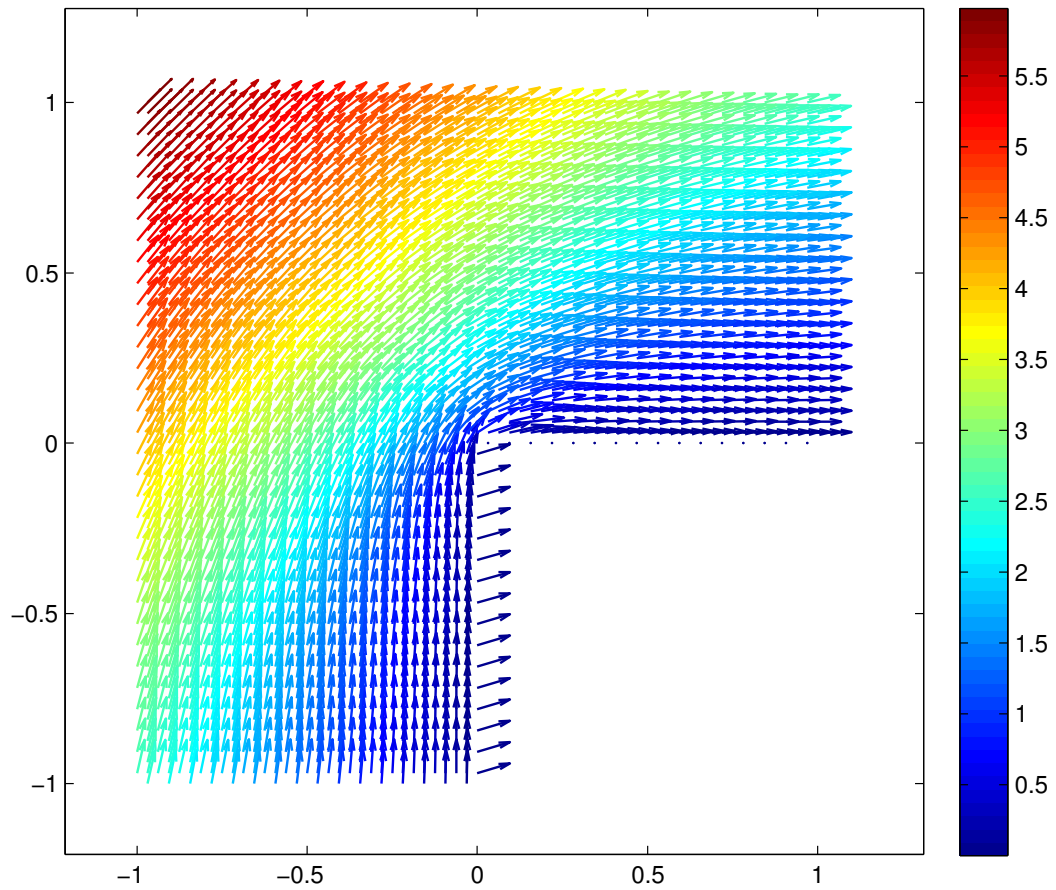


Abbildung 1.1: Zur Veranschaulichung des Untersuchungsgegenstands ein Flussbild der Lösung eines typischen Modellproblems auf dem L-Gebiet mittels der nichtkonformen Least-Squares-Finiten-Elemente-Methode, Rechnung bis 8 000 Freiheitsgrade mit uniformer Netzverfeinerung.

Beobachtung befinden sich im abschließenden Kapitel section 6.

## 2 Theoretische Grundlagen

### 2.1 Vorbereitung

In dieser Arbeit betrachten wir ausschließlich Lipschitz-Gebiete  $\Omega \subseteq \mathbb{R}^2$  mit polygonalem Rand  $\partial\Omega$ . Dazu sei  $\mathcal{T}$  eine reguläre Triangulierung im Sinne von Ciarlet in abgeschlossene Dreiecke  $T \in \mathcal{T}$  (vgl. [Car09b, Abschnitt 1.3.1]), die lokal entsprechend Abbildung 2.1 bezeichnet sind. Die Netzweite der Triangulierung wird allgemein mit  $h$  benannt. Insbesondere sei  $H$  die maximale Netzweite und  $h_T$  die Breite, d.h. die Länge der längste Seite, eines Dreiecks  $T \in \mathcal{T}$ . Letztere ermöglicht die Definition folgender stückweise konstanter Funktion  $h_{\mathcal{T}} \in P_0(\mathcal{T})$  mit  $h_{\mathcal{T}}|_T := h_T$ .

Der Schwerpunkt von  $T \in \mathcal{T}$  sei mit  $\text{mid}(T)$  bezeichnet. Auch dies definiert uns eine stückweise konstante Funktion  $\text{mid}(\mathcal{T}) \in P_0(\mathcal{T}; \mathbb{R}^2)$  durch  $\text{mid}(\mathcal{T})|_T = \text{mid}(T)$ . Dann sei  $\mathcal{N}$  die Menge der Knoten der Triangulierung mit den Teilmengen der Innenknoten  $\mathcal{N}(\Omega)$  und der Randknoten  $\mathcal{N}(\partial\Omega)$ . Analog seien die Mengen  $\mathcal{E}, \mathcal{E}(\Omega)$  und  $\mathcal{E}(\partial\Omega)$  für Kanten definiert. Die Menge der Kanten eines Dreiecks  $T \in \mathcal{T}$  bezeichnen wir mit  $\mathcal{E}(T)$ .  $\nu_T$  und  $\nu_E$  seien die Normaleneinheitsvektoren auf dem Rand von  $T \in \mathcal{T}$  bzw. auf  $E \in \mathcal{E}$ . Die beiden Nachbardreiecke einer Innenkante  $E$  erhalten die Bezeichnung  $T_+$  und  $T_-$  entsprechend der Orientierung von  $\nu_E$ , sodass stets  $\nu_E \cdot \nu_{T_+} = +1$ .  $a \cdot b$  sei das euklidische Skalarprodukt zweier Vektoren  $a, b \in \mathbb{R}^m$ . Für die Verallgemeinerung für Matrizen  $A, B \in \mathbb{R}^{m \times n}$  schreiben wir  $A : B$ . Der Sprung entlang der Innenkante  $E = T_+ \cap T_- \in \mathcal{E}(\Omega)$  sei definiert als  $[\cdot]_E := \cdot|_{T_+} - \cdot|_{T_-}$ .

Das Maß  $|\cdot|$  steht je nach Argument für eine bestimmte Größe, wie den Betrag einer Zahl, die euklidische Länge eines Vektors, die Frobenius-Norm einer Matrix, den Inhalt einer Fläche oder die Länge einer Kante.  $\text{diag}(\cdot)$  bezeichnet je nach Kontext eine Matrix mit ausschließlich den Diagonaleinträgen der Argumentenvektors oder aber einen Spaltenvektor, der die Diagonaleinträge einer quadratischen Matrix enthält.

Für generische, d.h. netzweitenunabhängige, Konstanten  $C$  in einer Ungleichung  $A \leq CB$  schreiben wir  $A \lesssim B$  sowie  $A \approx B$  für  $A \lesssim B \lesssim A$ .

Wir folgen der üblichen Notation für Sobolev-Räume  $H^k(\Omega)$  bzw.  $H_0^k(\Omega)$  und bezeichnen vektor- und tensorwertige Funktionenräume entsprechend mit  $H^k(\Omega; \mathbb{R}^{m \times n})$

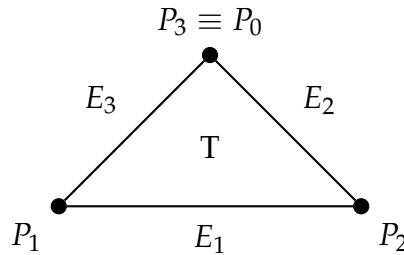


Abbildung 2.1: Lokale Nummerierung in einem Dreieck  $T$  der Triangulierung  $\mathcal{T}$

bzw.  $H_0^k(\Omega; \mathbb{R}^{m \times n})$  für  $m, n \in \mathbb{N}$ . Zu Gunsten einer übersichtlicheren Schreibweise verwenden wir folgende Abkürzungen für die Sobolev-Normen auf einer Menge  $\omega \subseteq \Omega$ ,  $\|f\|_{k,\omega} := \|f\|_{H^k(\omega; \mathbb{R}^{m \times n})}$ ,  $\|f\|_0 := \|f\|_{0,\Omega}$  und für die Energienorm  $\|f\| := \|Df\|_0$ . Außerdem sei

$$H(\operatorname{div}, \Omega) := \left\{ u \in L^2(\Omega; \mathbb{R}^2) \mid \operatorname{div} u \in L^2(\Omega) \right\}$$

versehen mit der Norm  $\|u\|_{H(\operatorname{div}, \Omega)} := \|u\|_0 + \|\operatorname{div} u\|_0$ . Für Tensoren sei

$$H(\operatorname{div}, \Omega; \mathbb{R}^{2 \times 2}) := \left\{ \tau \in L^2(\Omega; \mathbb{R}^{2 \times 2}) \mid \forall j = 1, 2, (\tau_{j1}, \tau_{j2}) \in H(\operatorname{div}, \Omega) \right\}$$

und  $\|\tau\|_{H(\operatorname{div}, \Omega)} := \|\operatorname{div} \tau\|_0 + \|\tau\|_0$ . Die Divergenz wenden wir bei Tensoren stets zeilenweise an.

Das Integral einer vektor- oder tensorwertigen Funktion ist komponentenweise zu verstehen.

Außerdem benötigen wir häufig die folgenden diskreten Räume für  $k \in \mathbb{N}_0$  und  $m, n \in \mathbb{N}$ ,

$$P_k(\mathcal{T}; \mathbb{R}^{m \times n}) := \{p \in L^2(\Omega; \mathbb{R}^{m \times n}) \mid \forall j = 1, \dots, m \forall \ell = 1, \dots, n \forall T \in \mathcal{T}, \\ p_{j\ell}|_T \in P_k(T)\}.$$

Dafür sei  $\Pi_0$  die orthogonale  $L^2$ -Projektion auf  $P_0(\mathcal{T}; \mathbb{R}^{m \times n})$ .

## 2.2 Least-Squares-Formulierung

Die klassische Formulierung der Stokes-Gleichungen haben wir bereits in der Einleitung eingeführt. Zur Vereinfachung gehen wir in diesem theoretischen Abschnitt von homogenen Randdaten  $g \equiv 0$  aus. Aus Praktikabilitätsgründen bei den Least-Squares Finite-Elemente-Methoden [BG09, siehe 2.2.2] sollten Systeme partieller Differentialgleichungen mit möglichst geringer Ordnung vorliegen. Um die Ordnung unseres Systems von zwei auf eins zu verringern, führen wir eine Hilfsvariable  $\sigma := Du - pI_{2 \times 2}$  für die sogenannte Pseudospannung ein.  $I_{2 \times 2} \in \mathbb{R}^{2 \times 2}$  bezeichne die  $2 \times 2$ -Einheitsmatrix. Für unsere Umformulierung benötigen wir den deviatorischen (spurfreen) Anteil eines Tensors.

**Definition 2.1 (Deviatorischer Anteil).** Für einen Tensor  $\tau \in \mathbb{R}^{2 \times 2}$  heißt  $\operatorname{dev} \tau := \tau - \frac{1}{2} \operatorname{tr}(\tau) I_{2 \times 2}$  deviatorischer Anteil von  $\tau$ .

**Lemma 2.2.** Für den deviatorischen Anteil von Tensoren  $\sigma, \tau \in \mathbb{R}^{2 \times 2}$  sind folgende Eigenschaften leicht nachzurechnen,

- (i)  $\operatorname{dev} \sigma : I_{2 \times 2} = \operatorname{tr}(\operatorname{dev} \sigma) = 0$ ,



$$(ii) \quad \operatorname{dev} \sigma : \operatorname{dev} \tau = \operatorname{dev} \sigma : \tau = \sigma : \tau - \frac{1}{2} \operatorname{tr}(\sigma) \operatorname{tr}(\tau).$$

Außerdem fordern wir nur noch eine schwache Differenzierbarkeit und führen deshalb folgende Funktionenräume ein

$$\begin{aligned} H(\operatorname{div}, \Omega; \mathbb{R}^{2 \times 2}) &:= \left\{ \tau \in L^2(\Omega; \mathbb{R}^{2 \times 2}) \mid \forall j = 1, 2, (\tau_{j1}, \tau_{j2}) \in H(\operatorname{div}, \Omega) \right\} \\ \Sigma_0(\Omega) &:= \left\{ \tau \in H(\operatorname{div}, \Omega; \mathbb{R}^{2 \times 2}) \mid \int_{\Omega} \operatorname{tr} \tau \, dx = 0 \right\} \\ V(\Omega) &:= H^1(\Omega; \mathbb{R}^2) \\ V_0(\Omega) &:= H_0^1(\Omega; \mathbb{R}^2). \end{aligned}$$

Damit lautet eine mögliche schwache Umformulierung der Stokes-Gleichungen als System partieller Differentialgleichungen erster Ordnung wie folgt,

**Problem 1 (Pseudospannungs-Geschwindigkeits-Formulierung).** Sei  $f \in L^2(\Omega; \mathbb{R}^2)$  gegeben. Finde  $\sigma \in \Sigma_0(\Omega)$  und  $u \in V_0(\Omega)$ , sodass

$$\operatorname{div} \sigma + f = 0 \quad \text{in } \Omega, \quad (2.1)$$

$$\operatorname{dev} \sigma - D u = 0 \quad \text{in } \Omega. \quad (2.2)$$

Mit den Gleichungen  $\operatorname{tr} D u = \operatorname{div} u = 0$  und  $\operatorname{tr} \sigma = -2p$  lässt sich die Äquivalenz zur klassischen Formulierung leicht zeigen.

Wir definieren nun das Least-Squares-Funktional auf die übliche Weise als Summe der  $L^2$ -Normen der Residuen unseres Problems,

$$\mathcal{LS} : L^2(\Omega; \mathbb{R}^2) \times \Sigma_0(\Omega) \times V(\Omega) \ni (f; \tau, v) \mapsto \|\operatorname{div} \tau + f\|_0^2 + \|\operatorname{dev} \tau - D v\|_0^2 \in \mathbb{R}.$$

Damit erhalten wir die folgende Least-Squares-Problemstellung.

**Problem 2 (Kontinuierliche Least-Squares-Formulierung).** Finde  $\sigma \in \Sigma_0(\Omega)$  und  $u \in V_0(\Omega)$ , sodass

$$\mathcal{LS}(f; \sigma, u) = \min_{(\tau, v) \in \Sigma_0(\Omega) \times V_0(\Omega)} \mathcal{LS}(f; \tau, v). \quad (2.3)$$

### 2.3 Wohlgestelltheit

Wir betrachten nun die zum Minimierungsproblem 2 gehörige Bilinearform

$$\begin{aligned} \mathcal{B} : (\Sigma_0(\Omega) \times V_0(\Omega)) \times (\Sigma_0(\Omega) \times V_0(\Omega)) &\rightarrow \mathbb{R}, \\ ((\sigma, u), (\tau, v)) &\mapsto \int_{\Omega} \operatorname{div} \sigma \cdot \operatorname{div} \tau \, dx + \int_{\Omega} (\operatorname{dev} \sigma - D u) : (\operatorname{dev} \tau - D v) \, dx \in \mathbb{R}. \end{aligned}$$

Die Wohlgestelltheit des Problems folgt nach dem Lax-Milgram-Lemma aus der Stetigkeit und Elliptizität dieser Bilinearform. Bevor wir diese beiden Eigenschaften nachweisen, benötigen wir noch einige Hilfsmittel.

In [BF91, Proposition 3.1] wird eine offensichtliche Modifikation des folgenden Lemmas bewiesen.

**Lemma 2.3 (tr-dev-div-Lemma).** *Es existiert eine Konstante  $C > 0$ , sodass für alle  $\tau \in L^2(\Omega; \mathbb{R}^{2 \times 2})$  mit  $\int_{\Omega} \operatorname{tr} \tau \, dx = 0$  gilt*

$$\|\tau\|_0 \leq C \left( \|\operatorname{dev} \tau\|_0 + \|\operatorname{div} \tau\|_{[H^{-1}(\Omega)]^2} \right).$$

*Bemerkung 2.4.* (i) Brezzi und Fortin zeigen diese Aussage nur für symmetrische Tensoren. Ihr Beweis ist jedoch wörtlich auch auf den allgemeineren Fall anwendbar.

(ii) Sie verwenden außerdem beim zweiten Summanden die  $L^2$ -Norm. Da diese für  $\operatorname{div} \tau$  jedoch nicht existieren muss, habe ich mich entschieden, die Norm des Dualraums  $H^{-1}$  zu verwenden. Für diese Aussage muss im Beweis von Brezzi und Fortin lediglich die Cauchy-Schwarz-Ungleichung durch die Abschätzung mit der Operatornorm ersetzt werden.

**Lemma 2.5 (Friedrichs-Ungleichung für  $H^{-1}$ -Funktionen).** *Für  $f \in L^2(\Omega; \mathbb{R}^2) \subseteq [H^{-1}(\Omega)]^2$  gilt*

$$\|f\|_{[H^{-1}(\Omega)]^2} \lesssim \|f\|_0.$$

*Beweis.* Sei  $f \in L^2(\Omega; \mathbb{R}^2) \subseteq [H^{-1}(\Omega)]^2$  und  $C_F > 0$  die Friedrichs-Konstante für  $H_0^1(\Omega; \mathbb{R}^2)$ -Funktionen. Damit sehen wir,

$$\begin{aligned} \|f\|_{[H^{-1}(\Omega)]^2} &= \sup_{v \in H_0^1(\Omega; \mathbb{R}^2)} \frac{\int_{\Omega} f \cdot v \, dx}{\|D v\|_0} \leq \sup_{v \in H_0^1(\Omega; \mathbb{R}^2)} \|f\|_0 \frac{\|v\|_0}{\|D v\|_0} \\ &\leq \sup_{v \in H_0^1(\Omega; \mathbb{R}^2)} C_F \|f\|_0 \frac{\|D v\|_0}{\|D v\|_0} = C_F \|f\|_0. \end{aligned} \quad \square$$

Außerdem benötigen wir die Youngsche Ungleichung in der folgenden Version.

**Lemma 2.6 (Youngsche Ungleichung).** Seien  $a, b \in \mathbb{R}$ . Dann gilt für alle  $\varepsilon > 0$ ,

$$ab \leq \frac{1}{2\varepsilon}a^2 + \frac{\varepsilon}{2}b^2.$$

**Theorem 2.7 (Stetigkeit und Elliptizität).** Für die obige Bilinearform  $\mathcal{B}$  und für alle  $(\sigma, u), (\tau, v) \in \Sigma_0(\Omega) \times V_0(\Omega)$  gilt,

$$(i) \quad |\mathcal{B}((\sigma, u), (\tau, v))| \lesssim \left( \|\sigma\|_{H(\operatorname{div}, \Omega)} + \|u\| \right) \left( \|\tau\|_{H(\operatorname{div}, \Omega)} + \|v\| \right),$$

$$(ii) \quad \|\tau\|_{H(\operatorname{div}, \Omega)}^2 + \|v\|^2 \lesssim \mathcal{B}((\tau, v), (\tau, v)).$$

*Beweis.* Ad (i): Das folgt unmittelbar aus der Dreiecksungleichung, aus  $\|\operatorname{dev} \sigma\|_0 \lesssim \|\sigma\|_0$  für alle  $\sigma \in \Sigma_0(\Omega)$  sowie der Cauchy-Schwarz-Ungleichung.

Ad (ii): Der Beweis folgt den Ideen von [CLW04], wendet diese jedoch auf die hiesige Situation an.

Wir betrachten zunächst die  $L^2$ -Norm von  $\tau$ . Nach Voraussetzung ist  $\int_{\Omega} \operatorname{tr} \tau \, dx = 0$  und  $\operatorname{div} \tau \in L^2(\Omega; \mathbb{R}^2)$ . Dies erlaubt uns die Anwendung von Lemma 2.3 und 2.5,

$$\|\tau\|_0 \lesssim \|\operatorname{dev} \tau\|_0 + \|\operatorname{div} \tau\|_{[H^{-1}(\Omega)]^2} \lesssim \|\operatorname{dev} \tau\|_0 + \|\operatorname{div} \tau\|_0.$$

Mit der Dreiecksungleichung schätzen wir weiterhin ab,

$$\|\tau\|_0 \lesssim \|Dv - \operatorname{dev} \tau\|_0 + \|Dv\|_0 + \|\operatorname{div} \tau\|_0. \quad (2.4)$$

Da die Norm der Spur eines Tensors durch die Norm des Tensors selbst beschränkt ist, gilt

$$\|\operatorname{div} v\|_0 = \|\operatorname{tr}(Dv - \operatorname{dev} \tau)\|_0 \lesssim \|Dv - \operatorname{dev} \tau\|_0.$$

Damit und mit (2.4) können wir folgendes Produkt abschätzen,

$$\|\operatorname{tr} \tau\|_0 \|\operatorname{div} v\|_0 \lesssim \|Dv - \operatorname{dev} \tau\|_0^2 + \|Dv\|_0 \|Dv - \operatorname{dev} \tau\|_0 + \|\operatorname{div} \tau\|_0 \|Dv - \operatorname{dev} \tau\|_0.$$

In dieser Situation benutzen wir die Youngsche Ungleichung (Lemma 2.6) für die letzten beiden Summanden und erhalten für alle  $\varepsilon > 0$ ,

$$\|\operatorname{tr} \tau\|_0 \|\operatorname{div} v\|_0 \lesssim \left(1 + \frac{1}{\varepsilon}\right) \|Dv - \operatorname{dev} \tau\|_0^2 + \frac{\varepsilon}{2} \|Dv\|_0^2 + \frac{\varepsilon}{2} \|\operatorname{div} \tau\|_0^2. \quad (2.5)$$

Nun wenden wir uns dem Gradienten von  $v$  zu. Zur Abschätzung werden wir die Gleichung  $I_{2 \times 2} : Dv = \operatorname{div} v$ , partielle Integration, die Cauchy-Schwarz- und die Poincaré-Ungleichung benutzen,

$$\begin{aligned} \|Dv\|_0^2 &= \int_{\Omega} (Dv - \operatorname{dev} \tau) : Dv \, dx + \int_{\Omega} \tau : Dv \, dx - \frac{1}{2} \int_{\Omega} \operatorname{tr} \tau \operatorname{div} v \, dx \\ &= \int_{\Omega} (Dv - \operatorname{dev} \tau) : Dv \, dx - \int_{\Omega} \operatorname{div} \tau \cdot v \, dx - \frac{1}{2} \int_{\Omega} \operatorname{tr} \tau \operatorname{div} v \, dx \\ &\lesssim \|Dv - \operatorname{dev} \tau\|_0 \|Dv\|_0 + \|\operatorname{div} \tau\|_0 \|Dv\|_0 + \|\operatorname{tr} \tau\|_0 \|\operatorname{div} v\|_0. \end{aligned}$$

Die Youngsche Ungleichung und die Abschätzung (2.5) zeigen für alle  $\varepsilon > 0$ , dass

$$\|Dv\|_0^2 \lesssim \left(1 + \frac{3}{2\varepsilon}\right) \|Dv - \operatorname{dev} \boldsymbol{\tau}\|_0^2 + \left(\frac{1}{2\varepsilon} + \frac{\varepsilon}{2}\right) \|\operatorname{div} \boldsymbol{\tau}\|_0^2 + \frac{3\varepsilon}{2} \|Dv\|_0^2.$$

Nun können wir ein  $\varepsilon > 0$  so klein wählen, dass der Term  $\|Dv\|_0^2$  auf der rechten Seite absorbiert werden kann, d.h. mit der generischen Konstante  $C > 0$  aus der Abschätzung gilt  $3C\varepsilon/2 < 1$ . Dabei ist zu beachten, dass  $\varepsilon$  dann lediglich von  $C$  abhängt. Deshalb folgt daraus, dass

$$\|Dv\|_0^2 \lesssim \|Dv - \operatorname{dev} \boldsymbol{\tau}\|_0^2 + \|\operatorname{div} \boldsymbol{\tau}\|_0^2. \quad (2.6)$$

Nun können wir die behauptete Abschätzung aus den Ungleichungen (2.4) und (2.6) folgern,

$$\begin{aligned} \|\boldsymbol{\tau}\|_0^2 + \|\operatorname{div} \boldsymbol{\tau}\|_0^2 + \|Dv\|_0^2 &\lesssim \|Dv - \operatorname{dev} \boldsymbol{\tau}\|_0^2 + \|\operatorname{div} \boldsymbol{\tau}\|_0^2 + \|Dv\|_0^2 \\ &\lesssim \|Dv - \operatorname{dev} \boldsymbol{\tau}\|_0^2 + \|\operatorname{div} \boldsymbol{\tau}\|_0^2. \end{aligned} \quad \square$$

### 3 Least-Squares Finite-Elemente-Methoden

#### 3.1 Diskretisierung

Wir wählen zunächst diskrete Räume  $\Sigma_0^h(\mathcal{T}) \subseteq \Sigma_0(\Omega)$  und  $V^h(\mathcal{T})$  nicht notwendigerweise in  $V(\Omega)$ , wobei wir für letzteren mehrere unterschiedliche Wahlen treffen wollen. Wir werden die folgenden Überlegungen nur für einen allgemeinen Raum  $V^h(\mathcal{T})$  anstellen und die Ergebnisse dann für die einzelnen Fälle anpassen. Stets seien

$$\begin{aligned}\Sigma^h(\mathcal{T}) &:= \{\tau_{LS} \in H(\operatorname{div}, \Omega; \mathbb{R}^{2 \times 2}) \mid \forall j = 1, 2, (\tau_{j1}, \tau_{j2}) \in RT_0(\mathcal{T})\}, \\ \Sigma_0^h(\mathcal{T}) &:= \Sigma^h(\mathcal{T}) \cap \Sigma_0(\Omega), \\ V_0^h(\mathcal{T}) &\subseteq V^h(\mathcal{T}) \subseteq P_1(\mathcal{T}),\end{aligned}$$

wobei  $V_0^h(\mathcal{T})$  Nullranddaten in geeigneter Form erfülle. Es bezeichne  $RT_0(\mathcal{T})$  den Raum der Raviart-Thomas-Funktionen geringster Ordnung.

**Definition 3.1 (Raviart-Thomas-Finite-Elemente-Raum).** Für  $T \in \mathcal{T}$  seien folgende diskrete Räume definiert,

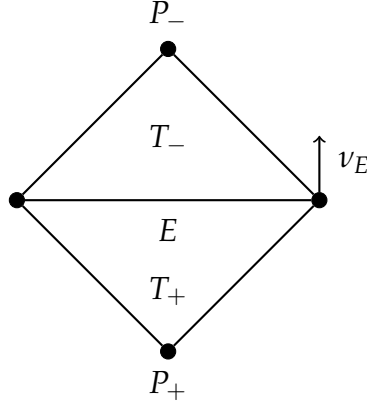
$$\begin{aligned}RT_0(T) &:= \{q \in P_1(T; \mathbb{R}^2) \mid \exists a, b, c \in \mathbb{R} \forall x \in T, q(x) = (a, b) + c(x_1, x_2)\}, \\ RT_0(\mathcal{T}) &:= \{q \in H(\operatorname{div}, \Omega) \mid \forall T \in \mathcal{T}, q|_T \in RT_0(T)\}.\end{aligned}$$

**Problem 3 (Diskrete Least-Squares-Formulierung).** Seien  $f \in L^2(\Omega; \mathbb{R}^2)$  und  $g \in H^1(\Omega; \mathbb{R}^2) \cap H^1(\mathcal{E}(\partial\Omega); \mathbb{R}^2)$  gegeben. Für  $g$  sei  $g_h \in V^h$  eine geeignete Approximation auf dem Rand von  $\Omega$ . Finde  $\sigma_{LS} \in \Sigma_0^h(\mathcal{T})$  und  $u_{LS} \in g_h + V_0^h(\mathcal{T})$ , sodass

$$\mathcal{LS}(f; \sigma_{LS}, u_{LS}) = \min_{(\tau_{LS}, v_{LS}) \in \Sigma_0^h(\mathcal{T}) \times (g_h + V_0^h(\mathcal{T}))} \mathcal{LS}(f; \tau_{LS}, v_{LS}). \quad (3.1)$$

Dies ist ein Minimierungsproblem mit Nebenbedingung  $\int_{\Omega} \operatorname{tr} \sigma_{LS} \, dx = 0$ . Wir erzwingen die Nebenbedingung durch Einführung eines Lagrange-Parameters  $\lambda \in \mathbb{R}$ . Dann ist das Lagrange-Funktional

$$\begin{aligned}\mathcal{L} : L^2(\Omega; \mathbb{R}^2) \times \Sigma^h(\mathcal{T}) \times V_0^h(\mathcal{T}) \times \mathbb{R} &\ni (f; \tau_{LS}, v_{LS}, \lambda) \\ &\mapsto \mathcal{LS}(f; \tau_{LS}, v_{LS}) + \lambda \int_{\Omega} \operatorname{tr} \tau_{LS} \, dx \in \mathbb{R}\end{aligned} \quad (3.2)$$


 Abbildung 3.1: Patch von  $E \in \mathcal{E}(\Omega)$ 

über  $\Sigma^h(\mathcal{T}) \times (g_h + V_0^h(\mathcal{T})) \times \mathbb{R}$  zu minimieren, wobei  $\Sigma^h(\mathcal{T})$  mit der Basis

$$(\Phi_1, \dots, \Phi_{2|\mathcal{E}|}) = \left( \begin{pmatrix} \Psi_1 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} \Psi_{|\mathcal{E}|} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \Psi_1 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \Psi_{|\mathcal{E}|} \end{pmatrix} \right)$$

versehen ist. Dabei setzen wir  $\forall j = 1, \dots, |\mathcal{E}|$ ,  $\Psi_j := \Psi_{E_j}$  für die kantenorientierten Raviart-Thomas-Basisfunktionen  $\Psi_E = (\Psi_{E,1}, \Psi_{E,2})$  aus der folgenden Definition.

**Definition 3.2 (Kantenorientierte Raviart-Thomas-Basisfunktionen).** Es sei eine Innenkante  $E \in \mathcal{E}(\Omega)$  mit gegenüber liegenden Punkten  $P_+$  und  $P_-$  (siehe Abbildung 3.1) und  $\zeta_{E,T_\pm} := \nu_E \cdot (\nu_{T_\pm})|_E$  gegeben. Dann lautet die zugehörige *kantenorientierte Raviart-Thomas-Basisfunktion*

$$\Psi_E(x) := \begin{cases} \zeta_{E,T_\pm} |E| / (2 |T_\pm|) (x - P_\pm) & \text{für } x \in T_\pm, \\ 0 & \text{sonst.} \end{cases} \quad (3.3)$$

Für Randkanten gelte eine analoge Definition mit der Einschränkung, dass der Beitrag auf  $T_-$  wegfalle.

Wir halten zunächst einige Eigenschaften dieser Basisfunktionen fest, die mit Ausnahme von Aussage (iii) in [BC05, Lemma 4.1] bewiesen werden, denn diese Aussage ist ein einfaches Korollar von (i).

**Lemma 3.3.** Auf einem festen Dreieck  $T = \text{conv}\{P, E\} \in \mathcal{T}$  gilt für die kantenorientierten Raviart-Thomas-Basisfunktionen  $\Psi_E$  mit den Baryzentrischen Koordinaten  $\varphi_j$  (siehe Definition 3.4) und mit  $\zeta_E := \zeta_{E,T}$ ,

$$(i) \quad \Psi_E \Big|_T = \zeta_E \frac{|E|}{2|T|} \sum_{\ell=1}^3 \varphi_\ell (P_\ell - P),$$

$$(ii) \quad \operatorname{div} \Psi_E = \zeta_E \frac{|E|}{|T|},$$

$$(iii) \quad \text{für } \alpha = 1, 2, \quad \int_T \Psi_{E,\alpha} \, dx = \zeta_E \frac{|E|}{6} \sum_{\ell=1}^3 (P_{\ell,\alpha} - P_\alpha).$$

Die Funktionen im Raum  $V^h(\mathcal{T})$  besitzen zwei identische Komponenten. Somit ist die Definition der einfachen Dimension  $d := \dim(V^h(\mathcal{T}))/2$  sinnvoll. Die Basis von  $V^h(\mathcal{T})$  sei allgemein bezeichnet mit

$$(\phi_1, \dots, \phi_{2d}) = \left( \begin{pmatrix} \psi_1 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} \psi_d \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \psi_1 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \psi_d \end{pmatrix} \right).$$

Wir verwenden stets folgende Basisdarstellungen,

$$\sigma_{LS} = \sum_{k=1}^{2|\mathcal{E}|} \sigma_k \Phi_k \quad \text{und} \quad u_{LS} = \sum_{m=1}^{2d} u_m \phi_m.$$

Durch Variation in den Komponenten des Lagrange-Funktional (3.2) erhalten wir folgende Formulierung als notwendige Bedingung für ein Minimum.

**Problem 4 (Diskrete Variationsformulierung).** Finde  $\sigma_{LS} \in \Sigma^h(\mathcal{T})$ ,  $u_{LS} \in g_h + V_0^h(\mathcal{T})$  und  $\lambda \in \mathbb{R}$ , sodass

$$\begin{aligned} \forall \tau_{LS} \in \Sigma^h(\mathcal{T}), \quad & \int_{\Omega} \operatorname{div} \sigma_{LS} \cdot \operatorname{div} \tau_{LS} \, dx + \int_{\Omega} (\operatorname{dev} \sigma_{LS} - D u_{LS}) : \operatorname{dev} \tau_{LS} \, dx \\ & + \lambda \int_{\Omega} \operatorname{tr} \tau_{LS} \, dx = - \int_{\Omega} f \cdot \operatorname{div} \tau_{LS} \, dx, \\ \forall v_{LS} \in V_0^h(\mathcal{T}), \quad & \int_{\Omega} (\operatorname{dev} \sigma_{LS} - D u_{LS}) : D v_{LS} \, dx = 0, \\ & \int_{\Omega} \operatorname{tr} \sigma_{LS} \, dx = 0. \end{aligned}$$

Zur allgemeinen Berechnung der Steifigkeitsmatrix setzen wir die Basisdarstellungen in die Gleichungen aus Problem 4 ein und testen fortan lediglich mit den

Basisfunktionen. Dadurch erhalten wir für  $j = 1, \dots, 2|\mathcal{E}|$  und  $\ell = 1, \dots, 2d$ ,

$$\begin{aligned} \sum_{k=1}^{2|\mathcal{E}|} \sigma_k \left( \int_{\Omega} \operatorname{div} \Phi_j \cdot \operatorname{div} \Phi_k \, dx + \int_{\Omega} \operatorname{dev} \Phi_j : \operatorname{dev} \Phi_k \, dx \right) \\ + \sum_{m=1}^{2d} u_m \left( - \int_{\Omega} \operatorname{dev} \Phi_j : D \phi_m \, dx \right) + \lambda \int_{\Omega} \operatorname{tr} \Phi_j \, dx = - \int_{\Omega} f \cdot \operatorname{div} \Phi_j \, dx, \\ \sum_{k=1}^{2|\mathcal{E}|} \sigma_k \left( - \int_{\Omega} \operatorname{dev} \Phi_k : D \phi_{\ell} \, dx \right) + \sum_{m=1}^{2d} u_m \int_{\Omega} D \phi_{\ell} : D \phi_m \, dx = 0, \\ \sum_{k=1}^{2|\mathcal{E}|} \sigma_k \int_{\Omega} \operatorname{tr} \Phi_k \, dx = 0. \end{aligned}$$

Wir setzen also für  $j, k = 1, \dots, 2|\mathcal{E}|$  und  $\ell, m = 1, \dots, 2d$ ,

$$A_{jk} := \int_{\Omega} \operatorname{div} \Phi_j \cdot \operatorname{div} \Phi_k \, dx + \int_{\Omega} \operatorname{dev} \Phi_j : \operatorname{dev} \Phi_k \, dx, \quad (3.4)$$

$$B_{jm} := - \int_{\Omega} \operatorname{dev} \Phi_j : D \phi_m \, dx, \quad (3.5)$$

$$C_{\ell m} := \int_{\Omega} D \phi_{\ell} : D \phi_m \, dx, \quad (3.6)$$

$$D_j := \int_{\Omega} \operatorname{tr} \Phi_j \, dx \quad (3.7)$$

sowie

$$F_j := - \int_{\Omega} f \cdot \operatorname{div} \Phi_j \, dx \quad (3.8)$$

und erkennen hier bereits folgende Blockstruktur für das resultierende lineare Gleichungssystem,

$$\begin{bmatrix} A & B & D \\ B^{\top} & C & 0 \\ D^{\top} & 0 & 0 \end{bmatrix} x = \begin{bmatrix} F \\ 0 \\ 0 \end{bmatrix},$$

wobei entsprechend den Komponenten der Basisfunktionen

$$\begin{aligned} A &= \begin{bmatrix} A_1^{(1D)}(\Omega) & A_2^{(1D)}(\Omega) \\ A_3^{(1D)}(\Omega) & A_4^{(1D)}(\Omega) \end{bmatrix} \in \mathbb{R}^{2|\mathcal{E}| \times 2|\mathcal{E}|} & B &= \begin{bmatrix} B_1^{(1D)}(\Omega) & B_2^{(1D)}(\Omega) \\ B_3^{(1D)}(\Omega) & B_4^{(1D)}(\Omega) \end{bmatrix} \in \mathbb{R}^{2|\mathcal{E}| \times 2d} \\ C &= \begin{bmatrix} C_1^{(1D)}(\Omega) & C_2^{(1D)}(\Omega) \\ C_3^{(1D)}(\Omega) & C_4^{(1D)}(\Omega) \end{bmatrix} \in \mathbb{R}^{2d \times 2d} & D &= \begin{bmatrix} D_1^{(1D)}(\Omega) \\ D_2^{(1D)}(\Omega) \end{bmatrix} \in \mathbb{R}^{2|\mathcal{E}|} \\ F &= \begin{bmatrix} F_1^{(1D)}(\Omega) \\ F_2^{(1D)}(\Omega) \end{bmatrix} \in \mathbb{R}^{2|\mathcal{E}|} \end{aligned}$$

und

$$x = \left( \sigma_1, \dots, \sigma_{2|\mathcal{E}|}, u_1, \dots, u_{2d}, \lambda \right)^{\top} \in \mathbb{R}^{2|\mathcal{E}|+2d+1}.$$



### 3.2 LSFEM mit konformer Wahl für die Geschwindigkeit

Für diese Methode (LS-CFEM) wählen wir für  $V^h(\mathcal{T})$  den konformen, d.h. in  $V(\Omega)$  enthaltenen, Raum der Courant-Funktionen

$$\begin{aligned} V^h(\mathcal{T}) &:= P_1(\mathcal{T}) \cap C^0(\overline{\Omega}), \\ V_0^h(\mathcal{T}) &:= \{v \in V^h(\mathcal{T}) \mid \forall z \in \mathcal{N}(\partial\Omega), v(z) = 0\}. \end{aligned}$$

mit  $d = |\mathcal{N}|$  und folgender Basis für  $V^h(\mathcal{T})$ ,

$$(\phi_1, \dots, \phi_{2|\mathcal{N}|}) = \left( \begin{pmatrix} \varphi_1 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} \varphi_{|\mathcal{N}|} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \varphi_1 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \varphi_{|\mathcal{N}|} \end{pmatrix} \right).$$

Dabei seien  $\varphi_\ell := \varphi_{z_\ell}$  die nodalen Basisfunktionen in den Knoten  $z_\ell \in \mathcal{N}$ .

**Definition 3.4 (Nodale Basisfunktionen).** Für alle Knoten  $z \in \mathcal{N}$  sind die nodalen Basisfunktionen  $\varphi_z \in P_1(\mathcal{T}) \cap C(\Omega)$  definiert vermöge der linearen Interpolation der Werte in  $y \in \mathcal{N}$ ,

$$\varphi_z(y) := \begin{cases} 1 & \text{falls } y = z, \\ 0 & \text{sonst.} \end{cases}$$

Eine mögliche Approximation  $g_h$  für die Randdaten aus Problem 3 ist  $g_h \in V^h(\mathcal{T})$  mit der Eigenschaft, dass  $\forall z \in \mathcal{N}(\partial\Omega), g_h(z) = g(z)$ .

Wir definieren folgende Indexmenge,

$$\mathcal{M} := \left\{ m \in \{1, \dots, 2|\mathcal{N}|\} \mid \phi_m \text{ gehört nicht zu einem Randknoten} \right\}.$$

Für die gegebenen Randdaten lautet dann das zu lösende Problem,

**Problem 5 (Gleichungssystem für LS-CFEM).** Finde  $x = (\sigma_1, \dots, \sigma_{2|\mathcal{E}|}, u_1, \dots, u_{2|\mathcal{N}|}, \lambda)^\top \in \mathbb{R}^{2|\mathcal{E}|+2|\mathcal{N}|+1}$ , sodass  $\forall j = 1, \dots, 2|\mathcal{E}| \forall \ell = 1, \dots, 2|\mathcal{N}|$ ,

$$\begin{aligned} \sum_{k \in \{1, \dots, 2|\mathcal{E}|\}} \sigma_k A_{jk} + \sum_{m \in \mathcal{M}} u_m B_{jm} + \lambda D_j &= F_j - \sum_{m \in \{1, \dots, 2|\mathcal{N}|\} \setminus \mathcal{M}} u_m B_{jm}, \\ \sum_{k \in \{1, \dots, 2|\mathcal{E}|\}} \sigma_k B_{\ell k}^\top + \sum_{m \in \mathcal{M}} u_m C_{\ell m} \, dx &= - \sum_{m \in \{1, \dots, 2|\mathcal{N}|\} \setminus \mathcal{M}} u_m C_{\ell m}, \\ \sum_{k \in \{1, \dots, 2|\mathcal{E}|\}} \sigma_k D_k &= 0. \end{aligned}$$

### 3.3 LSFEM mit nichtkonformer Wahl für die Geschwindigkeit

Für diese Methode (LS-NCFEM) wählen wir für  $V^h(\mathcal{T})$  den nichtkonformen, d.h. nicht in  $V(\Omega)$  enthaltenen Raum der Crouzeix-Raviart-Funktionen,

$$\begin{aligned} V^h(\mathcal{T}) &:= CR^1(\mathcal{T}; \mathbb{R}^2), \\ V_0^h(\mathcal{T}) &:= CR_0^1(\mathcal{T}; \mathbb{R}^2). \end{aligned}$$

**Definition 3.5 (Crouzeix-Raviart-Finite-Elemente-Räume).**

$$\begin{aligned} CR^1(\mathcal{T}) &:= \left\{ v \in P_1(\mathcal{T}) \mid \forall E \in \mathcal{E}, v \text{ ist stetig in } \text{mid}(E) \right\}, \\ CR_0^1(\mathcal{T}) &:= \left\{ v \in CR^1(\mathcal{T}) \mid \forall E \in \mathcal{E}(\partial\Omega), v(\text{mid } E) = 0 \right\}. \end{aligned}$$

Zu  $V^h(\mathcal{T})$  wählen wir die Basis

$$(\phi_1, \dots, \phi_{2|\mathcal{E}|}) = \left( \begin{pmatrix} \psi_1 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} \psi_{|\mathcal{E}|} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \psi_1 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \psi_{|\mathcal{E}|} \end{pmatrix} \right)$$

und es gilt  $d = |\mathcal{E}|$ . Dabei seien  $\psi_\ell := \psi_{E_\ell}$  die kantenorientierten Crouzeix-Raviart-Basisfunktionen.

**Definition 3.6 (Kantenorientierte Crouzeix-Raviart-Basisfunktionen).** Für  $E \in \mathcal{E}$  sei die kantenorientierte Crouzeix-Raviart-Basisfunktion  $\psi_E \in P_1(\mathcal{T})$  definiert durch lineare Interpolation der folgenden Werte in den Mittelpunkten der Kanten  $F \in \mathcal{E}$ ,

$$\psi_E(\text{mid}(F)) := \begin{cases} 1 & \text{falls } F = E, \\ 0 & \text{sonst.} \end{cases}$$

Eine mögliche Randapproximation zu den Daten aus Problem 3 ist  $g_h \in V^h(\mathcal{T})$  mit der Eigenschaft, dass  $\forall E \in \mathcal{E}(\partial\Omega), \int_E g_h \, ds = \int_E g \, ds$ .

Die Crouzeix-Raviart-Funktionen sind nicht notwendigerweise stetig und somit nicht alle schwach differenzierbar, d.h.  $V^h(\mathcal{T}) \not\subseteq V(\Omega)$ . Somit benötigen wir eine nichtkonforme Version von  $\mathcal{LS}$  mit dem stückweisen Gradienten  $D_{NC} : H^1(\mathcal{T}) \rightarrow L^2(\Omega; \mathbb{R}^2)$ ,  $(D_{NC} v)|_T := D(v|_T)$  für alle  $T \in \mathcal{T}$ ,

$$\begin{aligned} \mathcal{LS}_{NC} : L^2(\Omega; \mathbb{R}^2) \times \Sigma_0^h(\mathcal{T}) \times V^h(\mathcal{T}) &\ni (f; \tau_{LS-NC}, v_{LS-NC}) \\ &\mapsto \|\text{div } \tau_{LS-NC} + f\|_0^2 + \|\text{dev } \tau_{LS-NC} - D_{NC} v_{LS-NC}\|_0^2 \in \mathbb{R}. \end{aligned}$$

Die Probleme 3 und 4 können dann entsprechend umformuliert werden. Wir definieren folgende Indexmenge,

$$\mathcal{M}_{NC} := \left\{ m \in \{1, \dots, 2|\mathcal{E}|\} \mid \phi_m \text{ gehört nicht zu einer Randkante} \right\}.$$

Für die gegebenen Randdaten lautet dann das zu lösende Problem:

**Problem 6 (Gleichungssystem für LS-NCFEM).** Finde  $x = (\sigma_1, \dots, \sigma_{2|\mathcal{E}|}, u_1, \dots, u_{2|\mathcal{E}|}, \lambda)^\top \in \mathbb{R}^{4|\mathcal{E}|+1}$ , sodass  $\forall j, \ell = 1, \dots, 2|\mathcal{E}|$

$$\begin{aligned} \sum_{k \in \{1, \dots, 2|\mathcal{E}|\}} \sigma_k A_{jk} + \sum_{m \in \mathcal{M}_{NC}} u_m B_{jm} + \lambda D_j &= F_j - \sum_{m \in \{1, \dots, 2|\mathcal{E}|\} \setminus \mathcal{M}_{NC}} u_m B_{jm}, \\ \sum_{k \in \{1, \dots, 2|\mathcal{E}|\}} \sigma_k B_{\ell k}^\top + \sum_{m \in \mathcal{M}_{NC}} u_m C_{\ell m} \, dx &= - \sum_{m \in \{1, \dots, 2|\mathcal{E}|\} \setminus \mathcal{M}_{NC}} u_m C_{\ell m}, \\ \sum_{k \in \{1, \dots, 2|\mathcal{E}|\}} \sigma_k D_k &= 0. \end{aligned}$$

### 3.4 Über den Zusammenhang von LS-NCFEM, PSFEM und NCFEM

Wir rufen uns zunächst die diskreten Formulierungen der Pseudospannungs-Finite-Elemente-Methode (PSFEM) und der nichtkonformen Crouzeix-Raviart-Finite-Elemente-Methode (NCFEM) nach [CGS12] in Erinnerung.

**Problem 7 (PSFEM).** Seien  $f \in L^2(\Omega; \mathbb{R}^2)$  und  $g \in H^1(\Omega; \mathbb{R}^2) \cap H^1(\mathcal{E}(\partial\Omega); \mathbb{R}^2)$  mit  $\int_{\partial\Omega} g \cdot \nu \, ds = 0$  gegeben. Finde  $\sigma_{PS} \in \Sigma_0^h(\mathcal{T})$  und  $u_{PS} \in P_0(\mathcal{T}; \mathbb{R}^2)$ , sodass

$$\begin{aligned} \forall \tau_{PS} \in \Sigma_0^h(\mathcal{T}), \quad \int_{\Omega} \operatorname{dev} \sigma_{PS} : \tau_{PS} \, dx + \int_{\Omega} \operatorname{div} \tau_{PS} \cdot u_{PS} \, dx &= \langle g, \tau_{PS} \nu \rangle_{\partial\Omega} \\ \forall v_{PS} \in P_0(\mathcal{T}; \mathbb{R}^2), \quad \int_{\Omega} \operatorname{div} \sigma_{PS} \cdot v_{PS} \, dx &= - \int_{\Omega} f \cdot v_{PS} \, dx \end{aligned}$$

Dabei bezeichne  $\langle \cdot, \cdot \rangle_{\partial\Omega}$  die Dualitätsklammern von  $H^{-1/2}(\partial\Omega)$  auf  $H^{1/2}(\partial\Omega)$  auf dem Rand  $\partial\Omega$ . Für Formulierung der nichtkonformen Finite-Elemente-Methode benötigen wir noch den Funktionenraum

$$Q_{CR}(\mathcal{T}) := \left\{ q_{CR} \in P_0(\mathcal{T}) \mid \int_{\Omega} q_{CR} \, dx = 0 \right\}.$$

**Problem 8 (NCFEM).** Seien  $f \in L^2(\Omega; \mathbb{R}^2)$  und  $g \in H^1(\Omega; \mathbb{R}^2) \cap H^1(\mathcal{E}(\partial\Omega); \mathbb{R}^2)$  mit  $\int_{\partial\Omega} g \cdot \nu \, ds = 0$  gegeben. Sei  $g_{CR} \in CR^1(\mathcal{T}; \mathbb{R}^2)$  derart, dass für alle  $E \in$

$\mathcal{E}(\partial\Omega), \int_E g_{CR} \, ds = \int_E g \, ds$ . Finde  $u_{CR} \in g_{CR} + CR_0^1(\mathcal{T}; \mathbb{R}^2)$  und  $p_{CR} \in Q_{CR}(\mathcal{T})$ , sodass

$$\begin{aligned} \forall v_{CR} \in CR_0^1(\mathcal{T}; \mathbb{R}^2), \quad & \int_{\Omega} D_{NC} u_{CR} : D_{NC} v_{CR} \, dx - \int_{\Omega} p_{CR} \operatorname{div}_{NC} v_{CR} \, dx \\ & = \int_{\Omega} f \cdot v_{CR} \, dx, \\ \forall q_{CR} \in Q_{CR}(\mathcal{T}), \quad & \int_{\Omega} q_{CR} \operatorname{div}_{NC} u_{CR} \, dx = 0. \end{aligned}$$

Es ist bekannt, dass wir die Pseudospannungs-Finite-Elemente-Lösung mit Hilfe der nichtkonformen Crouzeix-Raviart-Lösung der Stokes-Gleichungen darstellen können. Seien  $\tilde{u}_{CR} \in CR^1(\mathcal{T}; \mathbb{R}^2)$  und  $\tilde{p}_{CR} \in Q_{CR}(\mathcal{T})$  die nichtkonforme Crouzeix-Raviart-Lösung zur rechten Seite  $f_{\mathcal{T}} := \Pi_0 f$  für  $f \in L^2(\Omega; \mathbb{R}^2)$ . Dann sind  $\sigma_{PS} \in RT_0(\mathcal{T}; \mathbb{R}^{2 \times 2})$  und  $u_{PS} \in P_0(\mathcal{T}; \mathbb{R}^2)$ , definiert vermöge

$$\begin{aligned} \sigma_{PS} &:= D_{NC} \tilde{u}_{CR} - \frac{\Pi_0 f}{2} \otimes (\bullet - \operatorname{mid} \mathcal{T}) - \tilde{p}_{CR} I_{2 \times 2}, \\ u_{PS} &:= \Pi_0 \tilde{u}_{CR} + \frac{1}{4} \Pi_0 (\operatorname{dev}(f_{\mathcal{T}} \otimes (\bullet - \operatorname{mid} \mathcal{T})) (\bullet - \operatorname{mid} \mathcal{T})), \end{aligned}$$

nach [CGS12, Theorem 2.1] Lösungen der Pseudospannungs-Finite-Elemente-Methode. Dabei bezeichne  $(\bullet - \operatorname{mid} \mathcal{T})$  die Funktion  $x \mapsto x - \operatorname{mid}(T)$  für  $x \in T \in \mathcal{T}$ . Wir sehen sofort, dass  $\operatorname{div} \sigma_{PS} = -\Pi_0 f$  gilt, da die restlichen Terme konstant sind. Außerdem ist  $\operatorname{dev} \sigma_{PS} = D_{NC} \tilde{u}_{CR} - \operatorname{dev}((\Pi_0 f) \otimes (\bullet - \operatorname{mid} \mathcal{T}))/2$ . Wenn wir nun die Funktionen aus der Lösung der PSFEM in das nichtkonforme Least-Squares-Funktional einsetzen, erhalten wir zusammen mit der Dreiecksungleichung,

$$\begin{aligned} \mathcal{LS}_{NC}(f; \sigma_{PS}, \tilde{u}_{CR}) &= \|f + \operatorname{div} \sigma_{PS}\|_0^2 + \|\operatorname{dev} \sigma_{PS} - D_{NC} \tilde{u}_{CR}\|_0^2 \\ &\leq \|f - \Pi_0 f\|_0^2 + \|\operatorname{dev}((\Pi_0 f) \otimes (\bullet - \operatorname{mid} \mathcal{T}))/2\|_0^2. \end{aligned}$$

Da für alle  $T \in \mathcal{T}$  und  $x \in T$ ,  $|x - \operatorname{mid} T| \leq h_T$  gilt, können wir schlussfolgern, dass

$$\mathcal{LS}_{NC}(f; \sigma_{PS}, \tilde{u}_{CR}) \lesssim \|f - \Pi_0 f\|_0^2 + \|h_{\mathcal{T}} \Pi_0 f\|_0^2. \quad (3.9)$$

Diese Abschätzung bedeutet, dass die nichtkonforme Least-Squares Finite-Elemente-Methode und die Pseudospannungs-Finite-Elemente-Methode bis auf einen Datenfehler identisch sind. Insbesondere sehen wir, dass das nichtkonforme Least-Squares-Funktional für konstante  $f$  nur durch einen  $h_{\mathcal{T}} f$ -Term beschränkt ist und es daher nicht als Fehlerschätzer geeignet ist.

### 3.5 Fehlerschätzer

#### 3.5.1 Least-Squares-Schätzer

Das Least-Squares-Funktional enthält bereits die Normen der Residuen. Somit eignet es sich zur a posteriori Fehlerabschätzung (vgl. hierzu [BG09, Abschnitt 12.14]). Wir werden also die lokalen Beiträge des Least-Squares-Funktional als Fehlerschätzer verwenden,

$$\begin{aligned} \mathcal{LS}_{\text{err}}^2(f; \sigma_{LS}, u_{LS}; T) &:= \|\operatorname{div} \sigma_{LS} + f\|_{0,T}^2 + \|\operatorname{dev} \sigma_{LS} - \mathbf{D} u_{LS}\|_{0,T}^2, \\ \mathcal{LS}_{\text{err}}(f; \sigma_{LS}, u_{LS}) &:= \left( \sum_{T \in \mathcal{T}} \mathcal{LS}_{\text{err}}^2(f; \sigma_{LS}, u_{LS}; T) \right)^{1/2}. \end{aligned} \quad (3.10)$$

Aus den Überlegungen aus Abschnitt 3.4 geht hervor, dass die Verwendung des Least-Squares-Funktional als Fehlerschätzer bei der nichtkonformen Methode keinen Sinn ergibt. Daher wenden wir den Schätzer (3.10) nur in der konformen Methode an.

#### 3.5.2 Drei weitere Fehlerschätzer

In [CP12, Abschnitt 6] werden für die Least-Squares Finite-Elemente-Methode im Poisson-Modellproblem Fehlerschätzer eingeführt, die auf Volumentermen und Tangentialsprüngen der diskreten Lösung sowie Kantensprünge der Normalenableitung basieren. Die dort bewiesenen Eigenschaften Effizienz und Zuverlässigkeit motivieren uns, analoge Schätzer bei unserer Lösung des Stokes-Problems zu untersuchen.

Dazu definieren wir zunächst folgende Beiträge für  $T \in \mathcal{T}, E \in \mathcal{E}$  und die Dirichlet-Randdaten  $g \in H^1(\Omega; \mathbb{R}^2) \cap H^1(\mathcal{E}(\partial\Omega); \mathbb{R}^2)$ ,

$$\begin{aligned} \alpha^2(T) &:= \|(1 - \Pi_0) \operatorname{dev} \sigma_{LS}\|_{0,T}^2, \\ \beta^2(E) &:= \begin{cases} |E| \|[\mathbf{D} u_{LS}]_E \cdot \nu_E\|_{0,E}^2 & , \text{ falls } E \in \mathcal{E}(\Omega) \\ 0 & , \text{ sonst} \end{cases}, \\ \gamma^2(E) &:= \begin{cases} |E| \|[\operatorname{dev} \sigma_{LS}]_E \cdot \tau_E\|_{0,E}^2 & , \text{ falls } E \in \mathcal{E}(\Omega) \\ |E| \|\operatorname{dev} \sigma_{LS} \cdot \tau_E - \partial g / \partial \tau_E\|_{0,E}^2 & , \text{ sonst} \end{cases}. \end{aligned}$$

Die Kantenterme fassen wir nun zu Beiträgen auf den einzelnen Dreiecken zusammen. Für  $T \in \mathcal{T}$  seien

$$\begin{aligned} \beta^2(T) &:= \sum_{E \in \mathcal{E}(T)} \beta^2(E), \\ \gamma^2(T) &:= \sum_{E \in \mathcal{E}(T)} \gamma^2(E). \end{aligned}$$

In Analogie zu [CP12] können wir damit die folgenden Fehlerschätzer definieren. Für alle  $T \in \mathcal{T}$ ,

$$\begin{aligned}\eta_1^2(T) &:= \alpha^2(T) + \beta^2(T) + \gamma^2(T), \\ \eta_2^2(T) &:= \alpha^2(T) + \beta^2(T), \\ \eta_3^2(T) &:= \alpha^2(T) + \gamma^2(T).\end{aligned}$$

Die globalen  $\eta$ -Schätzer werden dann entsprechend berechnet durch die folgende Summe für  $s = 1, 2, 3$ ,

$$\eta_s(\mathcal{T}) := \left( \sum_{T \in \mathcal{T}} \eta_s^2(T) \right)^{1/2}. \quad (3.11)$$

## 4 Implementierung

### 4.1 Berechnung der Steifigkeitsmatrix

#### 4.1.1 Block A

Wir berechnen nun die vier Blöcke von  $A$ . Für alle  $j', k' = 1, \dots, |\mathcal{E}|$  gilt nach (3.4) und unter Benutzung von Lemma 2.2 (ii),

$$\begin{aligned} \left( A_1^{(1D)}(\Omega) \right)_{j'k'} &= \int_{\Omega} \operatorname{div} \begin{pmatrix} \Psi_{j'} \\ 0 \end{pmatrix} \cdot \operatorname{div} \begin{pmatrix} \Psi_{k'} \\ 0 \end{pmatrix} dx + \int_{\Omega} \operatorname{dev} \begin{pmatrix} \Psi_{j'} \\ 0 \end{pmatrix} : \operatorname{dev} \begin{pmatrix} \Psi_{k'} \\ 0 \end{pmatrix} dx \\ &= \int_{\Omega} \operatorname{div} \Psi_{j'} \operatorname{div} \Psi_{k'} dx + \int_{\Omega} \Psi_{j'} \cdot \Psi_{k'} dx - \frac{1}{2} \int_{\Omega} \Psi_{j',1} \Psi_{k',1} dx. \end{aligned}$$

Analog berechnen wir

$$\begin{aligned} \left( A_2^{(1D)}(\Omega) \right)_{j'k'} &= -\frac{1}{2} \int_{\Omega} \Psi_{j',1} \Psi_{k',2} dx, \\ \left( A_3^{(1D)}(\Omega) \right)_{j'k'} &= -\frac{1}{2} \int_{\Omega} \Psi_{j',2} \Psi_{k',1} dx = \left( A_2^{(1D)}(\Omega) \right)_{j'k'}^{\top}, \\ \left( A_4^{(1D)}(\Omega) \right)_{j'k'} &= \int_{\Omega} \operatorname{div} \Psi_{j'} \operatorname{div} \Psi_{k'} dx + \int_{\Omega} \Psi_{j'} \cdot \Psi_{k'} dx - \frac{1}{2} \int_{\Omega} \Psi_{j',2} \Psi_{k',2} dx. \end{aligned}$$

Die lokalen Steifigkeitsmatrizen ergeben sich dann für ein Dreieck  $T \in \mathcal{T}$  (mit den Bezeichnungen aus Abbildung 2.1) für  $j, k = 1, 2, 3$  wie folgt,

$$\left( A_1^{(1D)}(T) \right)_{jk} = \underbrace{\int_T \operatorname{div} \Psi_j \operatorname{div} \Psi_k dx}_{=:\textcircled{1}} + \underbrace{\int_T \Psi_j \cdot \Psi_k dx}_{=:\textcircled{2}} - \underbrace{\frac{1}{2} \int_T \Psi_{j,1} \Psi_{k,1} dx}_{=:\textcircled{3}}.$$

Für den ersten Summanden gilt nach Lemma 3.3 (ii):

$$\textcircled{1} = \int_T \zeta_{E_j} \frac{|E_j|}{|T|} \zeta_{E_k} \frac{|E_k|}{|T|} dx = \zeta_{E_j} \zeta_{E_k} \frac{|E_j| |E_k|}{|T|}.$$

Mit der Matrix

$$\begin{aligned} L &:= \begin{bmatrix} \int_T \operatorname{div} \Psi_1 dx & 0 & 0 \\ 0 & \int_T \operatorname{div} \Psi_2 dx & 0 \\ 0 & 0 & \int_T \operatorname{div} \Psi_3 dx \end{bmatrix} \\ &= \begin{bmatrix} \zeta_{E_1} |E_1| & 0 & 0 \\ 0 & \zeta_{E_2} |E_2| & 0 \\ 0 & 0 & \zeta_{E_3} |E_3| \end{bmatrix} \in \mathbb{R}^{3 \times 3} \end{aligned}$$

lässt sich der Beitrag von ① für die lokale Steifigkeitsmatrix also berechnen durch

$$\frac{1}{|T|} L \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} L.$$

Für den zweiten Summanden gilt nach [BC05],

$$\begin{aligned} \textcircled{2} = \frac{\zeta_{E_j}|E_j|}{48|T|} & \left( \left( \sum_{\ell=1}^3 (P_\ell - P_{j-1}) \right) \cdot \left( \sum_{m=1}^3 (P_m - P_{k-1}) \right) \right. \\ & \left. + \sum_{\ell=1}^3 (P_\ell - P_{j-1}) \cdot (P_\ell - P_{k-1}) \right) \zeta_{E_k}|E_k|. \end{aligned}$$

Nach elementaren Rechnungen erhalten wir, dass sich der Beitrag von ② zur lokalen Steifigkeitsmatrix mit den Matrizen

$$\begin{aligned} N &:= \begin{bmatrix} P_1 - P_3 & 0 & P_1 - P_2 \\ P_2 - P_3 & P_2 - P_1 & 0 \\ 0 & P_3 - P_1 & P_3 - P_2 \end{bmatrix} = \begin{bmatrix} P_1 & P_1 & P_1 \\ P_2 & P_2 & P_2 \\ P_3 & P_3 & P_3 \end{bmatrix} - \begin{bmatrix} P_3 & P_1 & P_2 \\ P_3 & P_1 & P_2 \\ P_3 & P_1 & P_2 \end{bmatrix} \in \mathbb{R}^{6 \times 3}, \\ M^{(A)} &:= \begin{bmatrix} 2 & 0 & 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 & 0 & 1 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 & 0 & 2 \end{bmatrix} \in \mathbb{R}^{6 \times 6} \end{aligned}$$

berechnen lässt durch

$$\frac{1}{48|T|} L N^\top M^{(A)} N L.$$

Für den dritten Summanden rechnen wir auf analoge Weise nach, dass

$$\begin{aligned} \textcircled{3} = -\frac{1}{2} \frac{\zeta_{E_j}|E_j|}{48|T|} & \left( \left( \sum_{\ell=1}^3 (P_{\ell,1} - P_{j-1,1}) \right) \left( \sum_{m=1}^3 (P_{m,1} - P_{k-1,1}) \right) \right. \\ & \left. + \sum_{\ell=1}^3 (P_{\ell,1} - P_{j-1,1})(P_{\ell,1} - P_{k-1,1}) \right) \zeta_{E_k}|E_k|. \end{aligned}$$



Nach den selben Rechenschritten wie für ② wird klar, dass sich der Beitrag von ③ zur lokalen Steifigkeitsmatrix mit der Matrix

$$M_1^{(A)} := \begin{bmatrix} 2 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \text{diag}(1, 0, 1, 0, 1, 0)M \in \mathbb{R}^{6 \times 6}$$

darstellen lässt als

$$-\frac{1}{96|T|}LN^\top M_1^{(A)}NL.$$

Zusammen gilt dann für die lokale Steifigkeitsmatrix,

$$A_1^{(1D)}(T) = \frac{1}{|T|}L \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} L + \frac{1}{48|T|}LN^\top \left( M^{(A)} - \frac{1}{2}M_1^{(A)} \right) NL. \quad (4.1)$$

Mit den Matrizen

$$M_2^{(A)} := \begin{bmatrix} 0 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{6 \times 6}, \quad M_4^{(A)} := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 2 \end{bmatrix} \in \mathbb{R}^{6 \times 6}$$

erhalten wir durch analoge Berechnungen,

$$A_2^{(1D)}(T) = -\frac{1}{96|T|}LN^\top M_2^{(A)}NL, \quad (4.2)$$

$$A_4^{(1D)}(T) = \frac{1}{|T|}L \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} L + \frac{1}{48|T|}LN^\top \left( M^{(A)} - \frac{1}{2}M_4^{(A)} \right) NL. \quad (4.3)$$

#### 4.1.2 Block B

Wir betrachten nun die vier Blöcke von  $B$ . Für alle  $j' = 1, \dots, |\mathcal{E}|$  und  $m' = 1, \dots, d$  gilt nach Gleichung (3.5),

$$\begin{aligned} \left( B_1^{(1D)}(\Omega) \right)_{j'm'} &= - \int_{\Omega} \text{dev} \begin{pmatrix} \Psi_{j'} \\ 0 \end{pmatrix} : D_{NC} \begin{pmatrix} \psi_{m'} \\ 0 \end{pmatrix} dx \\ &= - \int_{\Omega} \Psi_{j'} \cdot D \psi_{m'} dx + \frac{1}{2} \int_{\Omega} \Psi_{j',1} \frac{\partial}{\partial x_1} \psi_{m'} dx. \end{aligned}$$

Analog berechnen wir,

$$\begin{aligned} \left( B_2^{(1D)}(\Omega) \right)_{j'm'} &= \frac{1}{2} \int_{\Omega} \Psi_{j',1} \frac{\partial}{\partial x_2} \psi_{m'} \, dx, \\ \left( B_3^{(1D)}(\Omega) \right)_{j'm'} &= \frac{1}{2} \int_{\Omega} \Psi_{j',2} \frac{\partial}{\partial x_1} \psi_{m'} \, dx, \\ \left( B_4^{(1D)}(\Omega) \right)_{j'm'} &= - \int_{\Omega} \Psi_{j'} \cdot D \psi_{m'} \, dx + \frac{1}{2} \int_{\Omega} \Psi_{j',2} \frac{\partial}{\partial x_2} \psi_{m'} \, dx. \end{aligned}$$

Die lokalen Steifigkeitsmatrizen ergeben sich dann für ein Dreieck  $T \in \mathcal{T}$  (mit den Bezeichnungen aus Abbildung 2.1) für  $j, m = 1, 2, 3$  wie folgt,

$$\left( B_1^{(1D)}(T) \right)_{jm} = \underbrace{- \int_{\Omega} \Psi_j \cdot D \psi_m \, dx}_{=: \textcircled{1}} + \underbrace{\frac{1}{2} \int_T \Psi_{j,1} \frac{\partial}{\partial x_1} \psi_m \, dx}_{=: \textcircled{2}}.$$

Für den ersten Summanden gilt nach Lemma 3.3 (iii), da  $D \psi_m = (\partial \psi_m / \partial x_1, \partial \psi_m / \partial x_2)$  konstant auf  $T$  ist,

$$\textcircled{1} = - \int_T \Psi_j \, dx \cdot D \psi_m = -\zeta_{E_j} \frac{|E_j|}{6} \sum_{\ell=1}^3 (P_{\ell} - P_{j-1}) \cdot D \psi_m.$$

Mit den Matrizen

$$G := \begin{bmatrix} D \psi_1 \\ D \psi_2 \\ D \psi_3 \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_1} \psi_1 & \frac{\partial}{\partial x_2} \psi_1 \\ \frac{\partial}{\partial x_1} \psi_2 & \frac{\partial}{\partial x_2} \psi_2 \\ \frac{\partial}{\partial x_1} \psi_3 & \frac{\partial}{\partial x_2} \psi_3 \end{bmatrix} \in \mathbb{R}^{3 \times 2}, \quad M^{(B)} := \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}^{\top} \in \mathbb{R}^{6 \times 2}$$

und  $L$  und  $N$  wie oben lässt sich der Beitrag von  $\textcircled{1}$  für die lokale Steifigkeitsmatrix berechnen durch

$$-\frac{1}{6} L N^{\top} M^{(B)} G^{\top}.$$

Auch dies wird nach elementarer Algebra ersichtlich. Für den zweiten Summanden gilt, da der Gradient konstant ist,

$$\textcircled{2} = \frac{1}{2} \int_T \Psi_{j,1} \, dx \frac{\partial}{\partial x_1} \psi_m = \frac{1}{2} \zeta_{E_j} \frac{|E_j|}{6} \sum_{\ell=1}^3 (P_{\ell,1} - P_{j-1,1}) \frac{\partial}{\partial x_1} \psi_m.$$

Mit der Matrix

$$M_1^{(B)} := \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^{\top} \in \mathbb{R}^{6 \times 2}$$

lässt sich der Beitrag von ② für die lokale Steifigkeitsmatrix analog berechnen durch

$$\frac{1}{12} L N^\top M_1^{(B)} G^\top.$$

Zusammen gilt dann für die lokale Steifigkeitsmatrix

$$B_1^{(1D)}(T) = -\frac{1}{6} L N^\top \left( M^{(B)} - \frac{1}{2} M_1^{(B)} \right) G^\top. \quad (4.4)$$

Mit den Matrizen

$$M_2^{(B)} := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}^\top \in \mathbb{R}^{6 \times 2}, \quad M_3^{(B)} := \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^\top \in \mathbb{R}^{6 \times 2},$$

$$M_4^{(B)} := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}^\top \in \mathbb{R}^{6 \times 2}.$$

erhalten wir nach analogen Berechnungen

$$B_2^{(1D)}(T) = \frac{1}{12} L N^\top M_2^{(B)} G^\top, \quad (4.5)$$

$$B_3^{(1D)}(T) = \frac{1}{12} L N^\top M_3^{(B)} G^\top, \quad (4.6)$$

$$B_4^{(1D)}(T) = -\frac{1}{6} L N^\top \left( M^{(B)} - \frac{1}{2} M_4^{(B)} \right) G^\top. \quad (4.7)$$

#### 4.1.3 Block C

Wir betrachten nun die vier Blöcke von  $C$ . Für alle  $\ell', m' = 1, \dots, d$  gilt nach Gleichung (3.6)

$$\left( C_1^{(1D)}(\Omega) \right)_{\ell' m'} = \int_{\Omega} D \begin{pmatrix} \psi_{\ell'} \\ 0 \end{pmatrix} : D \begin{pmatrix} \psi_{m'} \\ 0 \end{pmatrix} dx = \int_{\Omega} D \psi_{\ell'} \cdot D \psi_{m'} dx.$$

Analog berechnen wir

$$\begin{aligned} \left( C_2^{(1D)}(\Omega) \right)_{\ell' m'} &= \left( C_3^{(1D)}(\Omega) \right)_{\ell' m'} = 0, \\ \left( C_4^{(1D)}(\Omega) \right)_{\ell' m'} &= \int_{\Omega} D \psi_{\ell'} \cdot D \psi_{m'} dx = \left( C_1^{(1D)}(\Omega) \right)_{\ell' m'}. \end{aligned}$$

Die lokalen Steifigkeitsmatrizen ergeben sich dann für ein Dreieck  $T \in \mathcal{T}$  (mit den Bezeichnungen aus Abbildung 2.1) für  $\ell, m = 1, 2, 3$  wie folgt,

$$\left( C_1^{(1D)}(T) \right)_{\ell m} = \int_T D \psi_{\ell} \cdot D \psi_m dx = |T| D \psi_{\ell} \cdot D \psi_m.$$

Hier benutzen wir, dass die Gradienten konstant auf  $T$  sind. Mit den obigen Matrizen lässt sich die lokale Steifigkeitsmatrix berechnen durch

$$C_1^{(1D)}(T) = |T| G G^\top. \quad (4.8)$$

#### 4.1.4 Block D

Wir betrachten nun die zwei Blöcke von  $D$ . Für alle  $j' = 1, \dots, |\mathcal{E}|$  gilt nach (3.7),

$$\left(D_1^{(1D)}(\Omega)\right)_{j'} = \int_{\Omega} \operatorname{tr} \begin{pmatrix} \Psi_{j'} \\ 0 \end{pmatrix} dx = \int_{\Omega} \Psi_{j',1} dx$$

und analog

$$\left(D_2^{(1D)}(\Omega)\right)_{j'} = \int_{\Omega} \Psi_{j',2} dx.$$

Die lokalen Steifigkeitsmatrizen ergeben sich dann für ein Dreieck  $T \in \mathcal{T}$  (mit den Bezeichnungen aus Abbildung 2.1) für  $j = 1, 2, 3$  entsprechend Lemma 3.3 (iii) wie folgt,

$$\left(D_1^{(1D)}(T)\right)_j = \int_T \Psi_{j,1} dx = \zeta_{E_j} \frac{|E_j|}{6} \sum_{\ell=1}^3 (P_{\ell,1} - P_{j-1,1}).$$

Nach elementaren Berechnungen mit den obigen Matrizen lässt sich die lokale Steifigkeitsmatrix bestimmen,

$$D_1^{(1D)}(T) = \frac{1}{6} L N^{\top} M^{(B)} \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \quad (4.9)$$

Auf analogem Wege erhalten wir

$$D_2^{(1D)}(T) = \frac{1}{6} L N^{\top} M^{(B)} \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (4.10)$$

#### 4.1.5 Block F

Wir betrachten nun die zwei Blöcke von  $F$ . Für alle  $j' = 1, \dots, |\mathcal{E}|$  gilt nach (3.8),

$$\left(F_1^{(1D)}(\Omega)\right)_{j'} = - \int_{\Omega} f \cdot \operatorname{div} \begin{pmatrix} \Psi_{j'} \\ 0 \end{pmatrix} dx = \int_{\Omega} f_1 \operatorname{div} \Psi_{j'} dx$$

und analog

$$\left(F_2^{(1D)}(\Omega)\right)_{j'} = - \int_{\Omega} f_2 \operatorname{div} \Psi_{j'} dx.$$

Die lokalen rechten Seiten ergeben sich dann für ein Dreieck  $T \in \mathcal{T}$  (mit den Bezeichnungen aus Abbildung 2.1) für  $j = 1, 2, 3$  unter Benutzung von Lemma 3.3 (ii) wie folgt,

$$\left(F_1^{(1D)}(T)\right)_j = -\zeta_{E_j} \frac{|E_j|}{|T|} \int_T f_1 dx.$$

Mit den obigen Matrizen lässt sich die lokale rechte Seite berechnen durch

$$F_1^{(1D)}(T) = -\frac{1}{|T|} \int_T f_1 \, dx \, \text{diag}(L) \quad (4.11)$$

und analog

$$F_2^{(1D)}(T) = -\frac{1}{|T|} \int_T f_2 \, dx \, \text{diag}(L). \quad (4.12)$$

## 4.2 Berechnung der Fehlerschätzer

### 4.2.1 Least-Squares-Schätzer

Sei  $(\sigma_{LS-C}, u_{LS-C}) \in \Phi^h \times H^h$  die diskrete Lösung der konformen Least-Squares-Methode mit den Basisdarstellungen

$$\sigma_{LS-C} = \sum_{k=1}^{2|\mathcal{E}|} \sigma_k \Phi_k \quad \text{und} \quad u_{LS-C} = \sum_{m=1}^{2|\mathcal{N}|} u_m \phi_m.$$

Der Least-Squares-Schätzer lässt sich leicht mit den Daten der Steifigkeitsmatrix berechnen. Er besteht aus zwei Summanden,

$$\mathcal{LS}_{\text{err}}(f; \sigma_{LS-C}, u_{LS-C}) = \|\text{div } \sigma_{LS-C} + f\|_0^2 + \|\text{dev } \sigma_{LS-C} - D u_{LS-C}\|_0^2.$$

Um diese später qualitativ mit Hilfe des Quotienten

$$q := \|\text{div } \sigma_{LS-C} + \Pi_0 f\|_0^2 / \|\text{dev } \sigma_{LS-C} - D u_{LS-C}\|_0^2 \quad (4.13)$$

vergleichen zu können, sollten wir die beiden Normen getrennt berechnen. Insbesondere bedeutet dies, dass wir die Matrizen  $A_1^{(1D)}(T), \dots, A_4^{(1D)}(T)$  in einen Deviator- und einen Divergenzbeitrag additiv zerlegen müssen. Für  $T \in \mathcal{T}$  und  $s = 1, \dots, 4$  sei

$$A_s^{(\text{dev}), (1D)}(T) := \frac{1}{48|T|} L N^\top \left( M^{(A)} - \frac{1}{2} M_s^{(A)} \right) N L,$$

$$A^{(\text{div}), (1D)}(T) := L \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} L$$

und damit

$$A^{(\text{dev})}(T) := \begin{bmatrix} A_1^{(\text{dev}), (1D)}(T) & A_2^{(\text{dev}), (1D)}(T) \\ A_2^{(\text{dev}), (1D)}(T)^\top & A_4^{(\text{dev}), (1D)}(T) \end{bmatrix}, \quad A^{(\text{div})}(T) := \begin{bmatrix} A^{(\text{div}), (1D)}(T) & 0 \\ 0 & A^{(\text{div}), (1D)}(T) \end{bmatrix}$$

sowie in Analogie zu Abschnitt 4.1

$$\begin{aligned} B(T) &:= \begin{bmatrix} B_1^{(1D)}(T) & B_2^{(1D)}(T) \\ B_3^{(1D)}(T) & B_4^{(1D)}(T) \end{bmatrix}, & C(T) &:= \begin{bmatrix} C_1^{(1D)}(T) & 0 \\ 0 & C_1^{(1D)}(T) \end{bmatrix}, \\ F(T) &:= \begin{bmatrix} -F_1^{(1D)}(T) \\ -F_2^{(1D)}(T) \end{bmatrix}, & x_\sigma &:= (\sigma_1, \dots, \sigma_{2|\mathcal{E}|})^\top, \\ & & x_u &:= (u_1, \dots, u_{2|\mathcal{N}|})^\top. \end{aligned}$$

Zudem benötigen wir noch eine Bezeichnung für die lokalen Koeffizienten der Basisfunktionen, also für die Koeffizienten zu den Basisfunktionen, deren Träger ein bestimmtes Dreieck enthält. Für  $T \in \mathcal{T}$  mit  $\mathcal{E}(T) = \{E_{j_1}, E_{j_2}, E_{j_3}\}$  und  $\mathcal{N}(T) = \{P_{\ell_1}, P_{\ell_2}, P_{\ell_3}\}$  seien

$$\begin{aligned} x_\sigma(T) &:= (\sigma_{j_1}, \sigma_{j_2}, \sigma_{j_3}, \sigma_{|\mathcal{E}|+j_1}, \sigma_{|\mathcal{E}|+j_2}, \sigma_{|\mathcal{E}|+j_3})^\top \text{ und} \\ x_u(T) &:= (u_{\ell_1}, u_{\ell_2}, u_{\ell_3}, u_{|\mathcal{N}|+\ell_1}, u_{|\mathcal{N}|+\ell_2}, u_{|\mathcal{N}|+\ell_3})^\top. \end{aligned}$$

Damit berechnen sich die Beiträge des Funktional wie folgt,

$$\begin{aligned} & \|\operatorname{div} \sigma_{LS-C} + f\|_0^2 \\ &= \sum_{T \in \mathcal{T}} \left( \sum_{j,k=1}^{2|\mathcal{E}|} \sigma_j \int_T \operatorname{div} \Phi_j \cdot \operatorname{div} \Phi_k \, dx \sigma_k + 2 \sum_{j=1}^{2|\mathcal{E}|} \sigma_j \int_T \operatorname{div} \Phi_j \cdot f \, dx + \int_T f \cdot f \, dx \right) \\ &= \sum_{T \in \mathcal{T}} \left( x_\sigma(T)^\top A^{(\operatorname{div})}(T) x_\sigma(T) + 2x_\sigma(T) \cdot F(T) + \int_T f \cdot f \, dx \right) \end{aligned} \tag{4.14}$$

und

$$\begin{aligned} & \|\operatorname{dev} \sigma_{LS-C} - D u_{LS-C}\|_0^2 \\ &= \sum_{T \in \mathcal{T}} \left( \sum_{j,k=1}^{2|\mathcal{E}|} \sigma_j \int_T \operatorname{dev} \Phi_j : \operatorname{dev} \Phi_k \, dx \sigma_k - \sum_{j=1}^{2|\mathcal{E}|} \sum_{m=1}^{2|\mathcal{N}|} \sigma_j \int_T \operatorname{dev} \Phi_j : D \phi_m \, dx u_m \right. \\ & \quad \left. - \sum_{k=1}^{2|\mathcal{E}|} \sum_{\ell=1}^{2|\mathcal{N}|} u_\ell \int_T D \phi_\ell : \operatorname{dev} \Phi_k \, dx \sigma_k + \sum_{\ell,m=1}^{2|\mathcal{N}|} u_\ell \int_T D \phi_\ell : D \phi_m \, dx u_m \right) \\ &= \sum_{T \in \mathcal{T}} \left( x_\sigma(T)^\top A^{(\operatorname{dev})}(T) x_\sigma(T) + x_u(T)^\top B(T) x_\sigma(T) + \right. \\ & \quad \left. x_u(T)^\top B(T)^\top x_\sigma(T) + x_u(T)^\top C(T) x_u(T) \right). \end{aligned} \tag{4.15}$$

Der Zähler des Quotienten  $q$  lässt sich nach  $\int_{\Omega} \Pi_0 f \cdot \Pi_0 f \, dx = |\int_{\Omega} f \, dx|^2 / |T|$  ebenfalls leicht berechnen,

$$\begin{aligned} & \|\operatorname{div} \sigma_{LS-C} + \Pi_0 f\|_0^2 \\ &= \sum_{T \in \mathcal{T}} \left( x_{\sigma}(T)^{\top} A^{(\operatorname{div})}(T) x_{\sigma}(T) + 2x_{\sigma}(T) \cdot F(T) + \frac{1}{|T|} \left| \int_{\Omega} f \, dx \right|^2 \right). \end{aligned} \quad (4.16)$$

#### 4.2.2 Berechnung der eta-Schätzer

Es sind die drei Beiträge  $\alpha^2(T)$ ,  $\beta^2(E)$  und  $\gamma^2(E)$  zu berechnen. Für den Term  $\beta^2(E)$  benutzen wir Methoden aus dem AFEM-Paket [Car09a] und müssen daher hier keine weiteren Überlegungen vornehmen.

Nach dem Satz des Pythagoras gilt,

$$\begin{aligned} \alpha^2(T) &= \|(1 - \Pi_0) \operatorname{dev} \sigma_{LS}\|_{0,T}^2 \\ &= \|\operatorname{dev} \sigma_{LS}\|_{L^2(T; \mathbb{R}^2)}^2 - \|\Pi_0 \operatorname{dev} \sigma_{LS}\|_{L^2(T; \mathbb{R}^2)}^2 \\ &= \|\operatorname{dev} \sigma_{LS}\|_{L^2(T; \mathbb{R}^2)}^2 - \frac{1}{|T|} \left( \int_T \operatorname{dev} \sigma_{LS} \, dx \right)^2. \end{aligned}$$

Nach den Rechnungen aus (4.15) und der Linearität des Integrals gilt,

$$\begin{aligned} \alpha^2(T) &= x_{\sigma}(T)^{\top} A^{(\operatorname{dev})}(T) x_{\sigma}(T) \\ &\quad - \frac{1}{|T|} \left( \sum_{j=1}^3 \sigma_j \operatorname{dev} \int_T \Phi_j \, dx + \sigma_{j+|\mathcal{E}|} \operatorname{dev} \int_T \Phi_{j+|\mathcal{E}|} \, dx \right)^2. \end{aligned} \quad (4.17)$$

Mit Lemma 3.3 (iii) können wir die Integrale der Basisfunktionen leicht unter Verwendung der Matrizen  $L$  und  $N$  berechnen.

Für  $\gamma^2(E)$  benutzen wir zur Integration die Simpson-Formel, die für quadratische Polynome exakt ist, für  $E = \operatorname{conv}\{A, B\} \in \mathcal{E}$  lautet sie

$$\int_E v \, ds \approx |E|/6 (v(A) + 4v(\operatorname{mid}(E)) + v(B)).$$

Für jede Seite  $E = \operatorname{conv}\{A, B\} \in \mathcal{E}(\Omega)$  mit Mittelpunkt  $M = \operatorname{mid}(E)$ , kann demnach das gesuchte Integral direkt berechnet werden durch

$$\begin{aligned} \gamma^2(E) &= |E| \|\operatorname{dev} \sigma_{LS}|_E \cdot \tau_E\|_{0,E}^2 \\ &= |E|^2/6 \left( |\operatorname{dev} \sigma_{LS}|_E(A)|^2 + 4|\operatorname{dev} \sigma_{LS}|_E(M)|^2 + |\operatorname{dev} \sigma_{LS}|_E(B)|^2 \right), \end{aligned} \quad (4.18)$$

da die Funktionen  $\sigma_{LS}$  stückweise affin sind. Für  $E \in \mathcal{E}\partial\Omega$  gilt eine analoge Formel.

### 4.3 Realisierung

Die im Rahmen dieser Arbeit implementierten Methoden sind als Erweiterung des AFEM-Paketes der Arbeitsgruppe „Numerische Analysis“ von Prof. Carstensen am Institut für Mathematik der Humboldt-Universität zu Berlin [Car09a] gedacht und sind folglich zur korrekten Funktionsweise auf andere Methoden aus diesem Paket angewiesen.

Wichtig ist es, zu erwähnen, dass ich die Funktion `P0TangentJumps` aus dem Paket modifizieren musste. Sie berechnet die Tangentialsprünge von  $P_0(\mathcal{T})$ -Funktionen, bricht jedoch mit einem Fehler ab, wenn leere Neumann-Randdaten übergeben werden. Dies musste verändert werden, da wir stets von reinen Dirichlet-Daten ausgehen. Eine einfache Abfrage vor den entsprechenden Anweisungen, ob es überhaupt Neumann-Randstücke gibt, genügt, um das Problem zu beheben.

Die ausführbare Methode der Programme lautet

```
function StokesLS(problem,element,estimator,minNrDof,theta,
    foldername,subfolder) .
```

Sie realisiert den AFEM-Algorithmus (vgl. Abbildung 4.1). Wir gehen davon aus, dass dieser bekannt ist und verweisen für weitere Informationen auf [CGK<sup>+</sup>10]. Über die Argumente kann dem Programm mitgeteilt werden, welches Beispielp Problem und ob mit dem konformen oder dem nichtkonformen Element gerechnet werden soll. Es können ein Fehlerschätzer, ein „bulk chasing“-Parameter  $\theta$  und die minimale Anzahl an Freiheitsgraden als Abbruchbedingung angegeben werden. Tabelle 4.1 zeigt die möglichen Optionen an. Die entsprechenden Problem Daten, wie die Anfangstriangulierung, die rechte Seite  $f$ , die Dirichlet-Randdaten und ggf. die exakte Lösung, werden im Ablauf des Programmes aus den Dateien

```
Stokes_data_bf.m
Stokes_data_colliding.m
Stokes_data_Lshape.m
Stokes_data_noPress.m
Stokes_data_slit.m
```

ausgelesen. Sie wurden teils aus dem Softwarepaket [Car09a] und teils aus [BH12] übernommen.

Wir verwenden die Datenstrukturen aus Tabelle 4.2 aus dem AFEM-Paket. Für weitere Informationen verweisen wir auf die AFEM-Dokumentation [CGK<sup>+</sup>10] und die Praktikumsdokumentation [BH12].

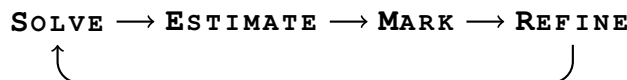


Abbildung 4.1: AFEM-Schleife



problem	Beispiel	element	Methode
bf	„backward facing step“	c	LS-CFEM
colliding	„colliding flow“	nc	LS-NCFEM
Lshape	L-Gebiet		
slit	Schlitzgebiet		
noPress	Poisson-Problem		

estimator	Schätzer	theta	Verfeinerung
ls	LS-Schätzer	$\theta = 1$	uniform
eta	$\eta$ -Schätzer	$0 < \theta < 1$	adaptiv

 Tabelle 4.1: Mögliche Argumente für die `StokesLS`-Methode

Im Folgenden wird die logische Struktur der implementierten Matlab-Funktionen erläutert. Sie beginnen mit der Initialisierung der benötigten geometrischen Datenstrukturen und der Berechnung weiterer Informationen, wie Integralen von Funktionen oder den Freiheitsgraden. Für die Integration wird stets die AFEM-Routine `integrate` benutzt. Wir wenden diese lediglich auf Polynome geringer Ordnung und trigonometrische Funktionen an. Ein Integrationsgrad von 10 erscheint daher als völlig ausreichend.

Es folgt die Definition von Konstanten, die während des Programmablaufs benötigt werden. Anschließend wird Speicherplatz für die Variablen reserviert, um die Performance nicht unnötig zu verringern. Schließlich werden die Berechnungen durchgeführt, häufig in for-Schleifen über die Dreiecke oder Kanten der Triangulierung. Dabei werden eingangs die lokalen Daten in gesonderte Variablen kopiert, um den Quelltext verständlicher zu gestalten.

Wo es möglich ist, werden zusätzlich zu den aus Funktionen ausgegebenen Vektoren, die lokale Werte für jedes Element oder jede Kante enthalten, auch eine Summe über alle Einträge optional ausgegeben. Wenn also an aufrufender Stelle nur die Summen benötigt werden, müssen nicht die gesamten Vektoren mit übergeben werden. Dadurch wird beispielsweise im Hauptprogramm viel Arbeitsspeicher eingespart.

### 4.3.1 Lösungsmethoden

Wir betrachten zunächst die beiden Lösungsmethoden, für die konforme Least-Squares-Methode

```
function [sigma,u,nrDof,STIMA,b] = solveLSCStokes(f,u4sDb,
    c4n,n4e,n4sDb)
```

und die nichtkonforme

Datum	Beschreibung
<i>Geometrie</i>	
c4n	Koordinaten für die Knoten der Triangulierung (zeilenweise)
n4e	Knotennummern für die Dreiecke der Triangulierung (zeilenweise), wobei die ersten beiden Knoten die Referenzkante bilden
n4s	Knotennummern für die Kanten der Triangulierung (zeilenweise)
s4e	Kantennummern für die Dreiecke der Triangulierung (zeilenweise), wobei die erste Kante die Referenzkante ist
s4n	Kantennummern für eine gemeinsame Kante zweier Knoten (0, falls Knoten keine gemeinsame Kante besitzen)
e4s	Nummern der angrenzenden Dreiecke einer Kante (Reihenfolge entsprechend der Richtung des Normalenvektors der Kante)
n4sDb	Knotennummern für die Randkanten der Triangulierung, auf denen Dirichlet-Daten gegeben sind (zeilenweise)
tangent4e	die drei Tangentialeinheitsvektoren für die Dreiecke der Triangulierung
normal4e	die drei äußeren Normaleneinheitsvektoren für die Dreiecke der Triangulierung
normal4s	Normaleneinheitsvektor für die Seiten der Triangulierung (als Zeilenvektoren)
mid4e	Mittelpunkte der Dreiecke der Triangulierung
mid4s	Mittelpunkte der Kanten der Triangulierung
area4e	Flächeninhalt der Dreiecke der Triangulierung
length4s	Länge der Kanten der Triangulierung
<i>Input Daten</i>	
f	rechte Seite aus dem zu lösenden Problem
u4Db	Dirichlet-Randdaten
<i>Gleichungssystem</i>	
STIMA	Steifigkeitsmatrix
b	rechte Seite des Gleichungssystems
x	Lösungsvektor des Gleichungssystems
<i>Integrale</i>	
intF4e	Integral der beiden Komponenten von $\mathbf{f}$ auf jedem Dreieck der Triangulierung
intFSq4e	Integral des Skalarproduktes von $\mathbf{f}$ mit sich selbst auf jedem Dreieck der Triangulierung

Tabelle 4.2: Datenstrukturen

```
function [sigma,u,nrDof,STIMA,b] = solveLSNCStokes(f,u4sDb,
    c4n,n4e,n4sDb) .
```

Wie bereits an der Berechnung der Steifigkeitsmatrix in Abschnitt 4.1 zu erkennen ist, sind die beiden Lösungsmethoden sehr ähnlich aufgebaut. Daher können wir hier beide gleichzeitig betrachten.

Aus den Formeln (4.1)–(4.12) wird direkt ersichtlich, wie diese zu implementieren sind. Dabei werden die Integrale der Komponenten von  $f$ , wie oben beschrieben, mit der `integrate`-Methode berechnet. Zur Herleitung der expliziten Berechnung der Gradientenmatrix  $G$  verweisen wir auf [CGK<sup>+</sup>10]. Diese ist neben der Dimension der einzelnen Matrixblöcke der einzige Unterschied zwischen konformer und nichtkonformer Methode.

Für jedes Dreieck werden die lokalen Beiträge zunächst in den Blöcken  $A_1^{(1D)}$  bis  $F_2^{(1D)}$  an den entsprechenden Stellen aufaddiert und anschließend zum linearen Gleichungssystem zusammengefasst. Vor dem Aufbau der globalen Steifigkeitsmatrix werden nicht mehr benötigte Daten mittels `clear`-Befehl gelöscht, um den Arbeitsspeicher nicht unnötig zu belasten. Abschließend werden die Dirichlet-Randdaten von der rechten Seite entsprechend den Problemen 5 und 6 abgezogen und das Gleichungssystem auf den eingangs berechneten Freiheitsgraden gelöst.

### 4.3.2 Schätzmethoden

Der Least-Squares-Schätzer aus (3.10) wird in

```
function [value4e,valueSum,quotient] =
    estimateStokesLSCValueElements(f,sigma,u,c4n,n4e)
```

implementiert. Zur Auswertung des Least-Squares-Funktionalen sind dieselben Blockmatrizen nötig, die auch zur Steifigkeitsmatrix gehören. Daher ähneln sich diese Methode und `solveLSCStokes` sehr stark. Die einzigen Unterschiede sind die Aufspaltung der Blöcke in Divergenz- und Deviatorbeiträge, wie sie schon in Abschnitt 4.2.1 dargestellt worden ist, und die Berechnung der zusätzlichen Terme für den Quotienten  $q$  aus (4.13). Die notwendigen Rechenschritte, werden wieder aus den Formeln (4.14)–(4.16) ersichtlich.

Zur Bestimmung der  $\eta$ -Schätzer benutzen wir

```
function [eta4e1,eta4e2,eta4e3,alpha4e,beta4e,gamma4e,
    etaSum1,etaSum2,etaSum3,alphaSum,betaSum,gammaSum] =
    estimateStokesLSCEtaElements(f,u4Db,Du4Db1,Du4Db2,sigma,u,
    c4n,n4e,n4sDb,n4sNb,problem)
```

und

```
function [eta4e1,eta4e2,eta4e3,alpha4e,beta4e,gamma4e,
    etaSum1,etaSum2,etaSum3,alphaSum,betaSum,gammaSum] =
```

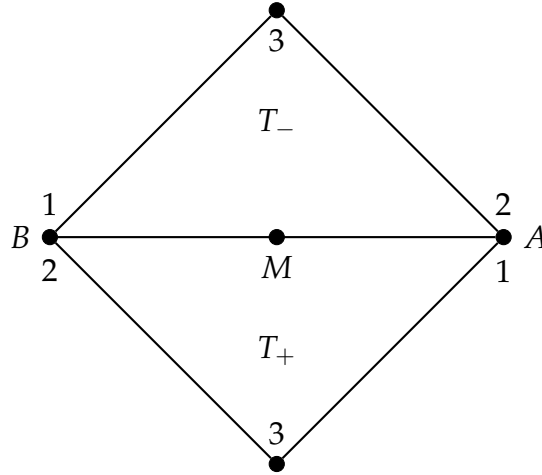


Abbildung 4.2: Lokale Umnummerierung der Knoten in den Funktionen `estimateStokesLSCetaElements` und `estimateStokesLSNCetaElements`

```
estimateStokesLSNCetaElements(f, u4Db, Du4Db1, Du4Db2, sigma,
u, c4n, n4e, n4sDb, n4sNb, problem) .
```

Die Berechnung von  $\alpha^2(T)$  erfolgt nach (4.17) wieder mit den bereits bekannten Rechenschritten unter Zuhilfenahme der Blöcke der Steifigkeitsmatrix. Die kantenorientierten Basisfunktionen  $\Phi_E$  werden dabei mit der Formel aus Lemma 3.3 (iii) integriert und neu geordnet, um den deviatorischen Anteil zu erhalten. Zusammen mit den lokalen Koeffizienten der Basisdarstellung  $x_\sigma(T)$  und dem Faktor  $1/|T|$  wird dann der zweite Beitrag berechnet.

Die  $\beta^2(E)$  werden direkt mit AFEM-Methoden bestimmt, im konformen Fall mit `estimatePlEtaSides` und im nichtkonformen mit `estimateCREtaSides`.

Auch bei der Berechnung der  $\gamma^2(E)$  werden einfache, bereits bekannte Rechnungen verwendet. Hier werden zunächst für jede Seite  $E = \text{conv}(A, B) \in \mathcal{E}$  die Beträge der Tangentialsprünge  $|\text{[dev } \sigma_{LS} \tau]_E|^2$  in den Punkten  $A, B$  und  $M = \text{mid}(E)$  berechnet und anschließend nach der Simpson-Regel (4.18) aufsummiert. Dabei ist zu beachten, dass für eine korrekte Anwendung der Integrationsregel eine lokale Umnummerierung entsprechend Abbildung 4.2 notwendig ist, um von beiden Dreiecken die jeweils richtige der drei Basisfunktionen zu verwenden.

Für Randkanten werden entsprechend der Definition von  $\gamma^2(E)$  die Tangentialableitungen der Dirichlet-Randdaten verwendet. Dies führt bei den Beispielen zum L- und Schlitzgebiet zu Problemen, da dort die Ableitung der Randdaten auf Grund einer Singularität nicht im Koordinatenursprung ausgewertet werden können. Diese Beiträge von Randkanten, die den Ursprung enthalten, vernachlässigen wir bei der Rechnung.

### 4.3.3 Exakter Fehler

Zur Berechnung des exakten Fehlers

$$\left( \|\sigma - \sigma_{LS-C}\|_{H(\operatorname{div}, \Omega)}^2 + \|D(u - u_{LS-C})\|_0^2 \right)^{1/2}$$

benutzen wir

```
function [error4e,errorSum] = error4eStokesLSC(f,c4n,n4e,
    sigmaApprox,uApprox,sigmaExact1,sigmaExact2,gradExact1,
    gradExact2)
```

und für

$$\left( \|\sigma - \sigma_{LS-NC}\|_{H(\operatorname{div}, \Omega)}^2 + \|D_{NC}(u - u_{LS-NC})\|_0^2 \right)^{1/2}.$$

die Funktion

```
function [error4e,errorSum] = error4eStokesLSNC(f,c4n,n4e,
    sigmaApprox,uApprox,sigmaExact1,sigmaExact2,gradExact1,
    gradExact2) .
```

Für den ersten Summanden spalten wir die Norm bei beiden Methoden in drei Beiträge auf,

$$\|\sigma - \sigma_{LS}\|_0^2 = \|\sigma\|_0^2 \, dx - 2 \int_{\Omega} \sigma : \sigma_{LS} \, dx + \|\sigma_{LS}\|_0^2.$$

Für den ersten Beitrag benutzen wir die `integrate`-Funktion des AFEM-Paketes [Car09a] und für den dritten einen Teil der Steifigkeitsmatrix. Der mittlere Summand ist nach Einsetzen der Basisdarstellungen leicht zu berechnen. Dazu müssen die Komponenten von  $\sigma$  jeweils im Produkt mit den drei Baryzentrischen Koordinaten integriert werden und dann mit den korrekten Vorfaktoren multipliziert und aufsummiert werden. Die Baryzentrischen Koordinaten bestimmen wir dabei durch die folgende lokale Funktion aus der Arbeit [Gal12]

```
function [Lambda1 Lambda2 Lambda3] = barycentricCoords(c4n,
    n4e) .
```

Der zweite Summand der  $H(\operatorname{div})$ -Norm gehört zum Least-Squares-Funktional und wird bereits in der Methode `estimateStokesLSCValueElements` in Abschnitt 4.3.2 berechnet. Dabei ist zu beachten, dass für die exakte Lösung  $-\operatorname{div} \sigma = f$  gilt.

Mit den AFEM-Methoden `error4eP1Energy` bzw. `error4eCREnergy` werden die Beiträge zur Energienorm direkt berechnet.

## 5 Numerische Beispiele

Die im vorigen Kapitel beschriebenen Methoden werden wir an fünf verschiedenen Benchmark-Problemen untersuchen,

1. „colliding flow“
2. L-Gebiet
3. Schlitzgebiet
4. „backward facing step“
5. gekoppeltes Poisson-Problem in zwei Komponenten

Wir benutzen bei jedem Beispiel die konforme Least-Squares-Methode (LS-CFEM) und die nichtkonforme (LS-NCFEM). Wir führen diese Rechnungen jeweils uniform und adaptiv aus, wobei wir zur adaptiven Verfeinerung der Triangulierung in der konformen Methode einmal den Least-Squares-Schätzer (LS-Schätzer) und einmal den  $\eta_1$ -Schätzer (kurz:  $\eta$ -Schätzer) verwenden. Da  $\eta_1$  alle Terme  $\alpha^2(T)$ ,  $\beta^2(E)$  und  $\gamma^2(E)$  enthält, reicht dessen Betrachtung aus, um auch Aussagen über  $\eta_2$  und  $\eta_3$  zu treffen. Auf Grund der Überlegungen aus Abschnitt 3.4 können wir in der nichtkonformen Methode sinnvollerweise nur den  $\eta$ -Schätzer betrachten. Als „bulk chasing“-Parameter haben wir, sofern nichts anderes angegeben worden ist,  $\theta = 0.5$  gewählt. Bei den adaptiven Methoden haben wir bis 300 000 und bei den uniformen bis 100 000 Freiheitsgraden rechnen lassen. Der Grund für diesen Unterschied sind Probleme mit dem Arbeitsspeicher bei uniformer Verfeinerung bis 300 000 Freiheitsgrade, die trotz der in Abschnitt 4.3 beschriebenen Vorkehrungen zur Optimierung nicht behoben werden konnten.

Um die Fehlerschätzer besser vergleichen zu können, werden wir stets alle verfügbaren Werte berechnen. Zur Kontrolle der Rechnungen geben wir stets den exakten Fehler in der  $H(\operatorname{div}, \Omega; \mathbb{R}^{2 \times 2}) \times H^1(\Omega; \mathbb{R}^2)$ -Norm mit an, falls dieser für das jeweilige Problem bekannt ist. Die Werte werden gemeinsam in „convergence history plots“ dargestellt. Von diesen Untersuchungen erhoffen wir uns Aussagen über die Güte der verwendeten Methoden und Fehlerschätzer.

Darüber hinaus werden wir den Quotienten  $q$  aus Gleichung (4.13) in einem gesonderten Diagramm betrachten. Auch hier werden wir die Ergebnisse von uniformer und adaptiver Rechnung gegenüberstellen. Dadurch könnten wir Hinweise auf supercloseness-Eigenschaften im Sinne von [CP12] erhalten.

## 5.1 „Colliding flow“-Beispiel

Sei  $\Omega = (-1, 1)^2$  das Einheitsquadrat und die rechte Seite  $f \equiv 0$ . Die exakte Lösung des „colliding flow“-Beispiels lautet dann für  $x \in \overline{\Omega}$ :

$$u(x) = (20x_1x_2^4 - 4x_1^5, 20x_1^4x_2 - 4x_2^5)^\top \quad \text{und} \quad p(x) = 120x_1^2x_2^2 - 20x_1^4 - 20x_2^4 - 16/3.$$

Die Randdaten werden direkt durch die exakte Lösung induziert. Für  $x \in \partial\Omega$  sei also  $g(x) := u(x)$ .

Die Konvergenzraten (siehe Abbildung 5.1 und 5.2) liegen bei den adaptiven Rechnungen stets bei 0.5. Die nichtkonforme uniforme Methode konvergiert ebenso schnell. Lediglich bei der konformen Methode beobachten wir im Fall der uniformen Verfeinerung eine leicht geringere Rate von 0.4.

Eine Besonderheit zeigt der exakte Fehler. Dort ist ein langer vorasymptotischer Bereich zu beobachten. Die Konvergenzrate liegt zunächst deutlich unterhalb der insgesamt zu beobachtenden 0.5 und wird ab ca. 2000 Freiheitsgraden sprunghaft größer. Der vorasymptotischen Effekt wird dort gebrochen und ist ab etwa 10000 Freiheitsgraden wieder ausgeglichen. Abschließend nähert sich der Graph wieder einer Rate von 0.5 an. Dieses „Locking“-artige Verhalten könnte eine Folge der schlechten Approximationseigenschaft der konformen Räume für die Geschwindigkeit sein. Besonders fällt jedoch auf, dass keiner der Fehlerschätzer dieses Verhalten widerspiegelt. Sie konvergieren stets mit annähernd gleicher Geschwindigkeit.

Um dieses Phänomen genauer zu untersuchen, werden wir in Abschnitt 5.4 ein Problem mit  $p \equiv 0$  betrachten. Dies entspricht einem Poisson-Problem, für das die konformen  $P_1$ -Funktionen bekanntermaßen gute Ergebnisse liefern [Car09b].

Nach dem vorasymptotischen Verhalten nähert sich der Least-Squares-Schätzer dem exakten Fehler jedoch stark an, was eine geringe Überschätzung des Fehlers bedeutet. Der Quotient  $q$  konvergiert mit der gleichen Rate von 0.5 bei adaptiver und uniformer Verfeinerung.

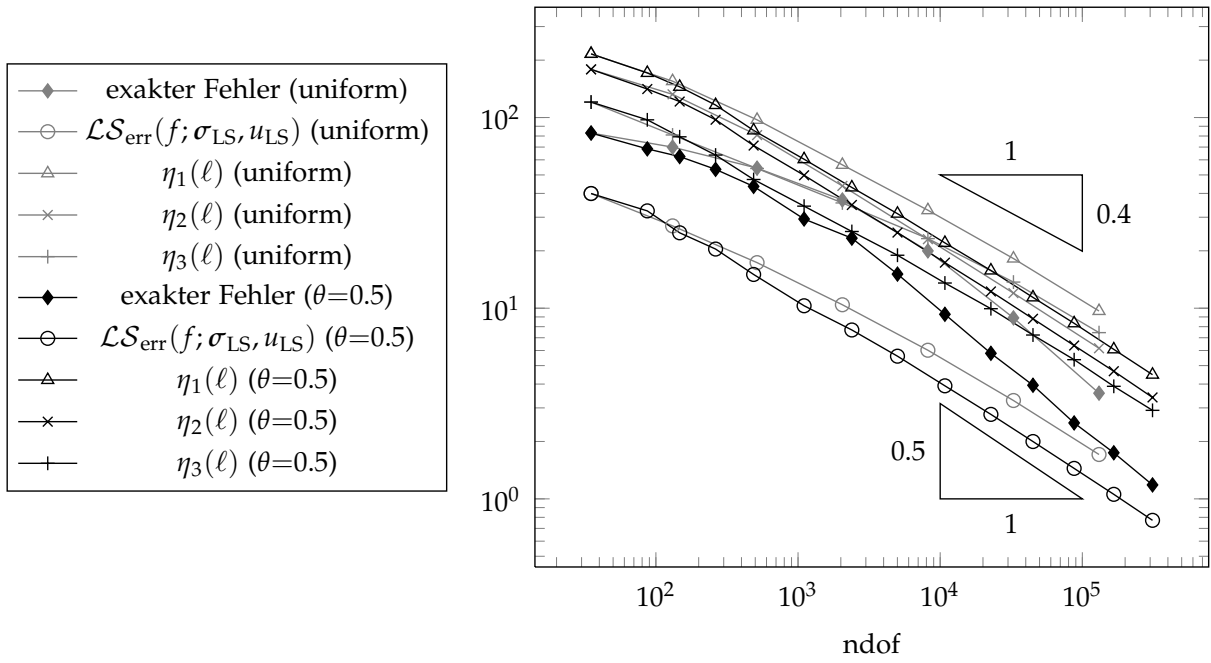
## 5.2 L-Gebiet-Beispiel

Das L-Gebiet lautet  $\Omega = (-1, 1)^2 \setminus [0, 1] \times [-1, 0]$ . Für die rechte Seite  $f \equiv 0$  betrachten wir das Problem zu den Randdaten  $g(x) = u(x)$  für  $x \in \partial\Omega$ . Dabei ist  $u$  die bekannte exakte Lösung, die sich für  $x \in \overline{\Omega}$  in Polarkoordinaten  $(r, \vartheta) \in [0, \infty) \times [0, 3\pi/2]$  wie folgt berechnet,

$$u(r, \vartheta) = r^\alpha \left( (1 + \alpha) \sin(\vartheta) w(\vartheta) + \cos(\vartheta) w'(\vartheta), -(1 + \alpha) \cos(\vartheta) w(\vartheta) + \sin(\vartheta) w'(\vartheta) \right)^\top,$$

wobei

$$\begin{aligned} w(\vartheta) = & \sin((1 + \alpha)\vartheta) \cos(\alpha\omega) / (1 + \alpha) - \cos((1 + \alpha)\vartheta) \\ & - \sin((1 - \alpha)\vartheta) \cos(\alpha\omega) / (1 - \alpha) + \cos((1 - \alpha)\vartheta), \end{aligned}$$



(a) LS-CFEM mit LS-Schätzer

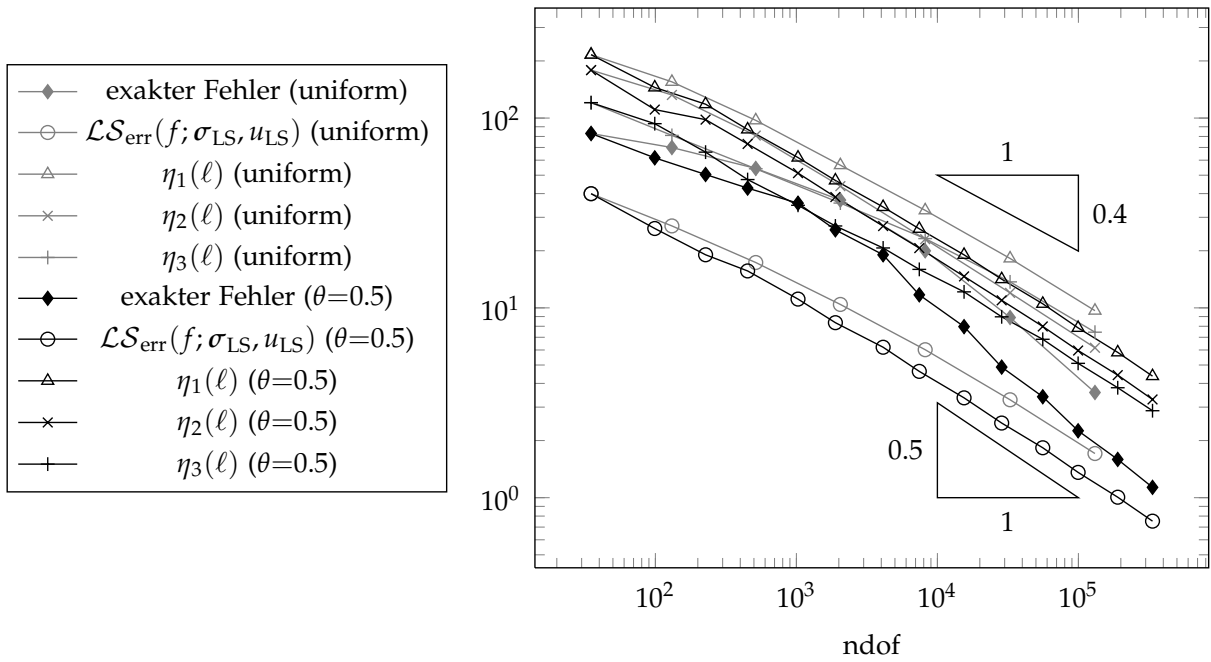

 (b) LS-CFEM mit  $\eta$ -Schätzer

Abbildung 5.1: Ergebnisse der Rechnungen aus dem „colliding flow“-Beispiel I



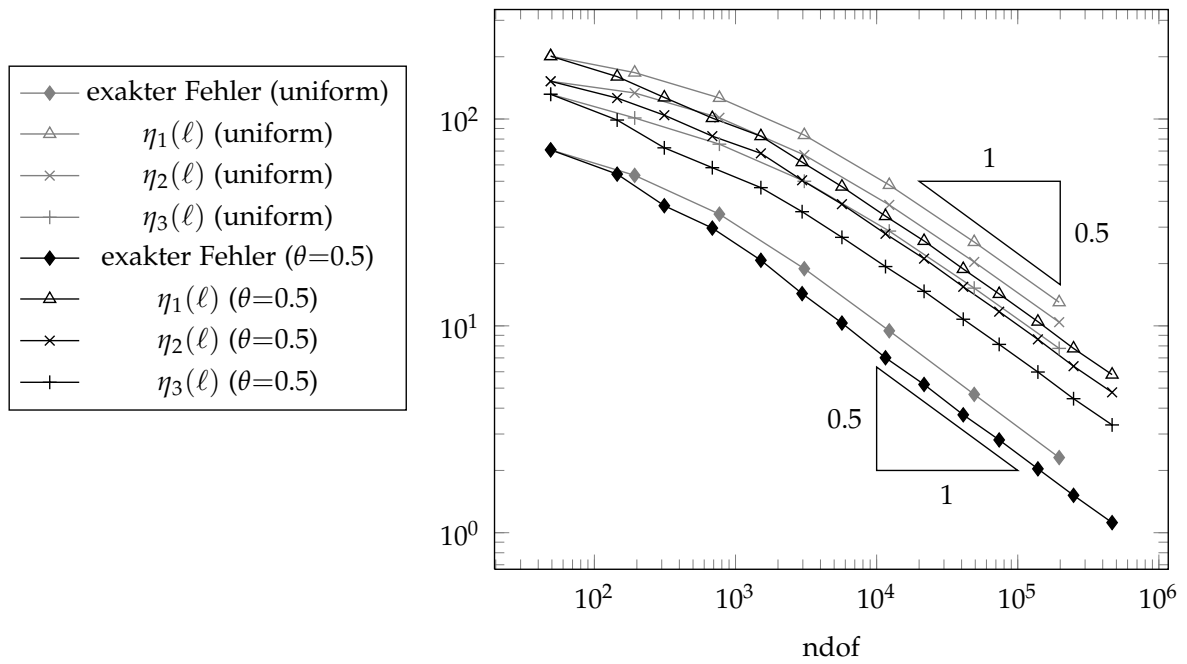
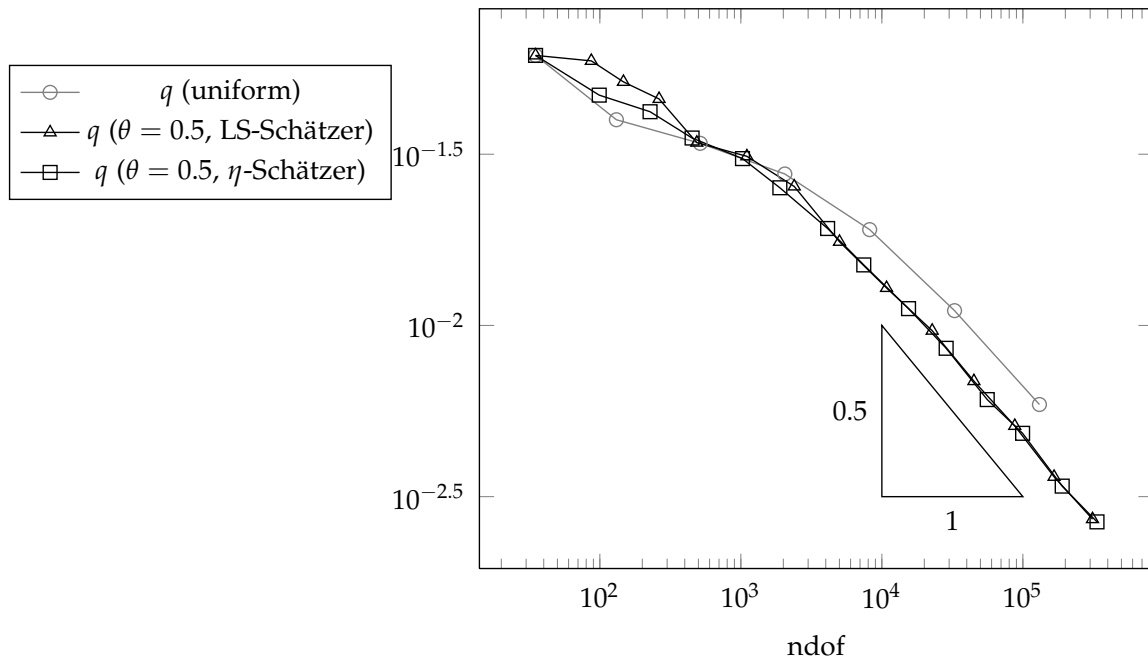
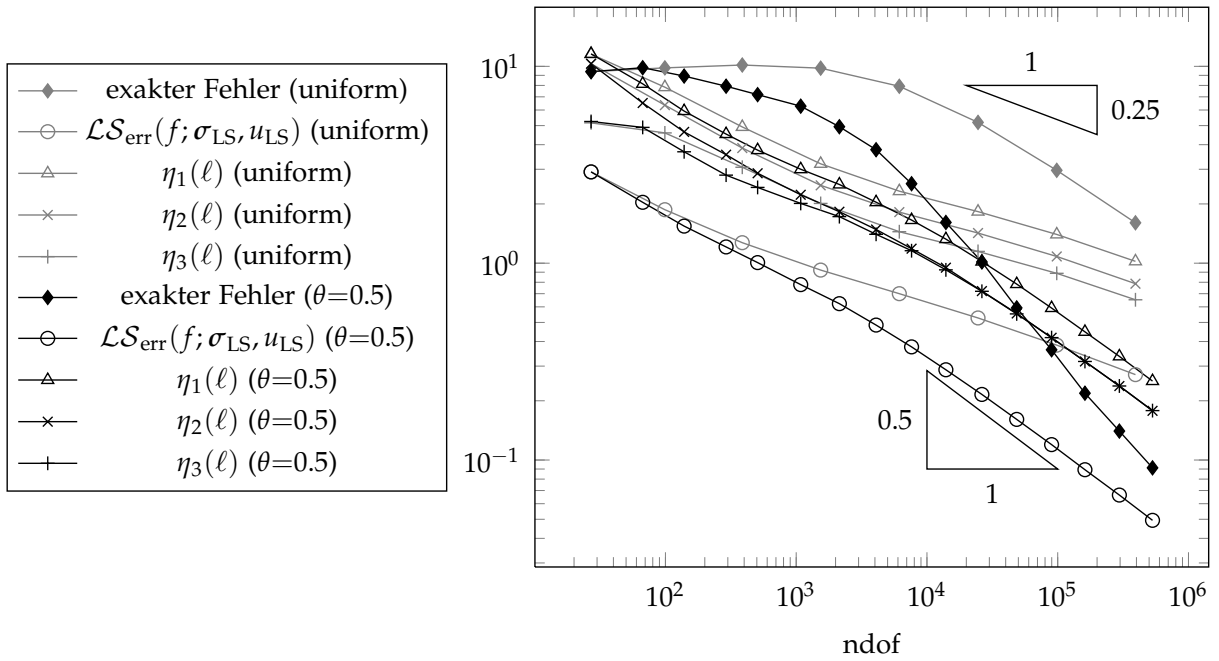

 (a) LS-NCFEM mit  $\eta$ -Schätzer

 (b) Quotient  $q$ 

Abbildung 5.2: Ergebnisse der Rechnungen aus dem „colliding flow“-Beispiel II



(a) LS-CFEM mit LS-Schätzer

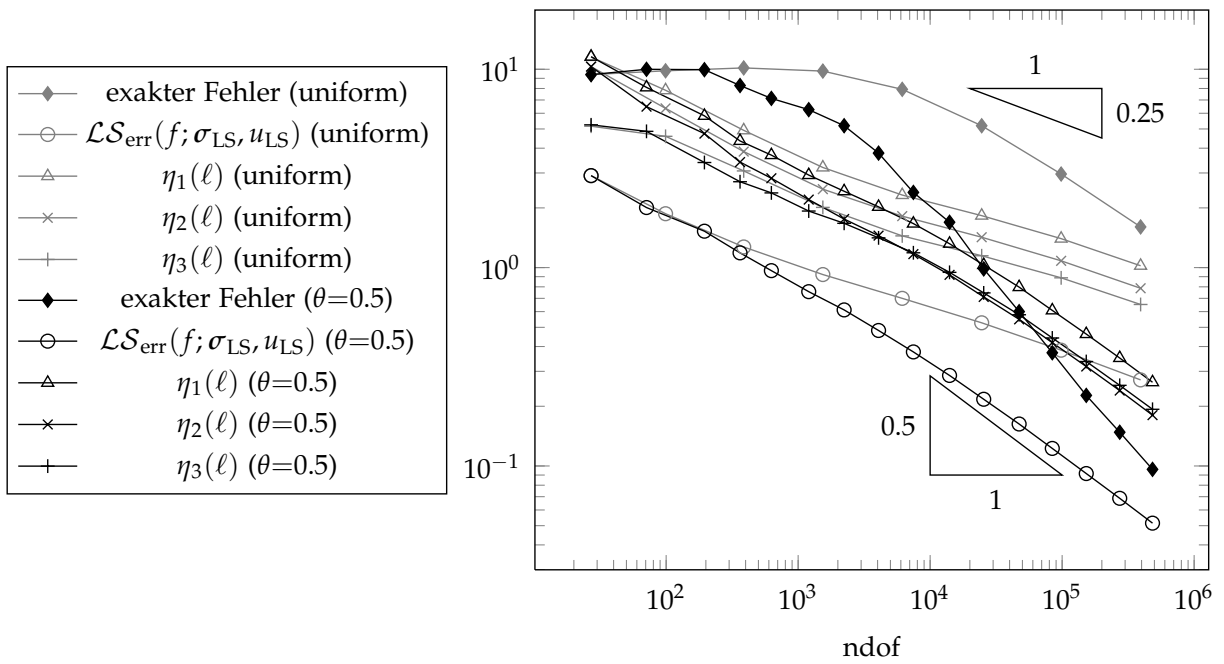

 (b) LS-CFEM mit  $\eta$ -Schätzer

Abbildung 5.3: Ergebnisse der Rechnungen aus dem L-Gebiet-Beispiel I

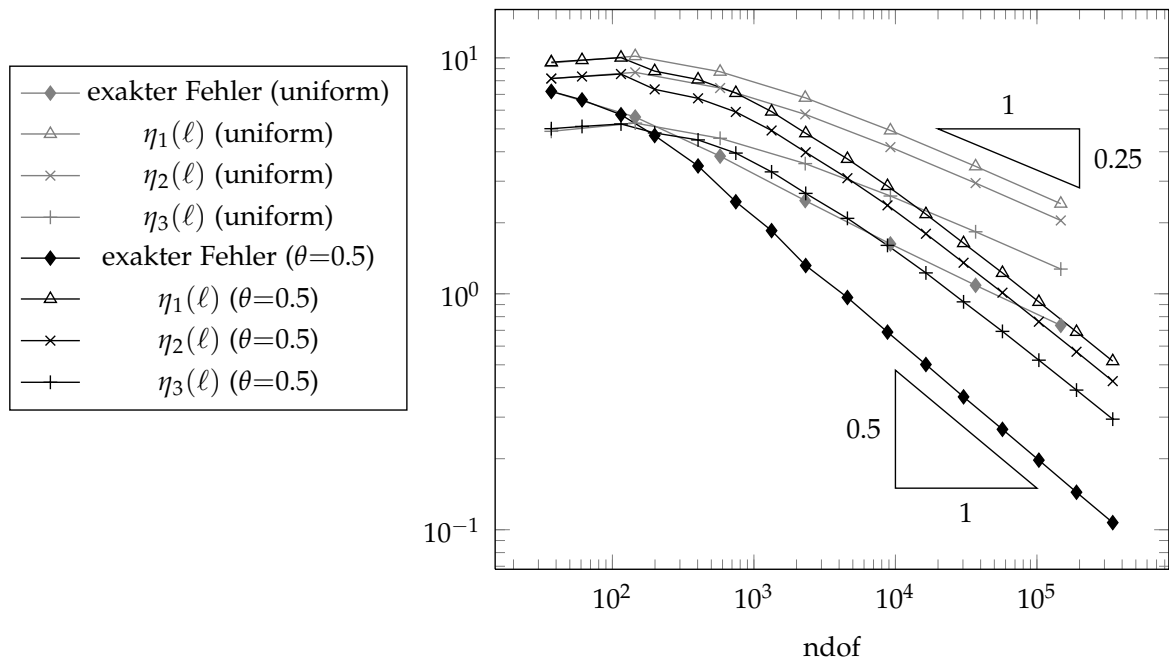
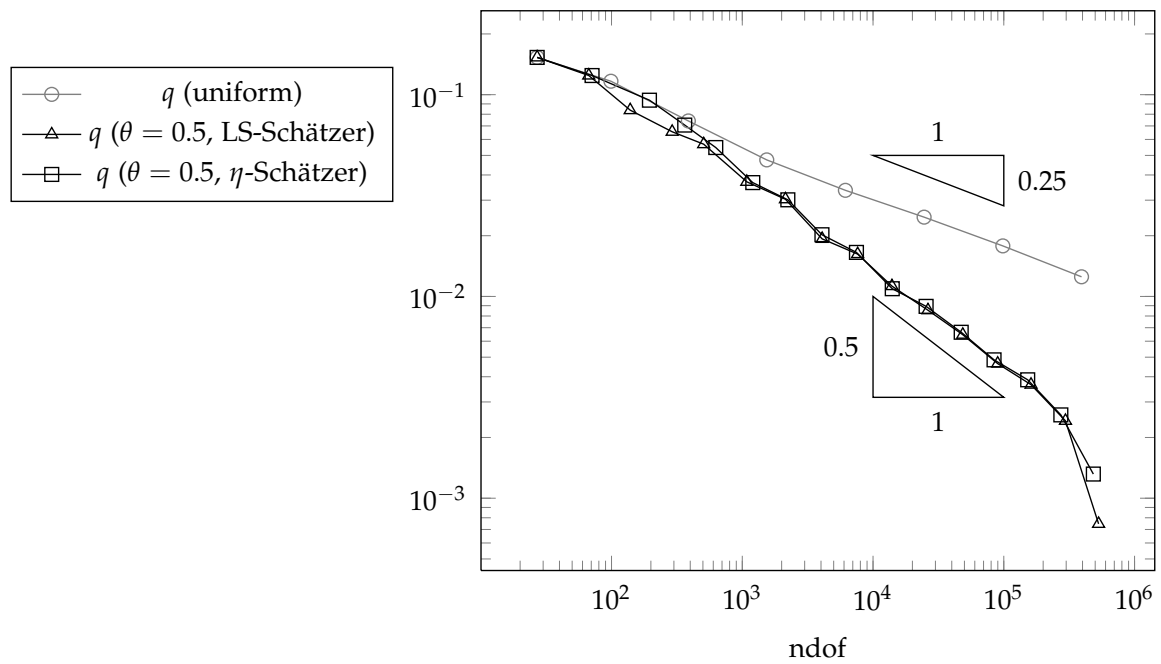

 (a) LS-NCFEM mit  $\eta$ -Schätzer

 (b) Quotient  $q$ 

Abbildung 5.4: Ergebnisse der Rechnungen aus dem L-Gebiet-Beispiel II

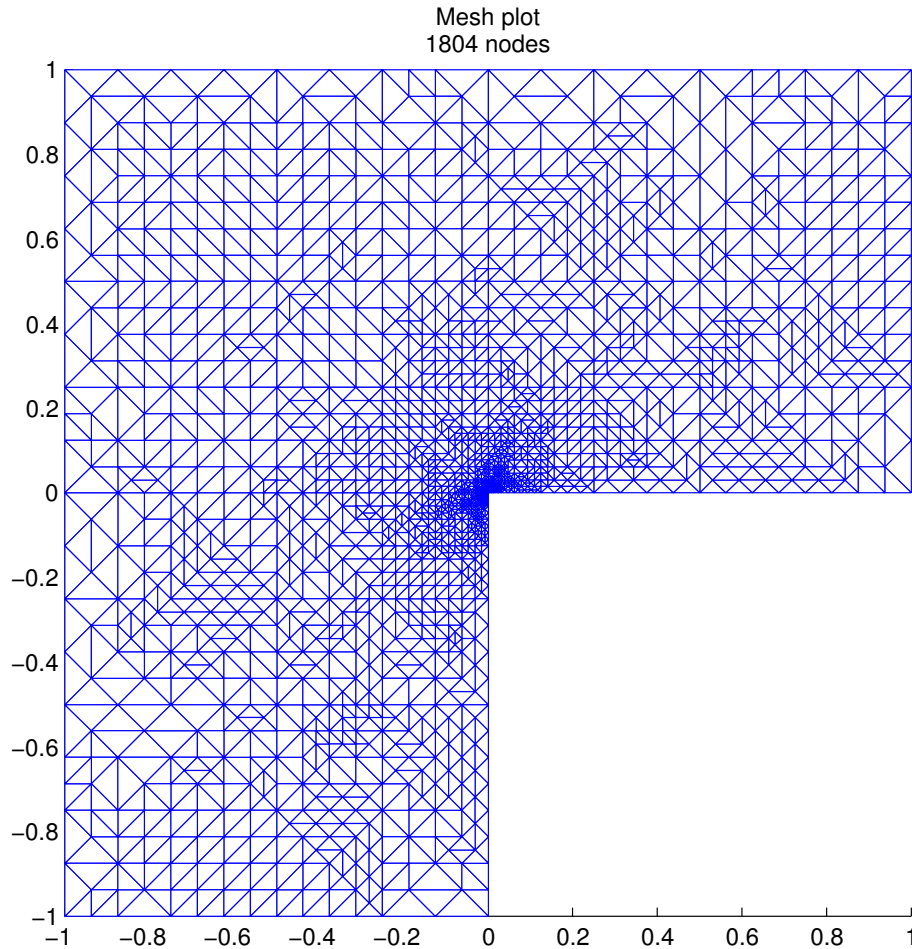


Abbildung 5.5: Triangulierung des L-Gebiets nach der konformen Rechnung mit adaptiver Verfeinerung mit dem LS-Schätzer ( $\theta = 0.5$ ) nach ca. 8 000 Freiheitsgraden

$\omega = 3\pi/2$  und  $\alpha = 0.54448373$  die positive Lösung der Gleichung  $\alpha \sin(2\omega) + \sin(2\alpha\omega) = 0$ .

Bei allen Methoden beobachten wir für den exakten Fehler, die Fehlerschätzer und den Quotienten  $q$  eine Konvergenzrate von 0.25 im Fall von uniformer Verfeinerung und 0.5 bei adaptiver (siehe Abbildung 5.3 und 5.4). Ein vorasymptotisches Verhalten ähnlich dem im „colliding flow“-Beispiel (Abschnitt 5.1) tritt ebenfalls auf. Auch hier scheint das Least-Squares-Funktional den exakten Fehler für eine hohe Zahl von Freiheitsgraden nur um einen geringen Faktor zu überschätzen.

Auf Abbildung 5.5 erkennen wir die Wirkungsweise der adaptiven Verfeinerung. Bei der einspringenden Ecke erwarten wir eine Singularität der Lösung, weshalb die gesuchte Funktion dort nur schwer approximiert werden kann. Die Fehlerschätzer berechnen demnach an dieser Stelle sehr hohe Werte, was der Grund dafür ist, dass auf Grund des AFEM-Algorithmus genau dort auch am stärksten verfeinert wird.

### 5.3 Schlitzgebiet-Beispiel

Wir betrachten das Schlitzgebiet  $\Omega = (-1, 1)^2 \setminus [0, 1] \times \{0\}$  und die rechte Seite  $f \equiv 0$ . Seien  $(r, \vartheta) \in [0, \infty) \times [0, 2\pi)$  die Polarkoordinaten zu  $x \in \overline{\Omega}$ . Die exakte Lösung

$$\begin{aligned} u(r, \vartheta) &= 2\sqrt{r}/3 (\cos(\vartheta/2) - \cos(3\vartheta/2), 3\sin(\vartheta/2) - \sin(3\vartheta/2)), \\ p(r, \vartheta) &= 6/\sqrt{r} \cos(\vartheta/2) \end{aligned}$$

gebe uns wieder die Randdaten vor,  $g(x) = u(x)$  für  $x \in \partial\Omega$ .

Die Konvergenzraten (siehe Abbildung 5.6 und 5.7) sind mit denen aus dem L-Gebiet-Beispiel nahezu identisch, 0.25 für uniforme Verfeinerung und 0.5 für adaptive sowohl in der konformen als auch in nichtkonformen Methode. Auch ein „Locking“-artiges Verhalten, welches mit dem aus dem „colliding flow“-Beispiel (Abschnitt 5.1) vergleichbar ist, ist zu beobachten.

Jedoch können wir über die konforme Methode mit uniformer Verfeinerung keine gesicherten Aussagen treffen, da sich die Konvergenzrate bis zu unserer Grenze von 300 000 Freiheitsgraden stets verschlechtert, sogar bis unter 0.25. Wenn wir uns den exakten Fehler dieser Rechnung anschauen, sehen wir das typische „Locking“-artige Verhalten. Der Graph beginnt jedoch erst bei ca. 10 000 Freiheitsgraden und damit deutlich später als bei den bisherigen Beispielen, sein Konvergenzverhalten zu verbessern. Ich vermute daher, dass der vorasymptotische Bereich bei dieser Rechnung weitaus größer ist als bisher. Das müsste jedoch noch weiter untersucht werden.

Das Konvergenzverhalten spiegelt sich ebenso in dem Graphen des Quotienten  $q$  wider.

### 5.4 Poisson-Problem-Beispiel

Für das *Poisson-Problem* betrachten wir das Einheitsquadrat  $\Omega = (-1, 1)^2$  mit der rechten Seite  $f$ , definiert für  $x = (x_1, x_2)^\top \in \Omega$  vermöge

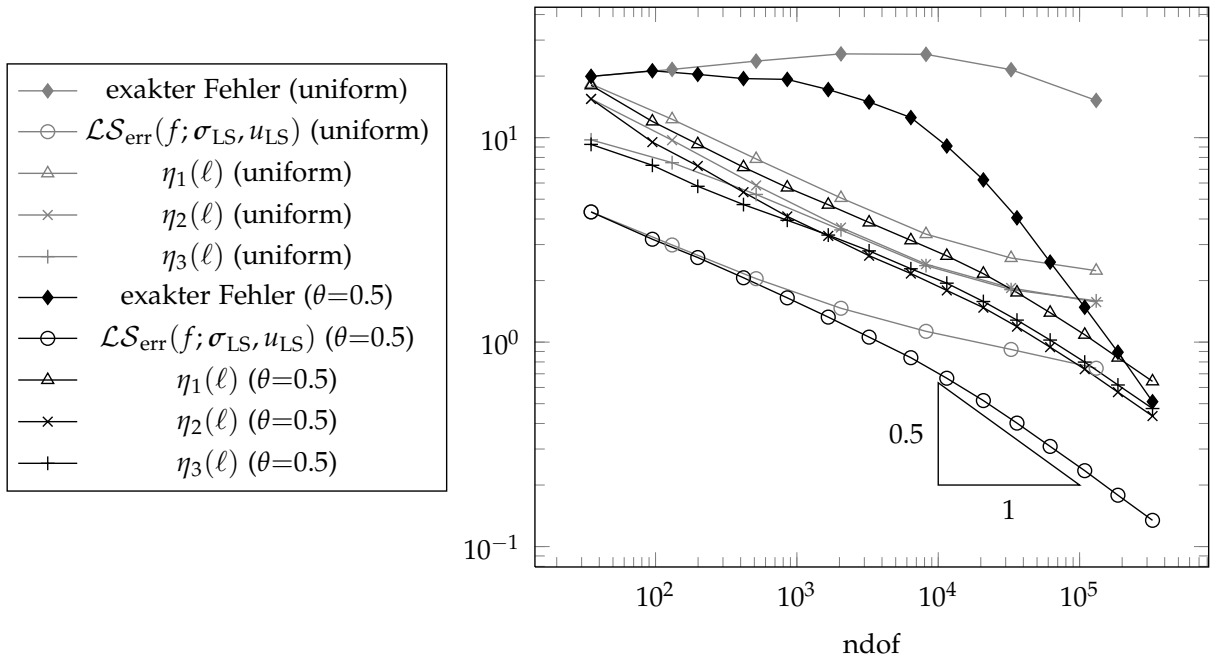
$$f(x) := (2x_2, -2x_1)^\top.$$

Die exakte Lösung lautet dann für  $x \in \overline{\Omega}$ ,

$$u(x) = (-x_1^2 x_2, x_1 x_2^2)^\top \quad \text{und} \quad p(x) = 0.$$

Da der Druck gleich Null ist, können wir dieses Problem als Poissonproblem betrachten. Dabei ist zu beachten, dass die Geschwindigkeit divergenzfrei ist.

Wie bereits in Abschnitt 5.4 angedeutet, wollen wir mit diesem Beispiel das bisher beobachtete veränderte vorasymptotische Verhalten genauer untersuchen. Bei allen Rechnungen konvergieren die Graphen zur Fehlerbetrachtung und des Quotienten  $q$  mit einer Rate von 0.5 (siehe Abbildung 5.8 und 5.9). Ein schlechtes vorasymptotisches Verhalten wie in den anderen Beispielen ist nicht zu erkennen.



(a) LS-CFEM mit LS-Schätzer

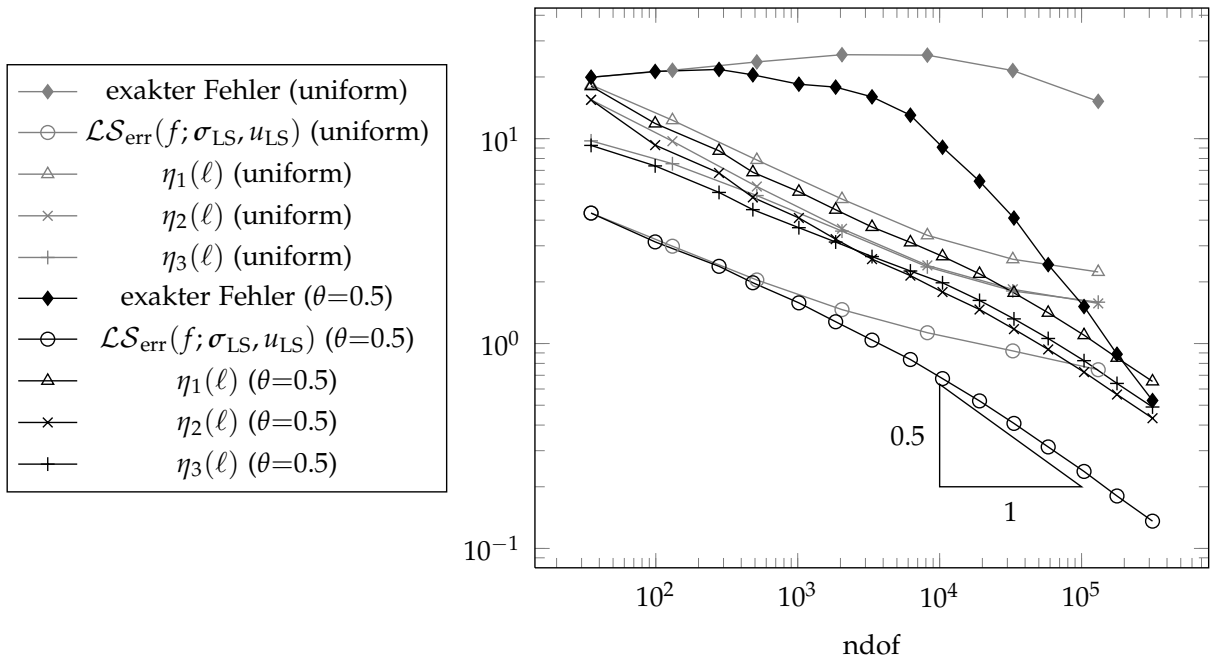

 (b) LS-CFEM mit  $\eta$ -Schätzer

Abbildung 5.6: Ergebnisse der Rechnungen aus dem Schlitzgebiet-Beispiel I

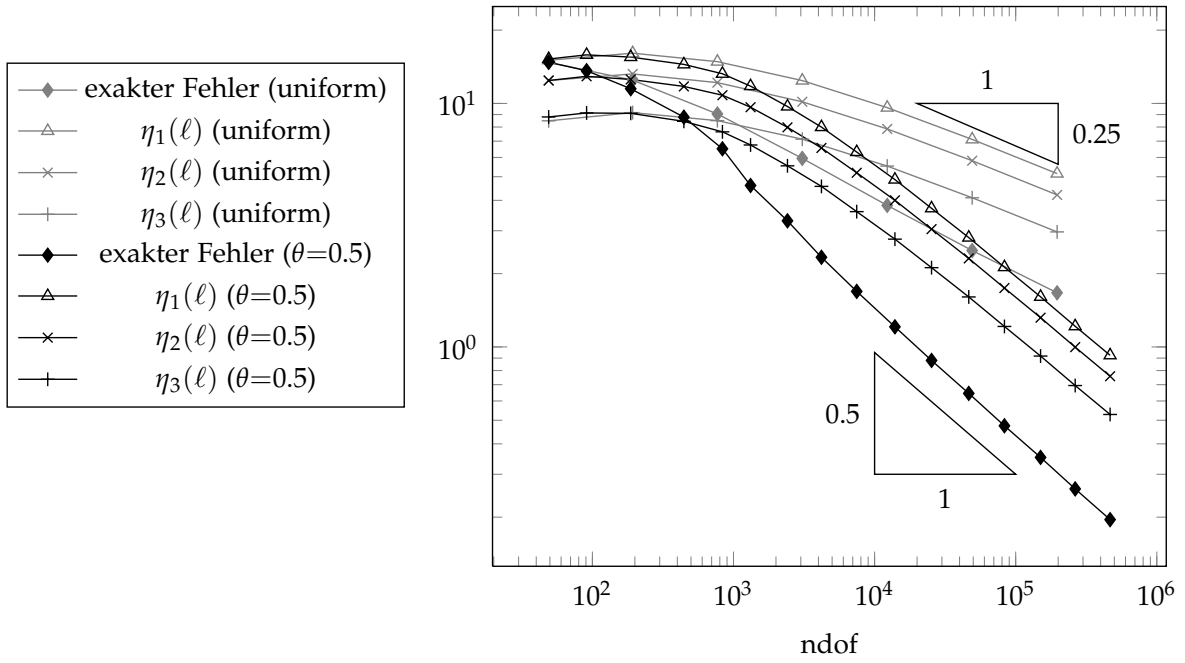
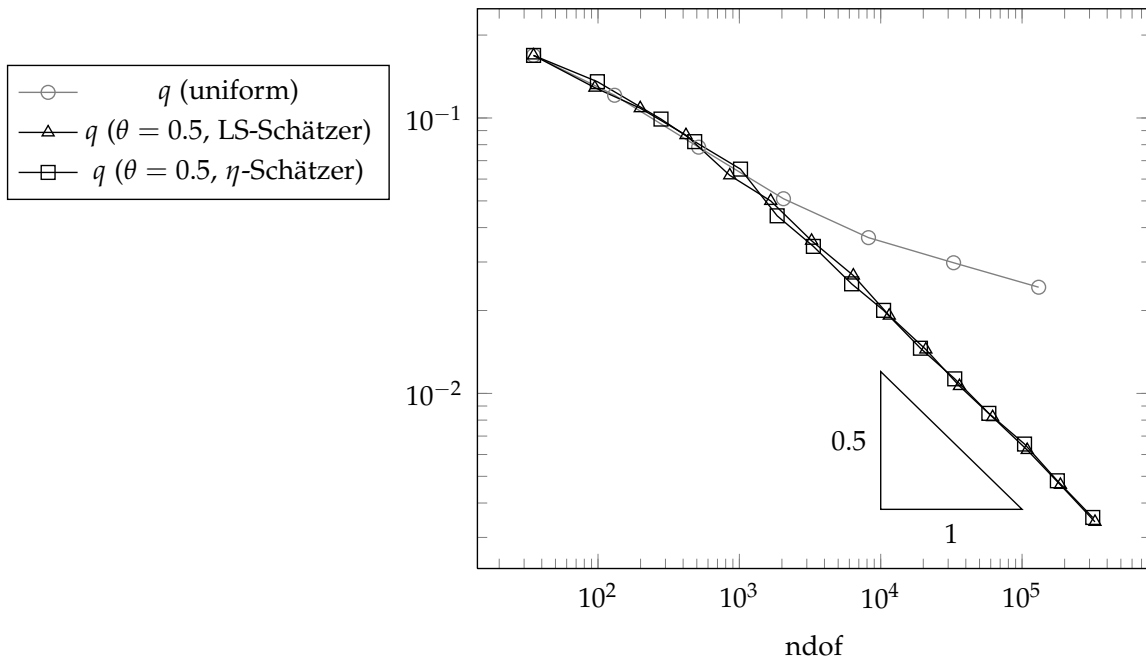
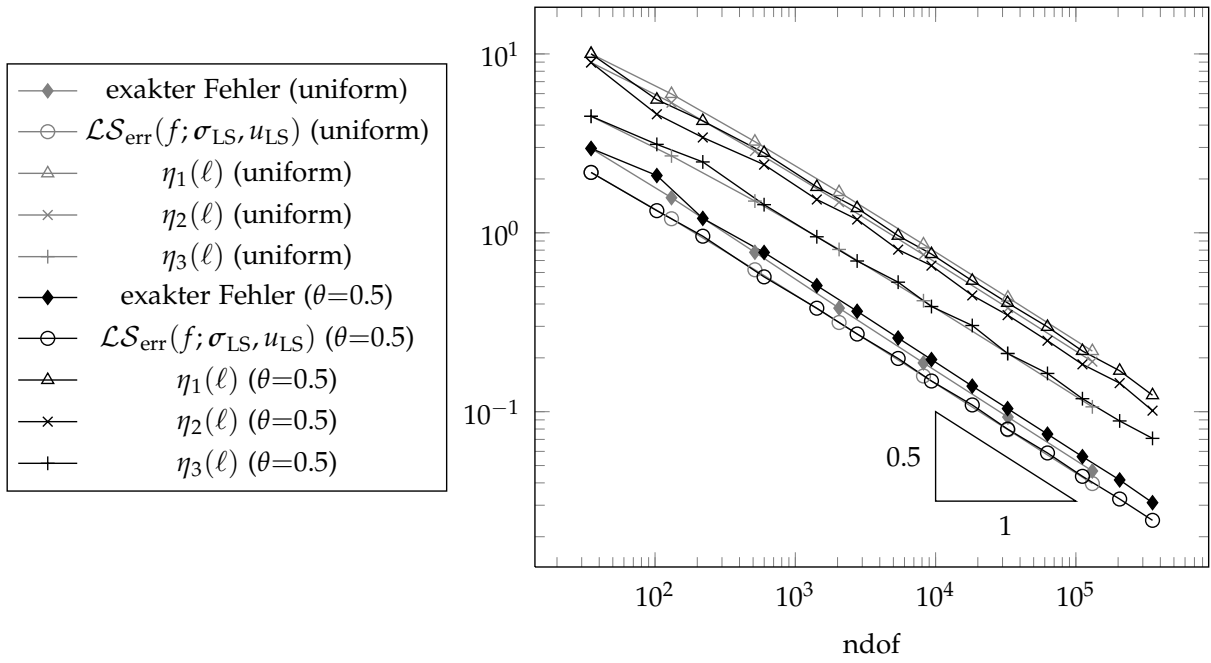

 (a) LS-NCFEM mit  $\eta$ -Schätzer

 (b) Quotient  $q$ 

Abbildung 5.7: Ergebnisse der Rechnungen aus dem Schlitzgebiet-Beispiel II



(a) LS-CFEM mit LS-Schätzer

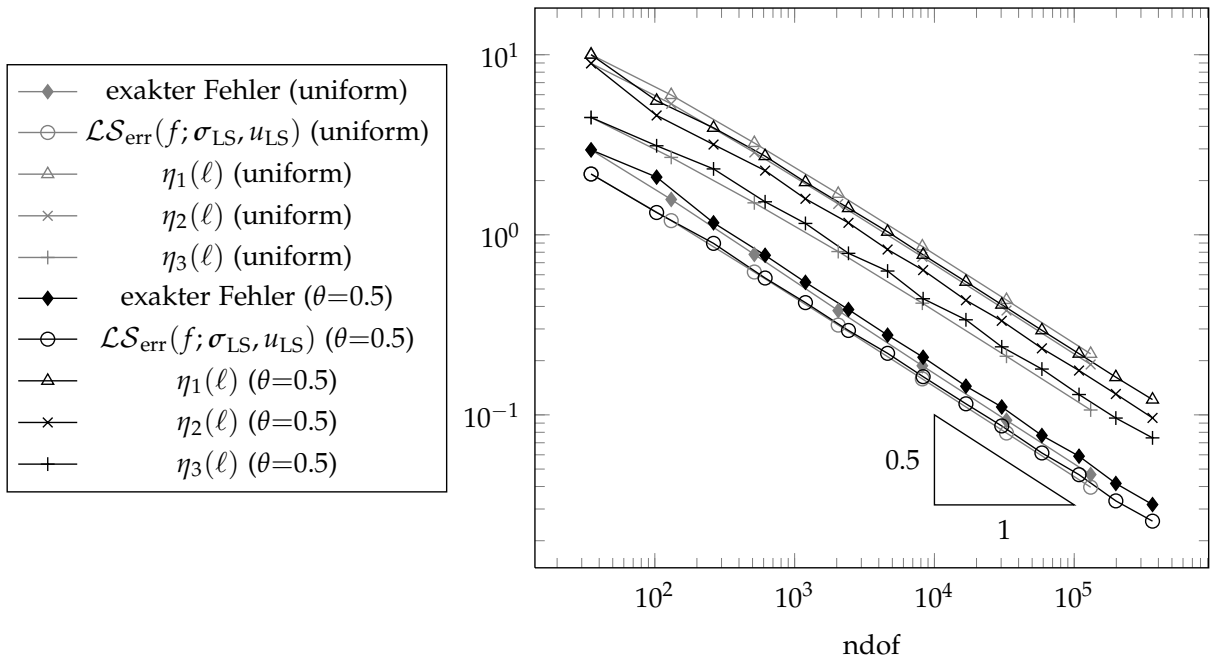

 (b) LS-CFEM mit  $\eta$ -Schätzer

Abbildung 5.8: Ergebnisse der Rechnungen aus dem Poisson-Beispiel I



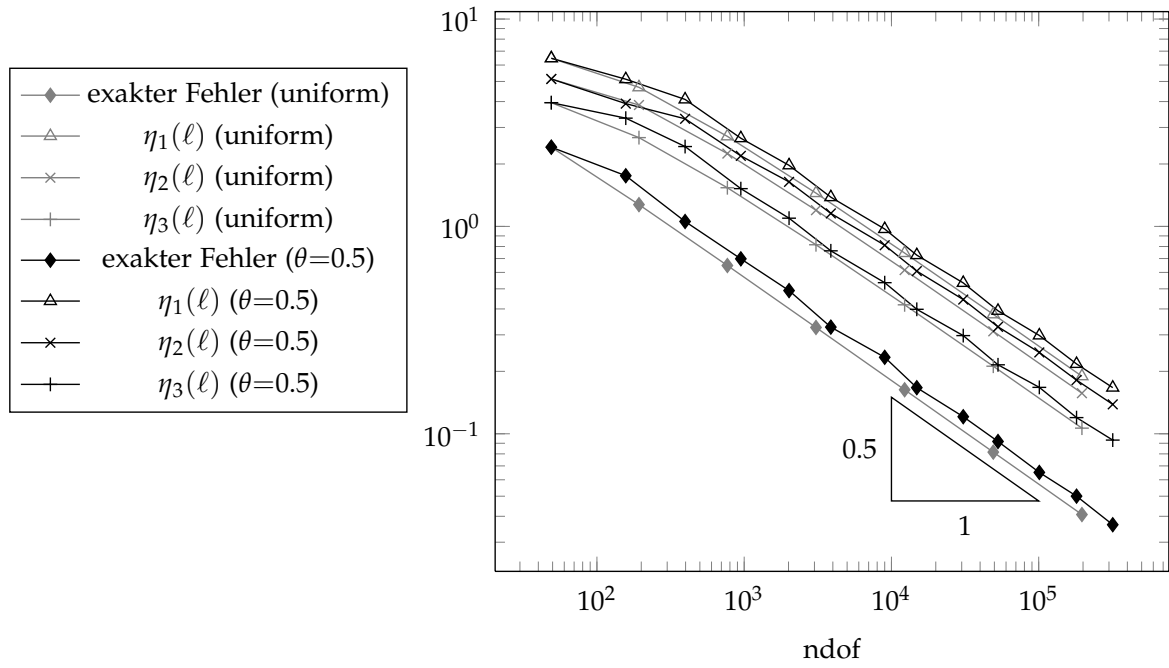
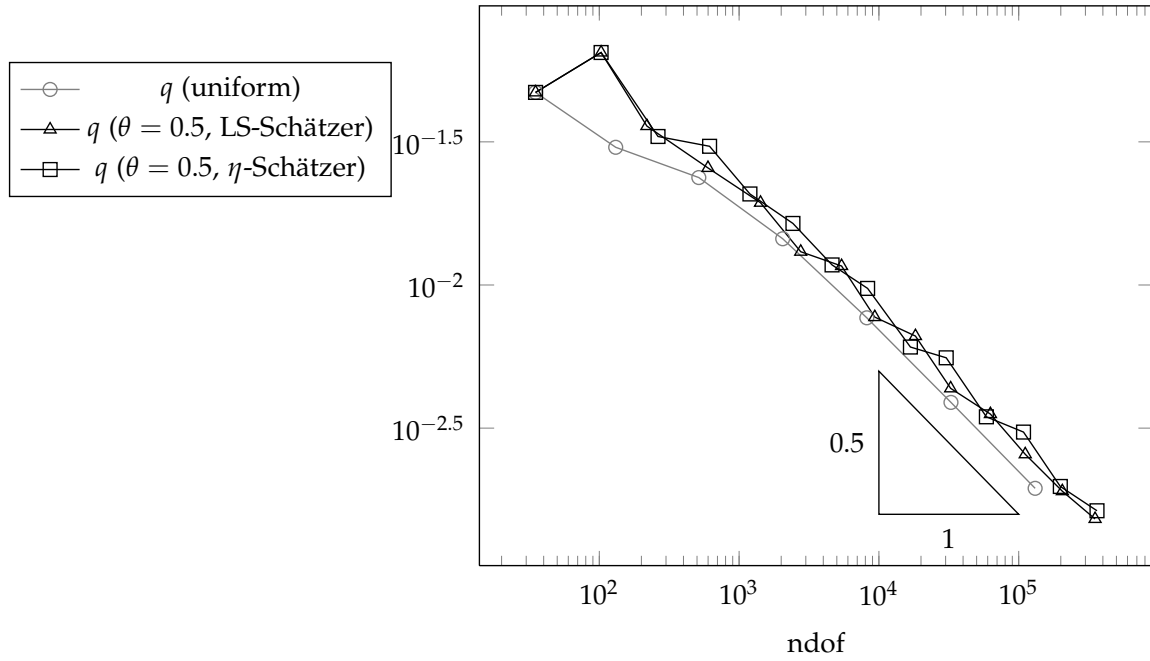

 (a) LS-NCFEM mit  $\eta$ -Schätzer

 (b) Quotient  $q$ 

Abbildung 5.9: Ergebnisse der Rechnungen aus dem Poisson-Beispiel II

### 5.5 „backward facing step“-Beispiel

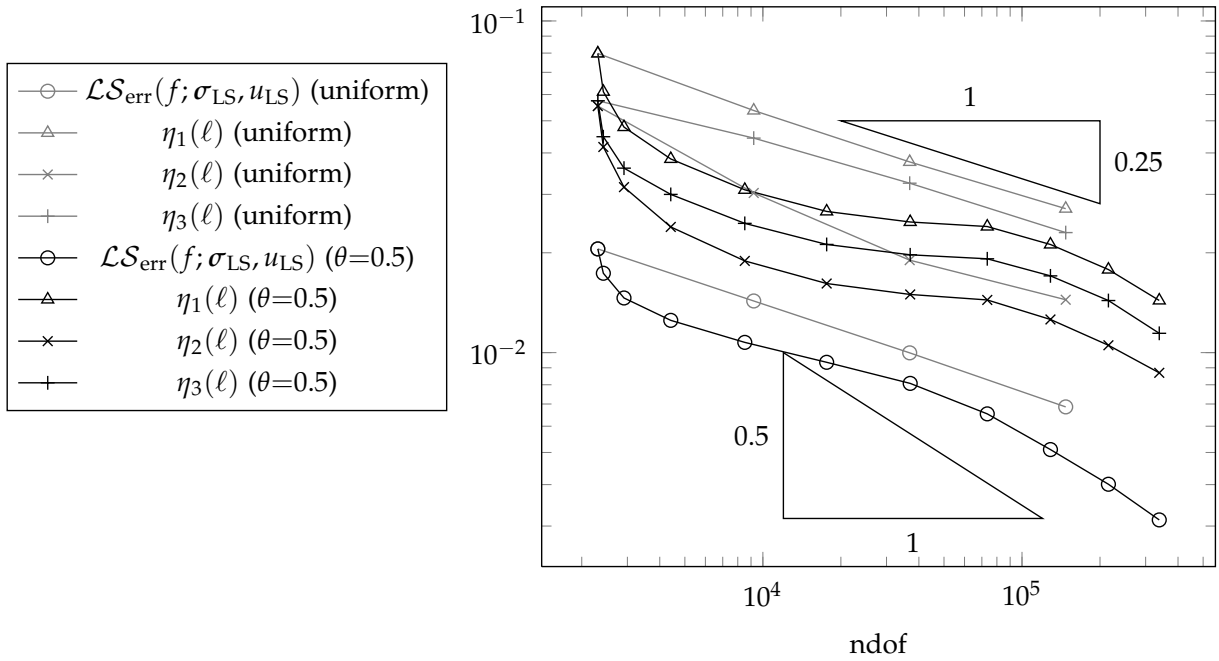
Zum Abschluss untersuchen wir noch ein Standardbeispiel ohne bekannte exakte Lösung. Dazu gehört das Gebiet  $\Omega = ((-2, 8) \times (-1, 1)) \setminus ((-2, 0) \times (-1, 0))$  und die rechte Seite  $f \equiv 0$ . Die Dirichlet-Randdaten lauten für  $x \in \partial\Omega$ ,

$$g(x) = \begin{cases} (0, 0)^\top & \text{für } -2 < x_1 < 8, \\ 1/10 (-x_2(x_2 - 1), 0)^\top & \text{für } x_1 = -2, \\ 1/80 (-(x_2 - 1)(x_2 + 1), 0)^\top & \text{für } x_1 = 8. \end{cases}$$

Wieder ist nach einem Abschnitt mit schlechterer Konvergenz eine Annäherung an die bekannte optimalen Konvergenzrate von 0.5 im adaptiven Fall in den „convergence history plots“ 5.10 und 5.11 zu beobachten. Auffällig ist, dass zunächst eine schnellere Konvergenzrate erreicht wird und diese dann stark abflacht. Bei uniformer Verfeinerung liegen die Werte bei 0.25 bei der konformen Methode und bei 0.4 bei der nichtkonformen. Wie auch schon beim „colliding flow“-Beispiel konvergiert hier die nichtkonforme Methode schneller. Der Quotient  $q$  spiegelt auch hier die Raten der Fehlerschätzer wider.

Das Verhalten der Lösung bei diesem Problem scheint jedoch stark von der Anfangstriangulierung abzuhängen. So hat es sich als sinnvoll herausgestellt, das Gebiet vor der Berechnung zweimal uniform zu verfeinern, um die Randdaten ausreichend aufzulösen. Da das Konvergenzverhalten in den Abbildungen 5.10 und 5.11 noch nicht ideal erscheint, habe ich mich entschlossen, das „backward facing step“-Beispiel ein zweites Mal mit einem anderen „bulk chasing“-Parameter  $\theta = 0.1$  berechnen zu lassen. Die Ergebnisse finden sich in den Graphen 5.12 und 5.13.

Das Konvergenzverhalten ist ähnlich dem aus der ersten Rechnung, jedoch beginnt die optimale Konvergenz etwas früher, bei circa  $10^{4.3}$  statt  $10^{4.7}$  Freiheitsgraden. Um das genauer zu untersuchen, betrachten wir die erzeugten Triangulierungen in Abbildung 5.14. Bereits nach 8 000 Freiheitsgraden ist zu erkennen, dass die Methode mit  $\theta = 0.1$  um die erwartete Singularität an der einspringenden Ecke stärker verfeinert. Beim ersten Beispiel ist eine Verfeinerung lediglich an den Rändern mit nichthomogenen Randdaten zu sehen.



(a) LS-CFEM mit LS-Schätzer

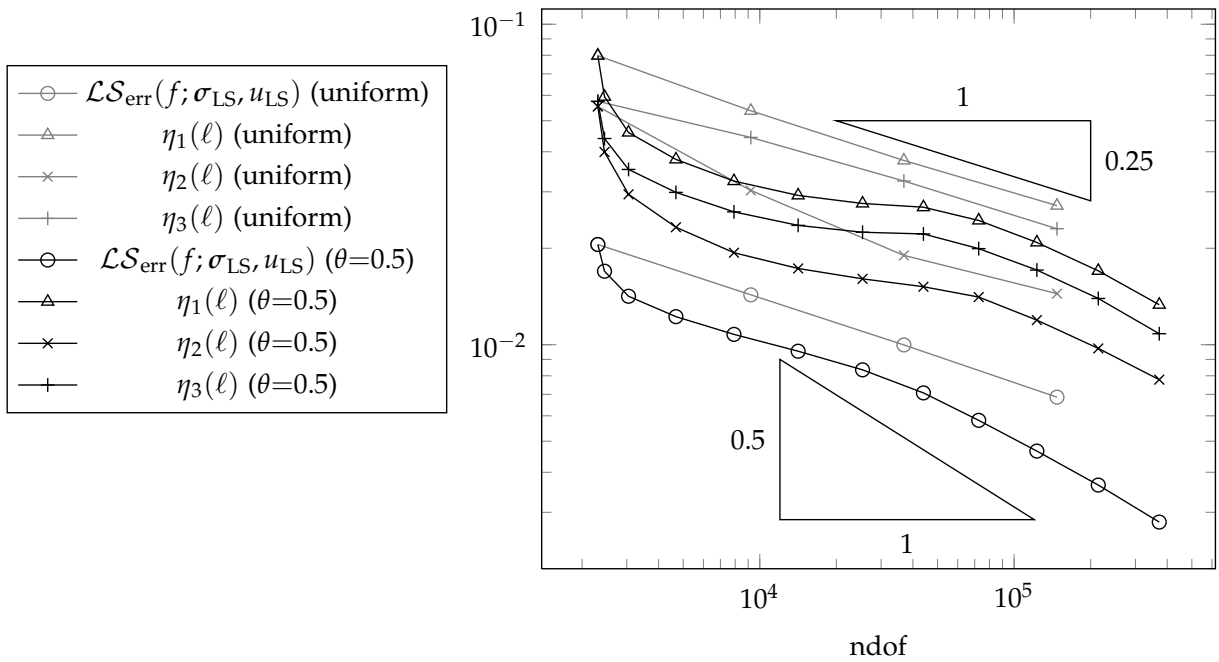

 (b) LS-CFEM mit  $\eta$ -Schätzer

 Abbildung 5.10: Ergebnisse der Rechnungen aus dem „backward facing step“-Beispiel mit bulk-Parameter  $\theta = 0.5$  I

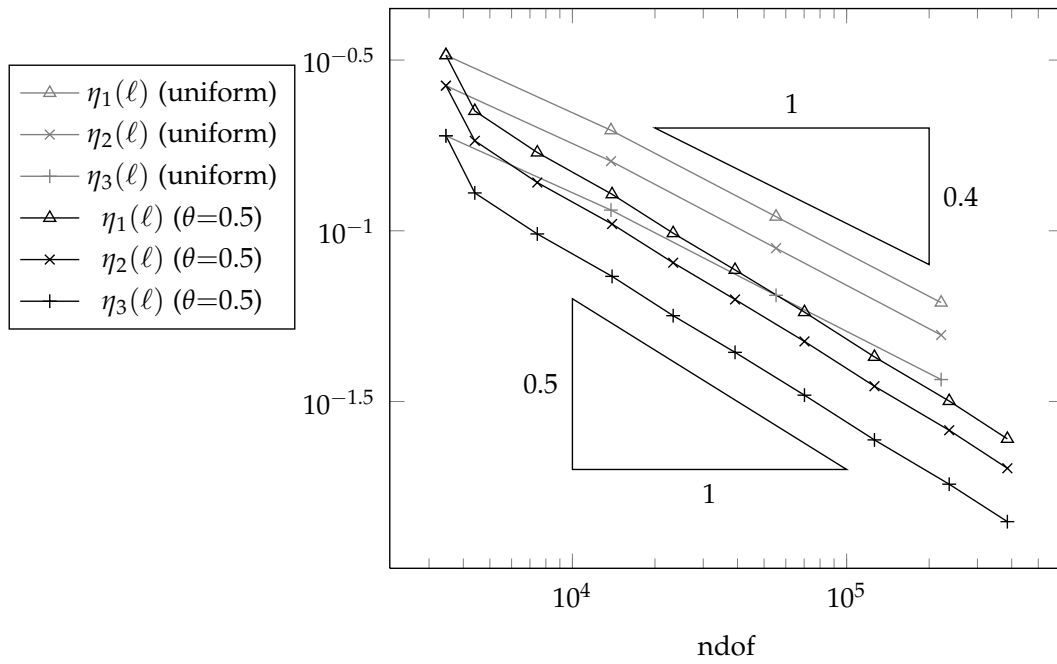
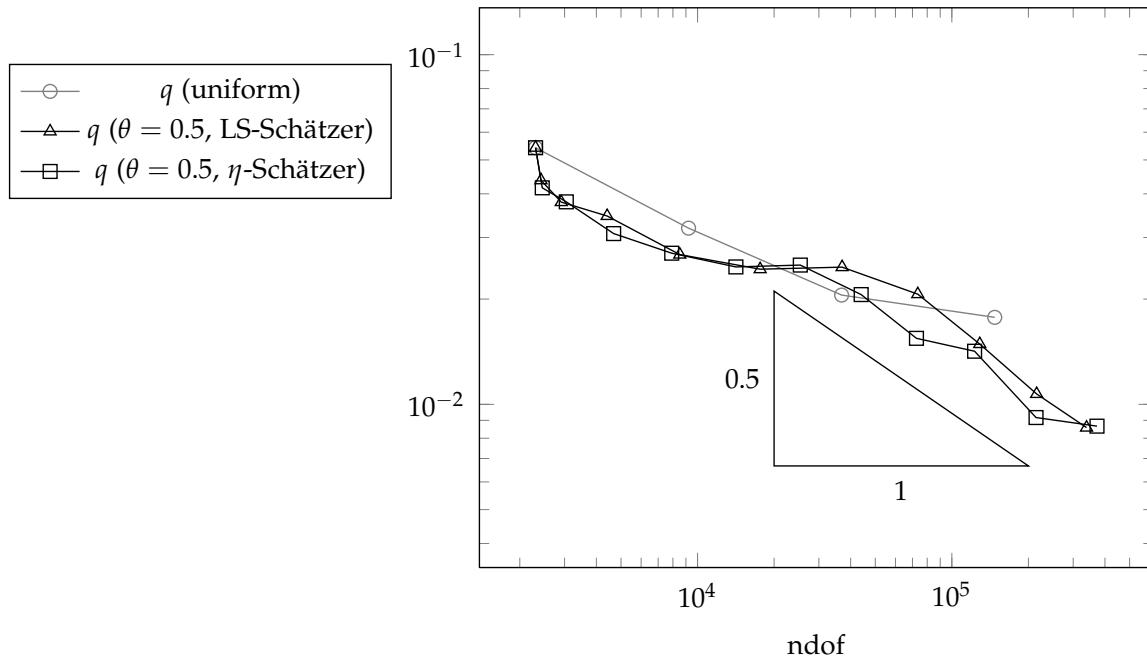
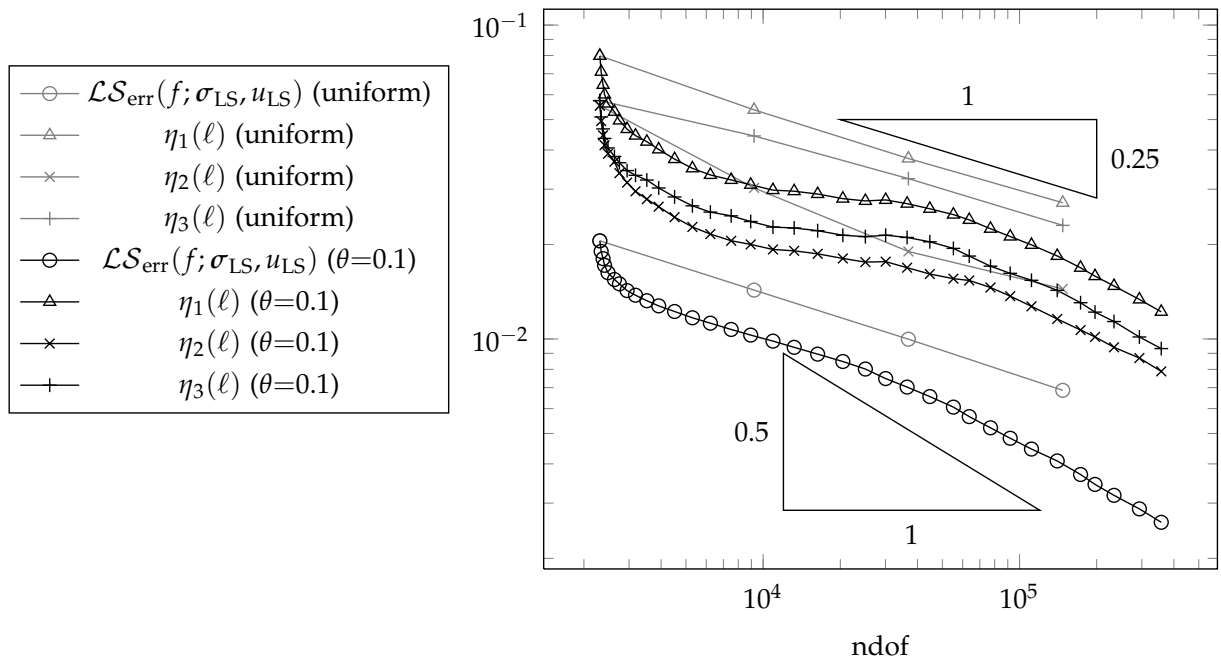

 (a) LS-NCFEM mit  $\eta$ -Schätzer

 (b) Quotient  $q$ 

 Abbildung 5.11: Ergebnisse der Rechnungen aus dem „backward facing step“-Beispiel mit bulk-Parameter  $\theta = 0.5$  II



(a) LS-CFEM mit LS-Schätzer

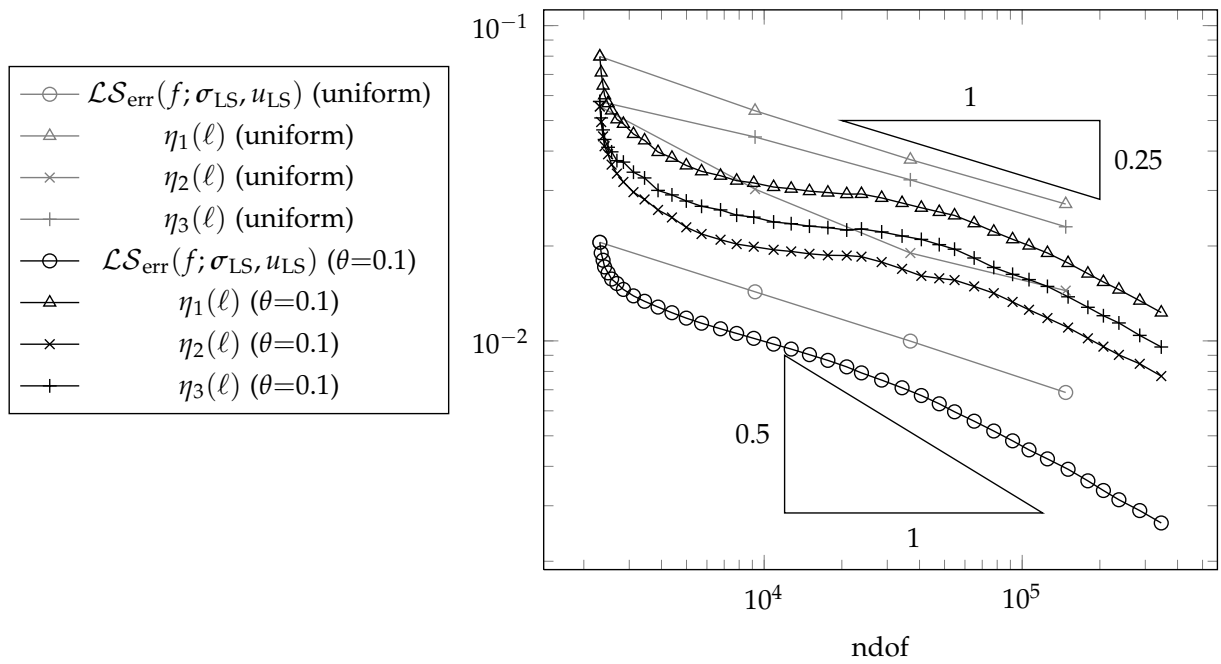

 (b) LS-CFEM mit  $\eta$ -Schätzer

 Abbildung 5.12: Ergebnisse der Rechnungen aus dem „backward facing step“-Beispiel mit bulk-Parameter  $\theta = 0.1$  I

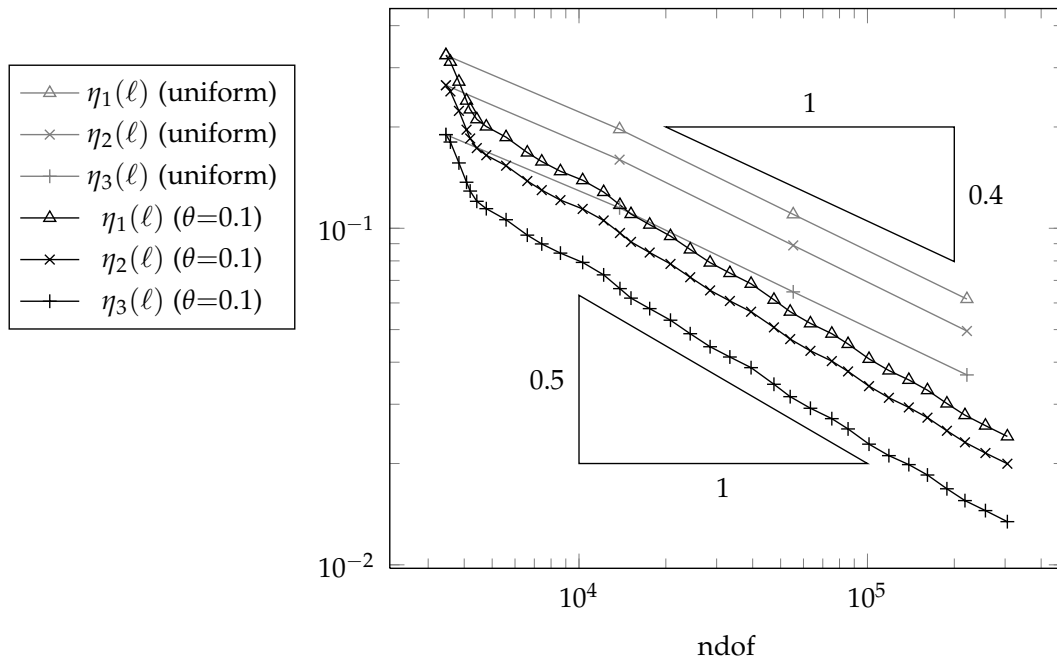
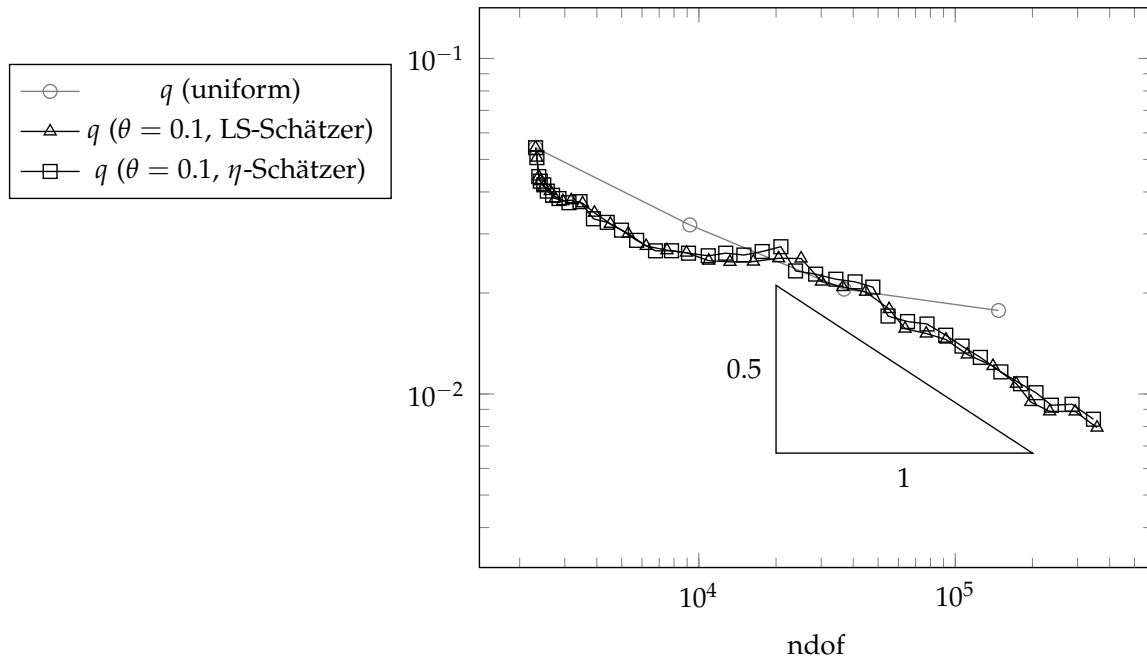

 (a) LS-NCFEM mit  $\eta$ -Schätzer

 (b) Quotient  $q$ 

 Abbildung 5.13: Ergebnisse der Rechnungen aus dem „backward facing step“-Beispiel mit bulk-Parameter  $\theta = 0.1$  II

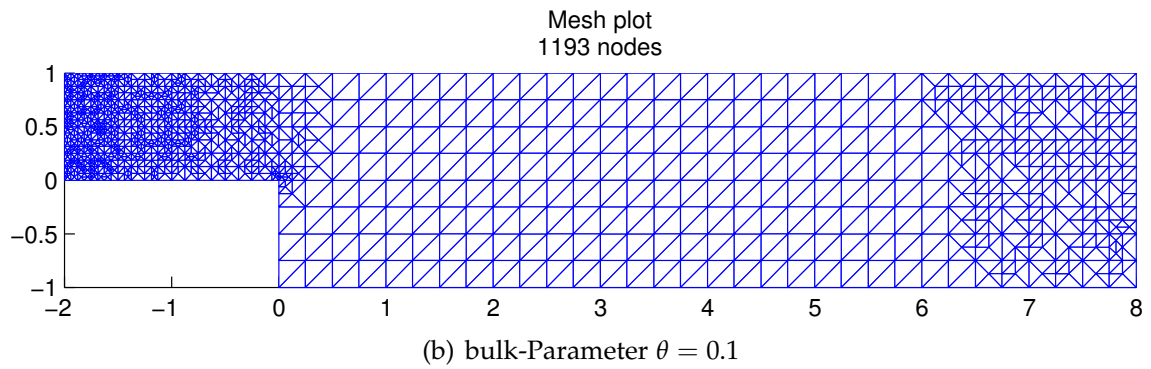
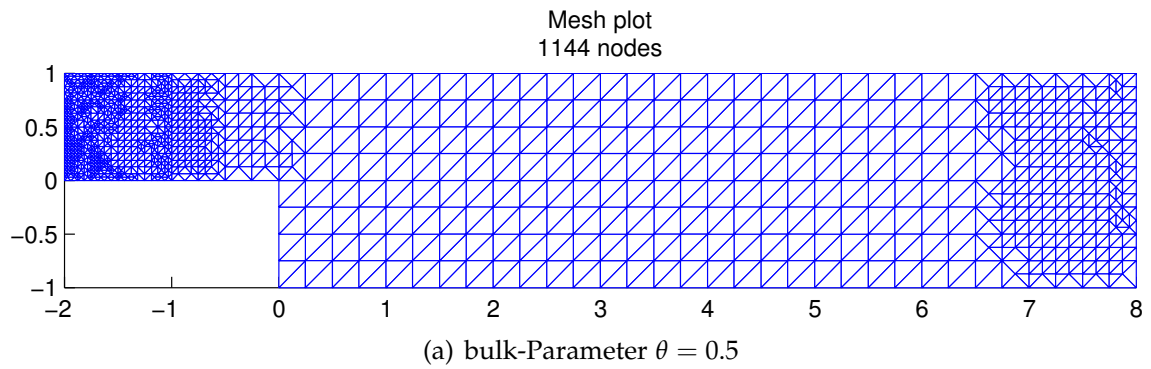


Abbildung 5.14: Unterschiedliche Triangulierungen aus dem „backward facing step“-Beispiel mit konformem Element und dem Least-Squares-Schätzer nach einer Rechnung mit bis zu 8 000 Freiheitsgraden

## 6 Diskussion der numerischen Ergebnisse

Zunächst stellen wir fest, dass die implementierten Lösungsmethoden qualitativ unterschiedliche Lösungen für die Probleme liefern. So sind die Konvergenzraten der nichtkonformen Least-Squares-Methode bei adaptiver Verfeinerung in jedem Beispiel optimal. Wie nach den Überlegungen aus Abschnitt 3.4 auch zu erwarten ist, stimmen sie mit denen der nichtkonformen Methode mit dem Crouzeix-Raviart Finiten Element aus [BH12] überein.

### 6.1 Adaptivität

Außerdem zeigt sich in den Ergebnissen, dass die adaptive Verfeinerungsstrategie stets zu optimalen Konvergenzraten und somit im Vergleich zur uniformen zu häufig besseren führt. Die gleichen Konvergenzraten im „colliding flow“ und Poisson-Problem können mit der Glattheit der exakten Lösung erklärt werden.

### 6.2 „Locking“-artiges vorasymptotisches Verhalten

Entgegen dieser positiven Ergebnisse zeichnet sich bei den Lösungen der konformen Methode für die Beispiele „colliding flow“, L-Gebiet und Schlitzgebiet in Form der sehr verschlechterten Konvergenz der Lösung bis zu den ersten 10 000 Freiheitsgraden ein ungewollter Nebeneffekt ab. Die Betrachtung des Poisson-Problem-Beispiels hat jedoch unseren Anfangsverdacht bestärkt, dass dieses Verhalten mit der schlechten Approximationseigenschaft des konformen  $P_1$ -Raumes für die Geschwindigkeit zusammenhängt. Bei dem Poisson-Problem tritt, wie auf Grund der guten Eigenschaften der konformen Funktionen für dieses Problem zu vermuten war, dieses Phänomen nicht auf. Darüber hinaus haben wir das „Locking“-artige Verhalten nur bei der konformen Least-Squares-Methode beobachtet, was diese Erklärung weiter untermauert. Es wäre jedoch auch möglich, dass dieser Effekt empfindlich von der Anfangstriangulierung abhängt, was wir im Rahmen dieser Arbeit aber nicht weiter untersuchen können.

### 6.3 Eigenschaften der Fehlerschätzer

Insbesondere ist zu beobachten, dass die vier untersuchten Schätzer stets die gleiche Konvergenzordnung wie der exakte Fehler besitzen. Das belegt die theoretisch gesicherte Effizienz und Zuverlässigkeit des Least-Squares-Schätzers und ist ein starkes Indiz für die Annahme, dass die untersuchten  $\eta$ -Schätzer ebenfalls diese Eigenschaften besitzen. Insbesondere liegt die Vermutung nahe, dass die für Schätzterme  $\eta_1$ ,  $\eta_2$  und  $\eta_3$  im Stokes-Problem ähnliche theoretische Aussagen bewiesen werden können wie für ihre Analoga aus [CP12].



Offen bleibt jedoch, warum die Fehlerschätzer nicht das „Locking“-artige Verhalten des exakten Fehlers widerspiegeln.

## 6.4 supercloseness

Ein weiteres Hauptanliegen dieser Arbeit, neben der Betrachtung der  $\eta$ -Schätzer, ist die Untersuchung der supercloseness-Eigenschaft der Least-Squares-Methoden. In [CP12] wurde für die Least-Squares-Lösungen  $\mathbf{p}_{LS} \in RT_0(\mathcal{T})$  und  $u_{LS} \in P_1(\mathcal{T}) \cap C_0(\bar{\Omega})$  des Poisson-Modellproblem folgende Eigenschaft bewiesen,

$$\|\operatorname{div} \mathbf{p}_{LS} + \Pi_0 f\|_0 \lesssim^{H^s} \|\mathbf{p}_{LS} - \mathbf{D} u_{LS}\|_0, \quad (6.1)$$

wobei  $H > 0$  die maximale Netzweite in  $\mathcal{T}$  ist und  $1/2 < s \leq \min\{1, \pi/\omega\}$  für den größten Innenwinkel  $\omega$  des polygonal berandeten Gebietes  $\Omega$  die Regularität der Lösung beschreibt. Diese Eigenschaft zeigt, dass der Flussterm den Divergenzterm dominiert, und stellt „supercloseness“-Ergebnisse wie [CP12, Korollar 3.2] in Aussicht.

Unsere Ergebnisse lassen einen solchen Zusammenhang auch für die Least-Squares-Methode bei den Stokes-Gleichungen als möglich erscheinen. Die Verbesserung der Konvergenz des Quotienten  $q$  legt zudem die Vermutung nahe, dass (6.1) nicht nur für die maximale Netzweite gilt, sondern auch die Konvergenzrate auf adaptiven Gittern widerspiegelt.

## 6.5 Fazit und Ausblick

Durch die Arbeit von [CP12] motiviert, haben wir uns zwei zentrale Untersuchungsgegenstände gewählt. Für beide haben wir durch die konforme und die nichtkonforme Least-Squares Finite-Elemente-Methode bei den Stokes-Gleichungen positive Ergebnisse erhalten. In einem möglichen nächsten Schritt gilt es, diese experimentell beobachteten Eigenschaften Effizienz und Zuverlässigkeit der  $\eta$ -Schätzer und „supercloseness“ der Least-Squares-Methode, sofern möglich, auch theoretisch zu beweisen.

Außerdem bieten die Überlegungen aus Abschnitt 3.4 Anlass dazu, den Zusammenhang der dort betrachteten Methoden näher zu untersuchen. So könnte versucht werden, für die Least-Squares-Methode ähnliche Repräsentationsformeln wie (??) zu finden.

Abschließend bleibt zu sagen, dass die starke Verzögerung bei der Erreichung von ordentlichen Konvergenzraten in der konformen Least-Squares-Methode sehr problematisch ist und die Methode für die Praxis untauglich erscheinen lässt. Die Methode der Wahl sollte daher eher die nichtkonforme Least-Square-Methode sein. Für sie haben wir ab Beginn der Rechnungen optimale Konvergenzraten beobachtet. Da sie äquivalent zur NCFEM ist, kann der erhöhte Rechenaufwand (vgl. Größe der Gleichungssysteme) nicht gerechtfertigt werden.

## **Danksagung**

Ich möchte mich sehr bei Prof. Carstensen und Dietmar Gallistl für die hilfreiche und umfangreiche Betreuung dieser Arbeit und die erhellenden und anregenden Gespräche bedanken.

# Literatur

- [BC05] BAHRIAWATI, C. ; CARSTENSEN, C.:  
Three MATLAB implementations of the lowest-order Raviart-Thomas  
MFEM with a posteriori error control.  
In: *Comput. Methods Appl. Math.* 5 (2005), Nr. 4, S. 333–361 (elektronisch).  
–  
ISSN 1609–4840
- [BF91] BREZZI, F. ; FORTIN, M.:  
*Computational Mathematics*. Bd. 15: *Mixed and Hybrid Finite Element Methods*.  
First Edition.  
New York : Springer, 1991. –  
ix+350 S.
- [BG09] BOCHEV, P. ; GUNZBURGER, M.:  
*Applied Mathematical Sciences*. Bd. 166: *Least-Squares-Finite Element Methods*.  
First Edition.  
New York : Springer, 2009. –  
xxii+660 S.
- [BH12] BRINGMANN, P. ; HELLWIG, F.:  
*Die Stokes-Gleichungen mit nicht-konformen Finiten Elementen nach Crouzeix  
und Raviart – Praktikumsdokumentation zu BZQ III: Numerik partieller  
Differentialgleichungen*.  
lokal verfügbar, 2012
- [Bra03] BRAESS, D.:  
*Finite Elemente. Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*.  
Vierte, erweiterte und überarbeitete Auflage.  
Berlin : Springer, 2003. –  
xvi+342 S.
- [Car09a] CARSTENSEN, C.:  
*AFEM-Softwarepaket der Arbeitsgruppe „Numerische Analysis“*.  
Institut für Mathematik der Humboldt-Universität zu Berlin, 2009
- [Car09b] CARSTENSEN, C.:  
*Yonsei Lectures at the WCU Department Computational Science and Engineering on Finite Element Method*.  
lokal verfügbar, 2009
- [CGK<sup>+</sup>10] CARSTENSEN, C. ; GEDICKE, J. ; KERN, L. ; NEUMANN, J. ; RABUS, H. ;  
ROZOVA, M.:

*AFEM-Dokumentation.*

lokal verfügbar, 2010

- [CGS12] CARSTENSEN, C. ; GALLISTL, D. ; SCHEDENSACK, M.:  
*Quasi optimal Adaptive Pseudostress approximation of the Stokes equations.*  
lokal verfügbar, 2012
- [CLW04] CAI, Z. ; LEE, B. ; WANG, P.:  
Least-squares methods for incompressible Newtonian fluid flow: linear  
stationary problems.  
In: *SIAM J. Numer. Anal.* 42 (2004), Nr. 2, S. 843–859 (elektronisch). –  
ISSN 0036–1429
- [CP12] CARSTENSEN, C. ; PARK, E.-J.:  
*Convergence and optimality of Adaptive Least Squares Finite Element Methods.*  
lokal verfügbar, 2012
- [Gal11] GALLISTL, D.:  
*Stokes-Dokumentation.*  
lokal verfügbar, 2011
- [Gal12] GALLISTL, D.:  
*Über nichtkonforme Finite-Elemente-Diskretisierungen der biharmonischen  
Gleichung,* Humboldt-Universität zu Berlin, Diplomarbeit, 2012

# Anhang

```
/
├── common
│   └── P0TangentJump.m
├── data
│   ├── Stokes_data_bf.m
│   ├── Stokes_data_cavity.m
│   ├── Stokes_data_colliding.m
│   ├── Stokes_data_Lshape.m
│   ├── Stokes_data_noPress.m
│   └── Stokes_data_slit.m
├── estimate
│   ├── estimateStokesLSCEtaElements.m
│   ├── estimateStokesLSCValueElements.m
│   └── estimateStokesLSNCEtaElements.m
├── integrate
│   ├── error4eStokesLSC.m
│   └── error4eStokesLSNC.m
├── solve
│   ├── solveLSCStokes.m
│   └── solveLSNCStokes.m
├── afem.zip.m
├── quicktest.m
├── README
└── StokesLS.m
```

Abbildung 6.1: Verzeichnisstruktur der Software im Anhang