

OPTIMAL FINITE ELEMENT APPROXIMATIONS OF MONOTONE SEMILINEAR ELLIPTIC PDE WITH SUBCRITICAL NONLINEARITIES

FLORIAN SPICHER[✉] AND THOMAS P. WIHLER*[✉]

Abstract. We study iterative finite element approximations for the numerical approximation of semilinear elliptic boundary value problems with monotone nonlinear reactions of subcritical growth. The focus of our contribution is on an optimal *a priori* error estimate for a contractive Picard type iteration scheme on meshes that are locally refined towards possible corner singularities in polygonal domains. Our analysis involves, in particular, an elliptic regularity result in weighted Sobolev spaces and the use of the Trudinger inequality, which is instrumental in dealing with subcritically growing nonlinearities. A series of numerical experiments confirm the accuracy and efficiency of our method.

Mathematics Subject Classification. 47J25, 65J15, 65N30.

Received April 15, 2025. Accepted October 24, 2025.

1. INTRODUCTION

On a bounded, open, non-degenerate polygonal domain $\Omega \subset \mathbb{R}^2$ with a finite number of straight edges $\Gamma_1, \dots, \Gamma_m$, we consider the semilinear boundary value problem

$$\begin{aligned} -\Delta u + g(\cdot, u) &= f && \text{in } \Omega \\ u &= 0 && \text{on } \Gamma_{\mathcal{D}} \\ \partial_{\mathbf{n}} u &= 0 && \text{on } \Gamma_{\mathcal{N}}. \end{aligned} \tag{1}$$

Here, $f \in L^p(\Omega)$, for $p \in (1, \infty)$, is a given source function, and $g : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ represents a possibly nonlinear reaction term that is measurable with respect to its first argument, and continuously differentiable with respect to its second one, *i.e.*, it is a Carathéodory function; we suppose that the partial derivative in the second variable, denoted by $g_u \equiv \partial_u g$, features *subcritical growth*, *i.e.*, for any $\sigma > 0$, it holds

$$\lim_{|\xi| \rightarrow \infty} \frac{|g_u(\mathbf{x}, \xi)|}{\exp(\sigma \xi^2)} = 0 \quad (\text{uniformly in } \mathbf{x} \in \Omega); \tag{2}$$

we also assume that g is *monotone* in the sense that

$$(g(\mathbf{x}, t_1) - g(\mathbf{x}, t_2))(t_1 - t_2) \geq 0 \quad \forall t_1, t_2 \in \mathbb{R} \quad \forall \mathbf{x} \in \Omega, \tag{3a}$$

Keywords and phrases. Semilinear elliptic boundary value problems, monotone operators, subcritical growth, corner-weighted Sobolev spaces, elliptic corner singularities, finite element methods, optimal convergence, graded meshes in polygons, Trudinger inequality.

Mathematics Institute, University of Bern, CH-3012 Bern, Switzerland.

*Corresponding author: thomas.wihler@unibe.ch

and that

$$g(\mathbf{x}, t)t \geq 0 \quad \forall t \in \mathbb{R} \quad \forall \mathbf{x} \in \Omega. \quad (3b)$$

Moreover, in order to specify the boundary data, we consider the subsets $\Gamma_{\mathcal{D}} = \bigcup_{j \in \mathcal{D}} \Gamma_j$ and $\Gamma_{\mathcal{N}} = \bigcup_{j \in \mathcal{N}} \Gamma_j$, which refer to Dirichlet and Neumann boundary conditions, respectively, where $\mathcal{D} \neq \emptyset$ and \mathcal{N} are two disjoint sets with $\mathcal{D} \cup \mathcal{N} = \{1, \dots, m\}$.

The focus of this paper is on *optimally convergent* iterative finite element approximations of the weak formulation of (1), which consists in finding a weak solution $u \in H_{\mathcal{D}}^1(\Omega)$ such that

$$a(u, v) + b(u; v) = \int_{\Omega} f v \, d\mathbf{x} \quad \forall v \in H_{\mathcal{D}}^1(\Omega), \quad (4)$$

where

$$a(u, v) := \int_{\Omega} \nabla u \cdot \nabla v \, d\mathbf{x}, \quad u, v \in H_{\mathcal{D}}^1(\Omega), \quad (5)$$

is the standard symmetric bilinear form for the Laplacian, and, for any given $w \in H_{\mathcal{D}}^1(\Omega)$, the form $b(w; \cdot)$ is given by

$$b(w; v) := \int_{\Omega} g(\mathbf{x}, w)v \, d\mathbf{x}, \quad v \in H_{\mathcal{D}}^1(\Omega). \quad (6)$$

Throughout, we denote by $H_{\mathcal{D}}^1(\Omega)$ the Sobolev space of all functions in $H^1(\Omega)$ with vanishing boundary trace along $\Gamma_{\mathcal{D}}$, equipped with the standard norm $\|\nabla(\cdot)\|_{L^2(\Omega)}$.

The *iterative* solution of monotone elliptic problems traces back to the early work of Zarantonello [26], where the convergence of a contractive fixed point iteration scheme for strongly monotone and (globally) Lipschitz continuous operators in Hilbert spaces has been established. In recent years, this approach has gained renewed interest for the purpose of numerical solution methods for monotone elliptic boundary value problems. More specifically, the key idea is to discretize the Zarantonello scheme (more generally referred to as Picard iteration) on suitably chosen discrete Galerkin spaces (*e.g.*, by applying finite elements), and thereby to deal with possible nonlinearities by naturally linking the iterative discretization to the underlying infinite-dimensional problem. This approach gives rise to the so-called *iterative linearized Galerkin (ILG)* methodology, which has been introduced as a general abstract framework in [17]; see also [11]. Furthermore, in combination with *adaptive* finite element discretizations for more specific nonlinear elliptic problems, the application of iterative linearization schemes has been developed, *e.g.*, in [2, 12, 13]. In addition, in the articles [14, 15, 18], the convergence and cost optimality of adaptive ILG discretizations for strongly monotone problems has been studied.

The fixed point technique applied in Zarantonello's original work [26] can be extended to monotone problems that are merely *locally* Lipschitz continuous. This observation has been pursued in the papers [7, 8, 16] for monotone semilinear elliptic problems with monomial reactions, or involving more general nonlinearities of *algebraic growth* in the works [6, 9]. Among the favorable properties of the Zarantonello iteration, which will be applied as a nonlinear solver in this paper, we point, on the one hand, to its ability to yield global convergence under mere monotonicity and subcritical growth assumptions, and, on the other hand, to the fact that the same symmetric stiffness matrix can be reused in each iteration step, thus avoiding expensive assemblies and extra regularity requirements from other nonlinear solvers; see [16] for details. Specifically, the novelty of the present paper is twofold:

- (1) Our analysis reaches far beyond the previously mentioned works, where algebraically growing nonlinearities have been considered, and enables nonlinear reaction terms of exponential and even of subcritical growth; cf. (2). Specifically, for a suitable parameter $\alpha > 0$, we will prove that the Picard iteration scheme (here written in strong form) given by

$$-\Delta u_{n+1} = -\Delta u_n + \alpha(\Delta u_n - g(\cdot, u_n) + f), \quad n \geq 0, \quad (7)$$

converges to the (unique) solution of (1) as $n \rightarrow \infty$; see Theorem 2.5. For this purpose, we derive a local Lipschitz bound (Lem. 2.3) whose proof crucially hinges on the application of the *Trudinger inequality* [24]. Our result immediately applies to discrete subspaces of $H^1_{\mathcal{D}}(\Omega)$, and thus, in particular, to any conforming finite element discretization of the iteration (7), thereby yielding a convergent ILG method for (1) that does not require to (directly) solve any nonlinear algebraic system; see Section 3.3.

- (2) In addition, we develop a regularity result for (1) in a Kondratiev-type setting [20], which takes into account possible corner singularities in the solution. More precisely, we make use of the regularity theory for linear elliptic problems developed in [3, 5] in order to show that the solution of (1) belongs to a family of corner-weighted Sobolev spaces (featuring classical H^2 -regularity in the interior of the domain Ω). This, in turn, motivates the application of so-called graded mesh refinements towards the corners of Ω , introduced in [5] (see also [22], Sect. 4.3), which allow to resolve the occurring singularities at an optimal rate. Our main result, Theorem 3.3, states that the proposed ILG scheme for (1) based on \mathbb{P}_1 -FE spaces with appropriate local mesh grading converges optimally and for finitely many iterations on each discrete space.

We remark that the subcritical growth assumption (SCG) used here is *not* needed for mere existence and uniqueness of solutions to monotone semilinear elliptic problems. In fact, classical approaches based on monotone operator theory or the Stampacchia truncation method (see, e.g., [19, 23]) yield existence (and uniqueness) under weaker hypotheses. Such more generally applicable ideas have been employed recently in the context of finite element discretizations of semilinear elliptic problems with non-Lipschitz nonlinearities; see [25]. The present work, however, deliberately adopts (SCG) because we aim for a *constructive* approach. Specifically, for the Picard-type iteration (7), we prove that the update map $u_n \mapsto u_{n+1}$ is contractive (see the proof of Thm. 2.5 for details), thereby giving rise to an explicit iterative solver with corresponding *a priori* stability bounds (cf. (22) and (26)); these estimates are essential for the convergence analysis of our ILG scheme and for proving that finitely many iterations on each discrete Galerkin space are sufficient to reach a prescribed accuracy.

Outline: We begin by recalling the classical Trudinger inequality in Section 2.1, and by deriving a suitable modification that serves our purposes (Prop. 2.1); in particular, this allows to show the convergence of the iteration (7) in Section 2.2, and thereby to provide a constructive proof for the existence of a (unique) solution of (1) that constitutes the basis of our ILG analysis. Furthermore, we derive a regularity result in weighted Sobolev spaces in Section 2.3. In Section 3 we turn our attention to finite element discretizations of both the nonlinear model problem (1) and the iterative Picard method (7) on graded meshes towards the corners of Ω in Sections 3.1, 3.2 and 3.3, respectively. Finally, we test our theoretical findings through a series of numerical experiments in Section 4, which highlight the practical performance of our ILG method, and draw some conclusions in Section 5.

2. WEAK SOLUTION AND REGULARITY

The existence of a weak solution of (1) under the subcritical growth condition (2) will be established based on the theory of monotone operators, see, e.g., [27]. In order to control the nonlinearity, we make use of the so-called Trudinger inequality.

2.1. Subcritical growth and the Trudinger inequality

For any function $\varphi : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ that is continuous in the second argument and has subcritical growth, i.e.,

$$\varphi(\cdot, \xi) = o(\exp(\sigma\xi^2)) \quad \text{for any } \sigma > 0 \text{ (and uniformly in } \mathbf{x} \in \Omega), \quad (8)$$

we recall the classical Trudinger inequality (in the two-dimensional case) states that there are constants $\mu_0 > 0$ and $C_9 > 0$ (depending on Ω) such that

$$\sup_{\substack{v \in H^1(\Omega) \\ \|v\|_{H^1(\Omega)} \leq 1}} \int_{\Omega} \exp(\mu_0 v^2) \, d\mathbf{x} \leq C_9; \quad (9)$$

see Trudinger's original paper ([24], Thm. 2). We note that the restriction on the H^1 -norm in the supremum above is usually not satisfied for weak solutions of (1), however, exploiting the subcritical growth property permits to circumvent this issue. To this end, for any φ as above, we notice that the expression

$$\vartheta(\varphi, \sigma) := \sup_{(\mathbf{x}, \xi) \in \Omega \times \mathbb{R}} \frac{|\varphi(\mathbf{x}, \xi)|}{\exp(\sigma \xi^2)} \quad (10)$$

is finite for any $\sigma > 0$. Then, we establish the following result, for which we first define

$$\mu(r) := \mu_0 \min\{1, r^{-2}\}, \quad r > 0, \quad (11)$$

with $\mu_0 > 0$ the constant from the Trudinger inequality (9).

Proposition 2.1 (Subcritical growth estimates). *Consider $u \in H^1(\Omega)$ and $\rho > 0$ such that $\|\nabla u\|_{L^2(\Omega)} \leq \rho$. If $0 \leq \sigma \leq \mu(\rho)$, then we have the bound*

$$\int_{\Omega} \exp(\sigma u^2) \, d\mathbf{x} \leq C_9. \quad (12)$$

Furthermore, for a function $\varphi : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ that has subcritical growth as in (8), and any $p \in [1, \infty)$, it holds that

$$\|\varphi(\cdot, u)\|_{L^p(\Omega)} \leq C_9^{1/p} \vartheta(\varphi, \mu(\rho)/p), \quad (13)$$

and, in particular, $\varphi(\cdot, u) \in L^p(\Omega)$.

Proof. Let $\rho > 0$ and $u \in H^1(\Omega)$ with $\|\nabla u\|_{L^2(\Omega)} \leq \rho$. If $\rho \leq 1$, then we have $\mu(\rho) = \mu_0$, and the estimate (12) follows immediately by monotonicity and from the Trudinger inequality (9):

$$\int_{\Omega} \exp(\sigma u^2) \, d\mathbf{x} \leq \int_{\Omega} \exp(\mu_0 u^2) \, d\mathbf{x} \leq C_9.$$

Otherwise, we introduce the auxiliary function $\tilde{u} := \rho^{-1}u$, and observe the bound $\|\nabla \tilde{u}\|_{L^2(\Omega)} \leq 1$. Thus, employing (9), it follows that

$$\int_{\Omega} \exp(\sigma u^2) \, d\mathbf{x} \leq \int_{\Omega} \exp(\mu_0 \rho^{-2} u^2) \, d\mathbf{x} = \int_{\Omega} \exp(\mu_0 \tilde{u}^2) \, d\mathbf{x} \leq C_9.$$

Moreover, for $p \in [1, \infty)$ and a continuous function φ with subcritical growth, applying (10), we conclude that

$$\int_{\Omega} |\varphi(\mathbf{x}, u)|^p \, d\mathbf{x} \leq \vartheta(\varphi, \mu(\rho)/p)^p \int_{\Omega} \exp(\mu(\rho) u^2) \, d\mathbf{x} \leq C_9 \vartheta(\varphi, \mu(\rho)/p)^p < \infty,$$

which yields (13). □

Remark 2.2. If the partial derivative g_u satisfies (2), then we immediately deduce that g and any of its anti-derivatives G (with $\partial_u G(\mathbf{x}, \xi) = g(\mathbf{x}, \xi)$), also fulfill (SCG); in addition, all three functions are L^p -integrable on $H_D^1(\Omega)$, for any $p \in [1, \infty)$. Indeed, by the fundamental theorem of calculus and the (SCG)-property of g_u , for any $\sigma \geq 0$, we infer that

$$\begin{aligned} |g(\mathbf{x}, \xi)| &\leq |g(\mathbf{x}, 0)| + \int_0^\xi |g_u(\mathbf{x}, s)| \, ds \leq |g(\mathbf{x}, 0)| + C \int_0^\xi e^{\sigma s^2} \, ds \\ &\leq |g(\mathbf{x}, 0)| + C|\xi| e^{\sigma \xi^2} \leq |g(\mathbf{x}, 0)| + C e^{2\sigma \xi^2}, \end{aligned}$$

for all $\mathbf{x} \in \Omega$ and $\xi \in \mathbb{R}$, which shows (2) for g , and analogously for G . The L^p -integrability follows from the previous Proposition 2.1.

We will now prove a few estimates for the form b from (6) that will be instrumental for the subsequent convergence analysis. To this end, for any $q \in [1, \infty)$, we recall the continuous Sobolev embedding $H_{\mathcal{D}}^1(\Omega) \hookrightarrow L^q(\Omega)$, which is expressed in terms of the bound

$$\|v\|_{L^q(\Omega)} \leq C_{14}(q) \|\nabla v\|_{L^2(\Omega)} \quad \forall v \in H_{\mathcal{D}}^1(\Omega), \quad (14)$$

for a constant $C_{14}(q) > 0$; see, e.g., Theorem 6.3 of [1].

Lemma 2.3. *Suppose that the nonlinearity g in (4) satisfies (2), and let $p, p^* \in (1, \infty)$ be given such that $1/p + 1/p^* = 1$. Then, for any $u, v, w \in H_{\mathcal{D}}^1(\Omega)$ and $\rho > 0$ such that*

$$\max\{\|\nabla u\|_{L^2(\Omega)}, \|\nabla w\|_{L^2(\Omega)}\} \leq \rho, \quad (15)$$

the estimates

$$|b(u; v)| \leq C_9^{1/p} C_{14}(p^*) \vartheta(g, \mu(\rho)/p) \|\nabla v\|_{L^2(\Omega)} \quad (16)$$

and

$$|b(u; v) - b(w; v)| \leq C_9^{1/p} C_{14}(2p^*)^2 \vartheta(g_u, \mu(\rho)/p) \|\nabla(u - w)\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} \quad (17)$$

hold true, with μ from (11).

Proof. Consider $u, v \in H_{\mathcal{D}}^1(\Omega)$, with $\|\nabla u\|_{L^2(\Omega)} \leq \rho$. Since g_u satisfies (2), we recall from Remark 2.2 that g is also of subcritical growth. Therefore, from (13), we note that

$$\|g(\cdot, u)\|_{L^p(\Omega)} \leq C_9^{1/p} \vartheta(g, \mu(\rho)/p).$$

Hence, applying Hölder's inequality and the Sobolev embedding (14), for each $v \in H_{\mathcal{D}}^1(\Omega)$, we establish (16):

$$|b(u; v)| \leq \|g(\cdot, u)\|_{L^p(\Omega)} \|v\|_{L^{p^*}(\Omega)} \leq C_9^{1/p} C_{14}(p^*) \vartheta(g, \mu(\rho)/p) \|\nabla v\|_{L^2(\Omega)}. \quad (18)$$

In order to derive (17), we use the fundamental theorem of calculus to infer that

$$\begin{aligned} |b(u; v) - b(w; v)| &= \left| \int_{\Omega} \int_0^1 \frac{d}{ds} g(\mathbf{x}, su + (1-s)w)v \, ds \, d\mathbf{x} \right| \\ &= \left| \int_{\Omega} \int_0^1 g_u(\mathbf{x}, su + (1-s)w) \, ds (u-w)v \, d\mathbf{x} \right|. \end{aligned}$$

Then, by employing a triple Hölder inequality with $q := 2p/(p-1) = 2p^* \in (2, \infty)$, so that $2/q + 1/p = 1$, we obtain

$$|b(u; v) - b(w; v)| \leq \|u - w\|_{L^q(\Omega)} \|v\|_{L^q(\Omega)} \left(\int_{\Omega} \left(\int_0^1 |g_u(\mathbf{x}, su + (1-s)w)| \, ds \right)^p \, d\mathbf{x} \right)^{1/p}.$$

Since the function $t \mapsto t^p$ is convex for $p > 1$, by Jensen's inequality, we deduce that

$$|b(u; v) - b(w; v)| \leq \|u - w\|_{L^q(\Omega)} \|v\|_{L^q(\Omega)} \left(\int_{\Omega} \int_0^1 |g_u(\mathbf{x}, su + (1-s)w)|^p \, ds \, d\mathbf{x} \right)^{1/p}. \quad (19)$$

Furthermore, in light of (2), letting $\sigma := \mu(\rho)/p$ and recalling (10), we have

$$|g_u(\mathbf{x}, \xi)| \leq \vartheta(g_u, \sigma) e^{\sigma \xi^2} \quad \forall (\mathbf{x}, \xi) \in \Omega \times \mathbb{R}.$$

This yields

$$\begin{aligned} \int_0^1 |g_u(\mathbf{x}, su + (1-s)w)|^p ds &\leq \vartheta(g_u, \sigma)^p \int_0^1 \exp(\sigma p(su + (1-s)w)^2) ds \\ &\leq \vartheta(g_u, \sigma)^p \exp(\sigma p \xi_{u,w}^2), \end{aligned}$$

where we define $\xi_{u,w} := \max_{s \in [0,1]}(su + (1-s)w)$. Moreover, invoking (15), which implies that $\|\nabla \xi_{u,w}\|_{L^2(\Omega)} \leq \rho$, we infer from Proposition 2.1 that

$$\int_{\Omega} \exp(\sigma p \xi_{u,w}^2) d\mathbf{x} = \int_{\Omega} \exp(\mu(\rho) \xi_{u,w}^2) d\mathbf{x} \leq C_9.$$

Hence, employing the Sobolev inequality (14) in (19), we arrive at (17). □

Remark 2.4. In the case of pure Dirichlet boundary conditions (*i.e.*, when $\mathcal{N} = \emptyset$ and $H_D^1(\Omega) = H_0^1(\Omega)$), we can refine Proposition 2.1 by invoking the so-called *Moser–Trudinger inequality* established in [21], which shows that $\mu_0 = 4\pi$ is optimal in (9) on $H_0^1(\Omega)$ (with the integral on the left-hand side of (9) arbitrary large for $\mu_0 > 4\pi$ and suitable functions $v \in H_0^1(\Omega)$). Consequently, Lemma 2.3 (and its implications) extends, in particular cases, to the situation of critical growth, *i.e.*, when there exists $\sigma_0 > 0$ such that

$$\lim_{|\xi| \rightarrow \infty} \frac{|g(\cdot, \xi)|}{\exp(\sigma|\xi|^2)} = \begin{cases} 0, & \text{if } \sigma > \sigma_0, \\ \infty, & \text{if } \sigma < \sigma_0, \end{cases} \quad \text{uniformly in } \mathbf{x} \in \Omega.$$

A sufficient condition for this extension is $\sigma_0 < \mu(\rho)/p \leq 4\pi/p$.

2.2. Existence and uniqueness of weak solution

The existence of a weak solution to (1) will be established in a constructive way using Banach’s contraction theorem. For this purpose, for a parameter $\alpha \in (0, 1]$ and an arbitrary closed linear subspace $\mathbb{W} \subset H_D^1(\Omega)$ (*e.g.*, $H_D^1(\Omega)$ itself, or a finite-dimensional subspace), we define an operator $T_\alpha : \mathbb{W} \rightarrow \mathbb{W}$ through the weak formulation

$$u \mapsto T_\alpha(u) : \quad a(T_\alpha(u), v) = (1 - \alpha)a(u, v) + \alpha \left(\int_{\Omega} f v d\mathbf{x} - b(u; v) \right) \quad \forall v \in \mathbb{W}, \tag{20}$$

where $a(\cdot, \cdot)$ and $b(\cdot; \cdot)$ are the forms from (5) and (6), respectively, cf. (7).

For an appropriate range of α , the ensuing result shows that T_α is a well-defined, self-mapping contraction on the closed ball

$$\mathcal{B}_\rho := \{v \in H_D^1(\Omega) : \|\nabla v\|_{L^2(\Omega)} \leq \rho\},$$

for $\rho > 0$ sufficiently large.

Theorem 2.5 (Existence and uniqueness). *Suppose that*

$$\rho > C_{14}(p^*) \|f\|_{L^p(\Omega)}. \tag{21}$$

Then, there is $\alpha_0 \in (0, 1]$ such that T_α , with $0 < \alpha \leq \alpha_0$, has a unique fixed-point in $\mathcal{B}_\rho \cap \mathbb{W}$, which is the (only) solution of (4). Furthermore, the stability bound

$$\|\nabla u\|_{L^2(\Omega)} \leq C_{14}(p^*) \|f\|_{L^p(\Omega)} \tag{22}$$

holds true.

Proof. We modify the proof of Theorem 25.B from [27], which is based on a global argument, to a local analysis on the ball $\mathcal{B}_\rho \cap \mathbb{W}$. We proceed in several steps.

- (1) For any given $u \in \mathcal{B}_\rho \cap \mathbb{W}$, with ρ as in (21), we begin by noticing that the right-hand side of the weak formulation (20) is a bounded linear functional on \mathbb{W} . Indeed, applying the Cauchy–Schwarz inequality, we see that $|a(u, v)| \leq \rho \|\nabla v\|_{L^2(\Omega)}$. Moreover, from Hölder’s inequality (with $1/p + 1/p^* = 1$) and the continuous Sobolev embedding (14), we obtain

$$\left| \int_{\Omega} f v \, d\mathbf{x} \right| \leq \|f\|_{L^p(\Omega)} \|v\|_{L^{p^*}(\Omega)} \leq C_{14}(p^*) \|f\|_{L^p(\Omega)} \|\nabla v\|_{L^2(\Omega)}. \quad (23)$$

Finally, the term $b(u; v)$ is bounded due to (16):

$$|b(u; v)| \leq C_9^{1/p} C_{14}(p^*) \vartheta(g, \mu(\rho)/p) \|\nabla v\|_{L^2(\Omega)}.$$

Hence, by the Riesz representation theorem, we conclude that $\mathsf{T}_\alpha(u) \in \mathbb{W}$ is well-defined.

- (2) We show that $\mathsf{T}_\alpha(u) \in \mathcal{B}_\rho \cap \mathbb{W}$ for any $u \in \mathcal{B}_\rho \cap \mathbb{W}$. For this purpose, we define $\zeta_u \in \mathbb{W}$ by

$$a(\zeta_u, v) = \int_{\Omega} f v \, d\mathbf{x} - b(u; v) \quad \forall v \in \mathbb{W}. \quad (24)$$

Testing with $v = \zeta_u$, and using the estimates derived above, we infer the stability bound

$$\|\nabla \zeta_u\|_{L^2(\Omega)} \leq L_\rho, \quad \text{with } L_\rho := C_{14}(p^*) \left(\|f\|_{L^p(\Omega)} + C_9^{1/p} \vartheta(g, \mu(\rho)/p) \right).$$

Furthermore, we observe the identity $\mathsf{T}_\alpha(u) = (1 - \alpha)u + \alpha\zeta_u$, from which we deduce that

$$\begin{aligned} \|\nabla \mathsf{T}_\alpha(u)\|_{L^2(\Omega)}^2 &= (1 - \alpha)^2 \|\nabla u\|_{L^2(\Omega)}^2 + 2\alpha(1 - \alpha)a(u, \zeta_u) + \alpha^2 \|\nabla \zeta_u\|_{L^2(\Omega)}^2 \\ &\leq (1 - \alpha)^2 \rho^2 + 2\alpha(1 - \alpha)a(u, \zeta_u) + \alpha^2 L_\rho^2. \end{aligned} \quad (25)$$

Employing (3b) and (23), we observe that

$$a(u, \zeta_u) = \int_{\Omega} f u \, d\mathbf{x} - b(u; u) \leq \int_{\Omega} f u \, d\mathbf{x} \leq C_{14}(p^*) \|f\|_{L^p(\Omega)} \|\nabla u\|_{L^2(\Omega)}.$$

Hence, we arrive at $\|\nabla \mathsf{T}_\alpha(u)\|_{L^2(\Omega)}^2 \leq \Psi(\alpha)$, where we let

$$\Psi(\alpha) := (1 - \alpha)^2 \rho^2 + 2\alpha(1 - \alpha) C_{14}(p^*) \|f\|_{L^p(\Omega)} \rho + \alpha^2 L_\rho^2.$$

Notice that $\Psi(0) = \rho^2$ and $\Psi'(0) = -2\rho(\rho - C_{14}(p^*) \|f\|_{L^p(\Omega)})$; the latter term is negative in view of (21). Consequently, for any $\alpha > 0$ small enough we have $0 \leq \Psi(\alpha) \leq \rho^2$, and it follows that $\|\nabla \mathsf{T}_\alpha(u)\|_{L^2(\Omega)} \leq \rho$.

- (3) Next, we establish a contraction property for T_α . For any $u, v \in \mathcal{B}_\rho \cap \mathbb{W}$, let us recall the corresponding functions $\zeta_u, \zeta_v \in \mathbb{W}$ defined in (24). Then, as in the previous step, we have the identity

$$\mathsf{T}_\alpha(u) - \mathsf{T}_\alpha(v) = (1 - \alpha)(u - v) + \alpha(\zeta_u - \zeta_v),$$

which leads to

$$\begin{aligned} &\|\nabla(\mathsf{T}_\alpha(u) - \mathsf{T}_\alpha(v))\|_{L^2(\Omega)}^2 \\ &= (1 - \alpha)^2 \|\nabla(u - v)\|_{L^2(\Omega)}^2 + 2\alpha(1 - \alpha)a(\zeta_u - \zeta_v, u - v) + \alpha^2 \|\nabla(\zeta_u - \zeta_v)\|_{L^2(\Omega)}^2, \end{aligned}$$

cf. (25). Using (3a), it holds that

$$a(\zeta_u - \zeta_v, u - v) = -b(u; u - v) + b(v; u - v) \leq 0;$$

furthermore, by observing that

$$\|\nabla(\zeta_u - \zeta_v)\|_{L^2(\Omega)}^2 = a(\zeta_u - \zeta_v, \zeta_u - \zeta_v) = -b(u; \zeta_u - \zeta_v) + b(v; \zeta_u - \zeta_v),$$

and upon applying (17), we obtain

$$\|\nabla(\zeta_u - \zeta_v)\|_{L^2(\Omega)} \leq C_9^{1/p} C_{14} (2p^*)^2 \vartheta(g_u, \mu(\rho)/p) \|\nabla(u - w)\|_{L^2(\Omega)}.$$

Thus,

$$\|\nabla(\mathsf{T}_\alpha(u) - \mathsf{T}_\alpha(v))\|_{L^2(\Omega)}^2 \leq \lambda(\alpha) \|\nabla(u - w)\|_{L^2(\Omega)}^2 \quad \forall u, v \in \mathcal{B}_\rho \cap \mathbb{W}, \tag{26a}$$

with

$$\lambda(\alpha) := (1 - \alpha)^2 + \alpha^2 C_9^{2/p} C_{14} (2p^*)^4 \vartheta(g_u, \mu(\rho)/p)^2. \tag{26b}$$

Similarly to the previous step, it can be verified that $\lambda(\alpha) < 1$ for all $\alpha > 0$ sufficiently small.

- (4) The existence and uniqueness of a fixed point $u \in \mathcal{B}_\rho \cap \mathbb{W}$ of T_α , and thus of a solution of (4), follows immediately from Banach’s contraction theorem. Since ρ can be arbitrary large, u is in fact the only fixed point in \mathbb{W} . Finally, the stability estimate (22) results from taking the limit $\rho \searrow C_{14}(p^*)\|f\|_{L^p(\Omega)}$.

This completes the proof. □

2.3. Regularity in weighted Sobolev spaces

In situations where the domain $\Omega \subset \mathbb{R}^2$ contains corners, as is the case for non-degenerate polygons, it is well-known that the inverse Laplacian $(-\Delta)^{-1}$ does not exhibit full elliptic regularity. Specifically, for right-hand side functions $f \in L^2(\Omega)$, solutions of Poisson-type boundary value problems are typically found to be H^2 away from the boundary $\partial\Omega$, however, H^2 -regularity does in general not extend uniformly to the corners. We will address this issue in terms of the corner-weighted Sobolev space $H_\beta^2(\Omega)$; see, e.g., [3, 5, 22]. To this end, we define the weight function

$$\Phi_\beta(\mathbf{x}) = \prod_{j=1}^m \text{dist}(\mathbf{x}, \mathbf{c}_j)^{\beta_j}, \tag{27}$$

which involves the distances from a point $\mathbf{x} \in \bar{\Omega}$ to any of the $m \geq 3$ corners $\mathbf{c}_1, \dots, \mathbf{c}_m$ of the polygon Ω , and associated exponents $\beta = (\beta_1, \dots, \beta_m)$, with $0 \leq \beta_1, \dots, \beta_m < 1$. We can then introduce the norm

$$\|v\|_{H_\beta^2(\Omega)}^2 := \|v\|_{H^1(\Omega)}^2 + \sum_{\alpha_1 + \alpha_2 = 2} \|\Phi_\beta |\partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} v|\|_{L^2(\Omega)}^2,$$

as well as the weighted Sobolev space

$$H_\beta^2(\Omega) := \left\{ v \in H^1(\Omega) : \|v\|_{H_\beta^2(\Omega)} < \infty \right\}.$$

We notice the continuous embedding $H_\beta^2(\Omega) \hookrightarrow C^0(\bar{\Omega})$; see, e.g., Section 2 of [5] for a proof.

Theorem 2.6 (Regularity of weak solution). *Suppose that the polygon Ω is non-degenerate, i.e., all interior angles at the corners $\mathbf{c}_1, \dots, \mathbf{c}_m$, which we signify by $\omega_1, \dots, \omega_m$, fulfill $\omega_j \in (0, 2\pi)$ for each $j = 1, \dots, m$. Furthermore, let the nonlinearity g satisfy the subcritical growth condition (2) as well as the monotonicity conditions (3), and $f \in L^2(\Omega)$. If the weight exponents β_1, \dots, β_m in (27) obey the bounds*

$$1 > \beta_j > \begin{cases} 1 - \min(1, \pi/\omega_j), & \text{if } \mathbf{c}_j \text{ is a Dirichlet–Dirichlet or Neumann–Neumann corner,} \\ 1 - \min(1, \pi/2\omega_j), & \text{if } \mathbf{c}_j \text{ is a Dirichlet–Neumann corner,} \end{cases} \tag{28}$$

then the weak solution of (1) belongs to $H_D^1(\Omega) \cap H_\beta^2(\Omega)$. Furthermore, the stability estimate

$$\|u\|_{H_\beta^2(\Omega)} \leq C(\|f\|_{L^2(\Omega)} + \vartheta(g, \mu(\rho)/2)) \quad (29)$$

is satisfied, where $C > 0$ is a constant independent of u , f , and g , and $\rho = C_{14}(2)\|f\|_{L^2(\Omega)}$.

Proof. Consider the (unique) solution $u \in H_D^1(\Omega)$ of the weak formulation (4) whose existence was established in Theorem 2.5. Then, the linear functional

$$\ell(v) = \int_{\Omega} f v \, d\mathbf{x} - b(u; v), \quad v \in H_D^1(\Omega),$$

with the form $v \mapsto b(u; v)$ from (6), is bounded in $L^2(\Omega)$; indeed, this follows immediately from (18) (with $p = p^* = 2$) and by means of (23). Then, using the regularity shift

$$(-\Delta)^{-1} : L^2(\Omega) \rightarrow H_\beta^2(\Omega),$$

see, e.g., Theorem 3.2 of [5] or the original work by Kondratiev ([20], Thm. 1.1), it follows from (1) that

$$u = (-\Delta)^{-1}(f - g(\cdot, u)) \in H_\beta^2(\Omega),$$

where the Laplacian is understood in the weak sense, i.e., in terms of the bilinear form a arising in the weak formulation (4). Moreover, it holds the regularity estimate

$$\|u\|_{H_\beta^2(\Omega)} \leq C\|f - g(\cdot, u)\|_{L^2(\Omega)}.$$

Using (13) with $p = p^* = 2$, we deduce that

$$\|u\|_{H_\beta^2(\Omega)} \leq C(\|f\|_{L^2(\Omega)} + \vartheta(g, \mu(\rho)/2)),$$

where we note that $\rho = C_{14}(2)\|f\|_{L^2(\Omega)} \geq \|\nabla u\|_{L^2(\Omega)}$, cf. (22). \square

Remark 2.7. We notice that the above regularity result and the convergence analysis in this paper can be generalized to inhomogeneous Neumann boundary conditions in (1), i.e., for $\partial_{\mathbf{n}} u = u_{\mathcal{N}}$ on $\Gamma_{\mathcal{N}}$, provided that the Neumann boundary data $u_{\mathcal{N}}$ belongs to an appropriate trace space; see [4] for details.

3. OPTIMAL CONVERGENCE OF FINITE ELEMENT APPROXIMATIONS

We propose a numerical approximation scheme for the solution of (1) that is based on a Picard iteration scheme (also referred to as Zarantonello iteration [26]) and a suitable finite element discretization thereof. This combination gives rise to the so-called iterative linearized Galerkin (ILG) methodology that has been introduced in [11, 17]. We begin by analyzing the discretization of the weak formulation (4) in Galerkin subspaces.

3.1. Quasi-optimal approximation in finite-dimensional Galerkin spaces

Let $\mathbb{W}_N \subset H_D^1(\Omega)$ denote any finite-dimensional (and thus closed) subspace, where the index N refers to its dimension, i.e., $N = \dim(\mathbb{W}_N)$. We begin by introducing the Galerkin discretization of the weak formulation (4) on \mathbb{W}_N : Find $U \in \mathbb{W}_N$ such that

$$a(U, v) + b(U; v) = \int_{\Omega} f v \, d\mathbf{x} \quad \forall v \in \mathbb{W}_N. \quad (30)$$

By virtue of Theorem 2.5, we conclude that U exists and is unique; furthermore, for $p, p^* \in (1, \infty)$ with $1/p + 1/p^* = 1$, we note that

$$\|\nabla U\|_{L^2(\Omega)} \leq C_{14}(p^*)\|f\|_{L^p(\Omega)}, \quad (31)$$

cf. (22).

Proposition 3.1 (Quasi-optimality). *Let $e_U = u - U$ signify the error between the exact solution $u \in H_{\mathcal{D}}^1(\Omega)$ of (4) and its Galerkin approximation $U \in \mathbb{W}_N$ from (30). Then, it holds the quasi-optimality bound*

$$\|\nabla e_U\|_{L^2(\Omega)} \leq C_{32} \inf_{w \in \mathbb{W}_N} \|\nabla(u - w)\|_{L^2(\Omega)}, \quad (32a)$$

with

$$C_{32} := 1 + C_9^{1/p} C_{14} (2p^*)^2 \vartheta(g_u, \mu^{(\rho)/p}), \quad (32b)$$

and $\rho = C_{14}(p^*) \|f\|_{L^p(\Omega)}$.

Proof. We first observe the Galerkin orthogonality

$$a(e_U, v) + b(u; v) - b(U; v) = 0 \quad \forall v \in \mathbb{W}_N.$$

For any $w \in \mathbb{W}_N$, observing that $v = w - U \in \mathbb{W}_N$, this yields

$$\|\nabla e_U\|_{L^2(\Omega)}^2 = a(e_U, u - w) + a(e_U, w - U) = a(e_U, u - w) - b(u; w - U) + b(U; w - U).$$

Moreover, making use of (3a), results in

$$\begin{aligned} \|\nabla e_U\|_{L^2(\Omega)}^2 &= a(e_U, u - w) - b(u; u - U) + b(U; u - U) - b(u; w - u) + b(U; w - u) \\ &\leq a(e_U, u - w) - b(u; w - u) + b(U; w - u). \end{aligned}$$

Applying the Cauchy–Schwarz inequality we have that

$$|a(e_U, u - w)| \leq \|\nabla e_U\|_{L^2(\Omega)} \|\nabla(u - w)\|_{L^2(\Omega)}. \quad (33)$$

Furthermore, recalling (22) and (31), we notice that

$$\max\{\|\nabla u\|_{L^2(\Omega)}, \|\nabla U\|_{L^2(\Omega)}\} \leq C_{14}(p^*) \|f\|_{L^p(\Omega)} = \rho.$$

Employing (17), we thus derive the local Lipschitz bound

$$|b(u; w - u) - b(U; w - u)| \leq C_9^{1/p} C_{14} (2p^*)^2 \vartheta(g_u, \mu^{(\rho)/p}) \|\nabla e_U\|_{L^2(\Omega)} \|\nabla(u - w)\|_{L^2(\Omega)}. \quad (34)$$

Since $w \in \mathbb{W}_N$ is arbitrary, from (33) and (34), we immediately obtain the estimate (32). \square

3.2. Finite element approximations on graded meshes

The approximation of functions in the weighted Sobolev space $H_{\beta}^2(\Omega)$ within a finite element setting mandates the use of suitably refined meshes that are able to properly resolve possible elliptic corner singularities. To this end, we recall the graded meshes introduced in [5].

Definition 3.2 (Finite spaces on graded meshes). Let $\beta = (\beta_1, \dots, \beta_m) \in [0, 1]^m$ be a weight vector associated with the corners $\mathbf{c}_1, \dots, \mathbf{c}_m$ of the polygon Ω , and Φ_{β} the corresponding weight function from (27). Then, a regular (conforming) and shape-regular triangulation $\mathcal{T}_h = \{T\}_{T \in \mathcal{T}_h}$ of mesh size $h = \max_{T \in \mathcal{T}_h} h_T$, where h_T denotes the diameter of any triangle $T \in \mathcal{T}_h$, is called a *graded mesh* if there exists a constant $\kappa \geq 1$ such that, for all $T \in \mathcal{T}_h$, it holds

$$\kappa^{-1} \sup_T \Phi_{\beta} \leq h_T/h \leq \kappa \inf_T \Phi_{\beta}, \quad \text{if } \Phi_{\beta} > 0 \text{ on } \overline{T},$$

and

$$\kappa^{-1} \leq \frac{h_T}{h \sup_T \Phi_\beta} \leq \kappa, \quad \text{if there is a corner } \mathbf{c}_i \text{ of } \Omega \text{ with } \mathbf{c}_i \in \bar{T}.$$

Furthermore, for a given triangulation \mathcal{T}_h , we define the associated \mathbb{P}_1 -finite element space by

$$\mathbb{S}_{\mathcal{D}}^1(\Omega, \mathcal{T}_h) := \{u \in H_{\mathcal{D}}^1(\Omega) : u|_T \in \mathbb{P}_1(T) \quad \forall T \in \mathcal{T}_h\}, \quad (35)$$

where $\mathbb{P}_1(T)$ denotes the set of all linear polynomials on a triangle $T \in \mathcal{T}_h$; this space consists of all continuous, element-wise linear functions on the graded mesh \mathcal{T}_h , with zero boundary values along $\Gamma_{\mathcal{D}}$.

The finite-dimensional spaces $\mathbb{S}_{\mathcal{D}}^1(\Omega, \mathcal{T}_h)$, which are based on graded mesh families, are able to approximate functions in $H_{\beta}^2(\Omega)$ at an optimal rate, *i.e.*, qualitatively comparable to the approximation of H^2 -functions on uniform meshes. More precisely, referring to Lemma 4.5 of [5], we have the interpolation bound

$$\|\nabla(u - I_h u)\|_{L^2(\Omega)} \leq Ch \|u\|_{H_{\beta}^2(\Omega)},$$

where $I_h : H_{\mathcal{D}}^1(\Omega) \rightarrow \mathbb{S}_{\mathcal{D}}^1(\Omega, \mathcal{T}_h)$ is the standard nodal interpolant on \mathcal{T}_h , and $C > 0$ depends on Ω , and on the triangulation parameters β and κ only. Moreover, it can be seen that

$$N := \dim \mathbb{S}_{\mathcal{D}}^1(\Omega, \mathcal{T}_h) \lesssim h^{-2},$$

see Lemma 4.1 of [5], which means that the corner refinements applied in graded meshes are sufficiently local as to preserve the order of the number of degrees of freedom occurring in uniform meshes. In particular, recalling the quasi-optimality bound (32) (with $p = p^* = 2$) together with the regularity estimate (29), we obtain the following convergence result.

Theorem 3.3 (Optimal convergence of FEM). *Suppose that the nonlinearity g from (1) satisfies the subcritical growth condition (2) as well as the monotonicity conditions (3), and $f \in L^2(\Omega)$. Then, on the finite element space $\mathbb{W}_N := \mathbb{S}_{\mathcal{D}}^1(\Omega, \mathcal{T}_h)$ from (35), where \mathcal{T}_h is a graded mesh of mesh size $h > 0$ (cf. Def. 3.2 above), the Galerkin approximation $U \in \mathbb{W}_N$ from (30) satisfies the error bound*

$$\|\nabla(u - U)\|_{L^2(\Omega)} \leq CN^{-1/2} (1 + \vartheta(g_u, \mu(\rho)/2)) (\|f\|_{L^2(\Omega)} + \vartheta(g, \mu(\rho)/2)), \quad (36)$$

for a constant $C > 0$ depending on Ω , and on the triangulation parameters β and κ , but independent of $N := \dim \mathbb{S}_{\mathcal{D}}^1(\Omega, \mathcal{T}_h)$, and $\rho = C_{14}(2) \|f\|_{L^2(\Omega)}$.

3.3. Iterative linearized finite element solution

The operator \mathbb{T}_α from (20) motivates a fixed point iteration

$$U_{n+1} = \mathbb{T}_{N,\alpha}(U_n), \quad n \geq 0, \quad (37)$$

where $\mathbb{T}_{N,\alpha}$ is the Galerkin discretization of the operator \mathbb{T}_α in a suitable finite dimensional linear subspace \mathbb{W}_N , which, for $u \in H_{\mathcal{D}}^1(\Omega)$, is given by

$$a(\mathbb{T}_\alpha(u) - \mathbb{T}_{N,\alpha}(u), v) = 0 \quad \forall v \in \mathbb{W}_N.$$

We can thereby approximate the nonlinear discrete formulation (30) by an iterative process that was previously proposed in equation (12) of [17] for strongly monotone and uniformly Lipschitz continuous problems: Specifically, for an (arbitrary) initial guess $U_0 \in \mathbb{W}_N$, the iteration (37) generates a sequence $\{U_n\}_{n \geq 0} \subset \mathbb{W}_N$ through the following scheme:

$$U_{n+1} \in \mathbb{W}_N : \quad a(U_{n+1}, v) = (1 - \alpha)a(U_n, v) + \alpha \left(\int_{\Omega} f v \, dx - b(U_n; v) \right) \quad \forall v \in \mathbb{W}_N, \quad (38)$$

for all $n \geq 0$ and all $v \in H^1_{\mathcal{D}}(\Omega)$, cf. (20). Here, α is a fixed parameter satisfying $0 < \alpha \leq 1$.

For the case where \mathbb{W}_N are chosen to be finite element spaces based on graded meshes, see Section 3.2, we aim to prove that the resulting iterative approximations converge to the exact solution at an optimal rate. To this end, we first observe the following result.

Proposition 3.4. *Let $\mathbb{W}_N \subset H^1_{\mathcal{D}}(\Omega)$ be a closed linear subspace. Then, for $\alpha \in (0, 1]$ sufficiently small and for any initial guess $U_0 \in \mathbb{W}_N$, the iteration (38) converges to the (unique) solution $U \in \mathbb{W}_N$ of (30) as $n \rightarrow \infty$. Furthermore, there is a constant $0 \leq r_\alpha < 1$ such that the a priori error bound*

$$\|\nabla(U - U_n)\|_{L^2(\Omega)} \leq r_\alpha^n \|\nabla(U - U_0)\|_{L^2(\Omega)} \quad \forall n \geq 0, \tag{39}$$

holds true.

Proof. Letting $\mathbb{W} = \mathbb{W}_N$ in Theorem 2.5 and $\rho > \max\{\|\nabla U_0\|_{L^2(\Omega)}, C_{14}(p^*)\|f\|_{L^p(\Omega)}\}$, cf. (21), the operator $T_{N,\alpha}$ above is a self-mapping contraction on the ball $\mathbb{W}_N \cap \mathcal{B}_\rho$ (for $\alpha > 0$ sufficiently small), and the fixed-point iteration (37) converges to some $U \in \mathbb{W}_N \cap \mathcal{B}_\rho$, which, in turn, is the unique solution of (30). The estimate (39) is the classical a priori bound for contractive fixed point iterations (see, e.g., [10], Thm. 3.7-1) with the contraction constant $r_\alpha = \lambda(\alpha)^{1/2}$ from (26). \square

We are now ready to present the main result of this work which shows that the finite element approximations generated by $n = \mathcal{O}(\log(\dim \mathbb{S}^1_{\mathcal{D}}(\Omega, \mathcal{T}_h)))$ iterations (on graded meshes \mathcal{T}_h), yields optimally converging discrete solutions of (4).

Theorem 3.5 (Optimal convergence of iterative linearized FEM). *Suppose that the assumptions of Theorem 3.3 hold, and let $\alpha > 0$ be a sufficiently small parameter. Then, there is a constant $\gamma > 0$ (depending on the contraction constant r_α from (39)) such that, for any initial guess $U_0 \in \mathbb{S}^1_{\mathcal{D}}(\Omega, \mathcal{T}_h)$, where $\{\mathcal{T}_h\}_{h>0}$ is the graded mesh family from Section 3.2, performing $n \geq \gamma \log N$ iterations of (38) on $\mathbb{S}^1_{\mathcal{D}}(\Omega, \mathcal{T}_h)$, it holds the a priori error bound*

$$\|\nabla(u - U_n)\|_{L^2(\Omega)} \leq CN^{-1/2},$$

where $u \in H^1_{\mathcal{D}}(\Omega)$ is the exact solution of (4), and $N = \dim \mathbb{S}^1_{\mathcal{D}}(\Omega, \mathcal{T}_h)$; the constant $C > 0$ depends on Ω , and on the triangulation parameters β and κ , but is independent of N .

Proof. Applying the triangle inequality, and using the bounds (36) and (39), implies

$$\|\nabla(u - U_n)\|_{L^2(\Omega)} \leq \|\nabla(u - U)\|_{L^2(\Omega)} + \|\nabla(U - U_n)\|_{L^2(\Omega)} \leq C(N^{-1/2} + r_\alpha^n).$$

Then, for $n \geq \lceil \log N / 2 \rceil / |\log(r_\alpha)|$, we obtain $r_\alpha^n \leq N^{-1/2}$, which completes the argument. \square

4. NUMERICAL EXPERIMENTS

In this section, we conduct several numerical experiments to showcase the optimal convergence rate (OCR) of the proposed iterative linearized Galerkin approach for semilinear elliptic boundary value problems (1) with (SCG)-nonlinearities, cf. (2). In particular, these experiments aim to validate our theoretical results and to illustrate the performance of graded meshes to resolve possible corner singularities in polygons. We employ the Picard iteration (38) with the initial guess $U_0 \equiv 0$ being the zero function in $H^1_{\mathcal{D}}(\Omega)$ and with a sufficiently small parameter $\alpha \in (0, 1]$, cf. Proposition 3.4. Throughout, we consider the L-shaped domain

$$\Omega = (-1, 1)^2 \setminus ([-1, 0] \times [0, 1]), \tag{40}$$

with a re-entrant corner at the origin. We use a uniform initial mesh consisting of 12 triangles and 3 interior mesh nodes, and then apply successive red-green-blue (RGB) refinements on all (or a suitable subset of all)

triangles to generate subdivisions that are uniform (or graded towards the re-entrant corner, respectively). In order to compute an approximate finite element solution U_n in line with Theorem 3.5, for a suitable value $\gamma(\alpha) \in \mathbb{N}$, we impose a maximal number of iterations for (38) given by

$$\tilde{n} = \gamma \lceil \log(N) \rceil. \quad (41)$$

Any integration on triangles is performed using the standard 3-point quadrature rule at the midpoints of the edges.

4.1. Experiment 1: Exponential nonlinearity and uniform meshes

In this first example, we aim to test the attainability of the OCR for the exponential nonlinearity $g(u) = e^u$. Specifically, we solve the Dirichlet problem on the L-shaped domain Ω from (40):

$$\begin{aligned} -\Delta u + e^u &= f && \text{in } \Omega \\ u &= 0 && \text{on } \partial\Omega. \end{aligned}$$

Here, we choose the right-hand side f so that the exact solution is given by the smooth function $u(x, y) = \sin(\pi x) \sin(\pi y)$. We use a sequence of uniform meshes corresponding to $\beta = 0$ in Definition 3.2; in our plots, the associated finite element spaces are characterized by their numbers of dofs, denoted by N . Our focus in this experiment is on the iterative numerical scheme applied to the above problem.

Figure 1 illustrates the global error between the exact solution u and the iterative solution U_n against the number of degrees of freedom (dofs) N for each uniform mesh. The iterative scheme (38) is applied to each mesh until the (logarithmic) error slope between two consecutive meshes is less than -0.49 , or the maximal number of iterations according to (41) is reached; we choose $\gamma = 4$ to obtain the desired convergence rate for all the selected values of α , although our numerical results indicate that $\gamma = 1$ would indeed suffice for larger values of α . In the legend of Figure 1, for each choice of α , we display the total number of iterations over all meshes by the variable ‘‘It’’. The convergence plot clearly shows that the expected convergence rate of $N^{-1/2}$ is achieved for the iterative linearized Galerkin scheme. Not surprisingly, we observe that larger α -values lead to a more efficient performance of the Picard iteration.

The sensitivity with respect to the choice of α prompts the question of optimal parameter selection. To determine the best value of α for the given example, *i.e.*, to minimize the number of iterations to reach an error of $\|\nabla e\|_{L^2(\Omega)} \leq 2 \cdot 10^{-2}$, we employ a golden section method on the parameter interval $[0, 1]$. The plot in Figure 2 illustrate this optimization process yielding $\alpha_{\text{opt}} \approx 0.8924$ (with 2 iterations); here, the optimization is designed towards a minimal number of iterations needed to achieve the prescribed error tolerance on the last mesh (corresponding to approximately $3.9 \cdot 10^5$ degrees of freedom).

4.2. Experiment 2: Cubic nonlinearity on graded meshes

In our second experiment, we aim to demonstrate the necessity of utilizing graded meshes for the accurate approximation of a solution with a singularity at the origin in the L-shaped domain Ω from (40). We consider the Dirichlet problem

$$\begin{aligned} -\Delta u + u^3 &= f && \text{in } \Omega \\ u &= 0 && \text{on } \partial\Omega, \end{aligned}$$

where $g(u) = u^3$ is a cubic monomial nonlinearity. We choose f such that the exact solution is given by

$$u(x, y) = 2r^{-4/3}xy(1-x^2)(1-y^2),$$

with the radial variable $r = (x^2 + y^2)^{1/2}$, which has a singularity at the origin $\mathbf{c}_1 = (0, 0)$; indeed, it can be verified that $u \in H_{\beta}^2(\Omega)$, for the weight function $\Phi_{\beta} = r^{\beta}$, with $\beta > 1/3$, which is in line with the pure Dirichlet case

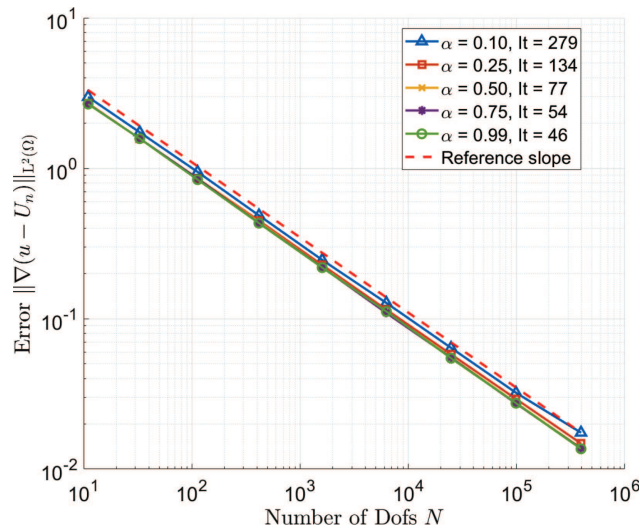


FIGURE 1. Example 1: Convergence behavior on uniform meshes for various values of α , and using $\gamma = 4$.

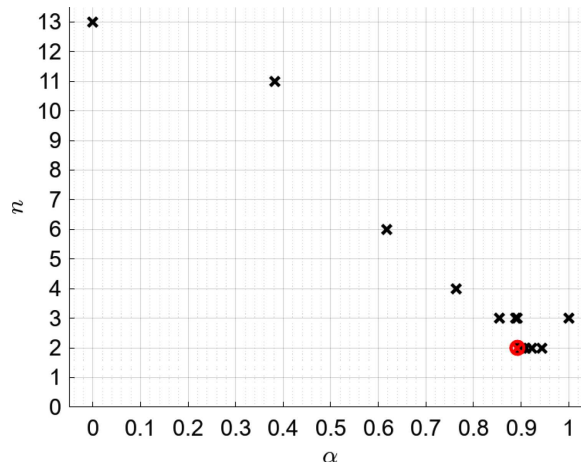


FIGURE 2. Experiment 1: Golden section optimization for the parameter α in the Picard iteration (with $\gamma = 1$). Identification of α_{opt} (with the optimal value circled).

in (28). The re-entrant corner singularity necessitates the use of graded mesh refinements towards $(0, 0)$, where we use $\beta = 0.4$. The corresponding graded meshes are constructed based on appropriate RGB refinements of triangles close to \mathbf{c}_1 , with values $h \in \{0.25, 0.15, 0.08, 0.035, 0.016, 0.008, 0.0038, 0.0019\}$; we present two examples in Figure 3.

Figure 4 compares the convergence rates on graded (left) and on uniform meshes (right). Whilst the OCR is reached on graded meshes, we clearly see a suboptimal convergence regime for uniform meshes (utilizing the maximal number of iterations at hand for each α). This is because uniform meshes do not effectively distribute the available degrees of freedom in the presence of the occurring corner singularity.

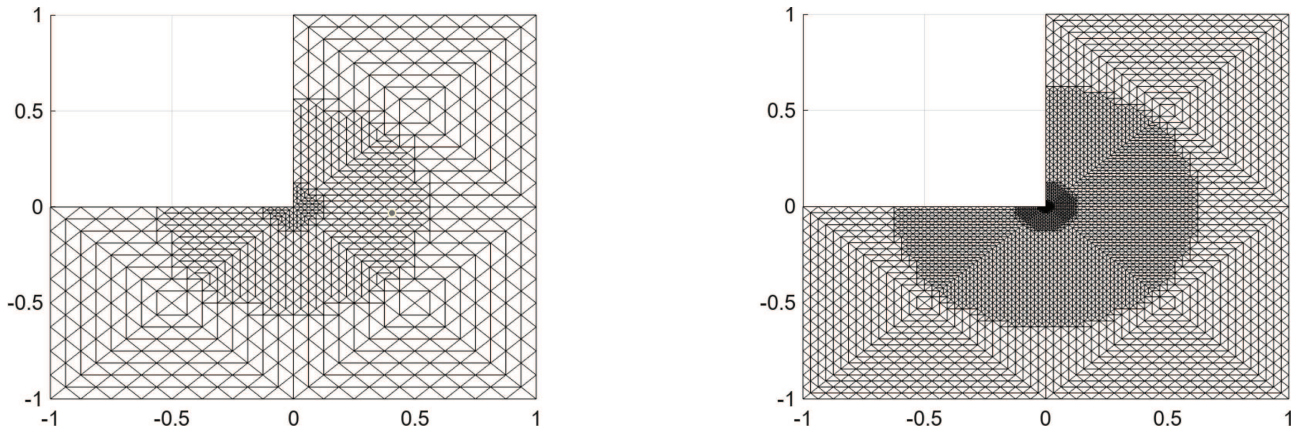


FIGURE 3. Graded meshes towards the origin in an L-shaped domain with $\beta = 0.40$ for mesh size $h = 0.035$ (left) and $h = 0.016$ (right).

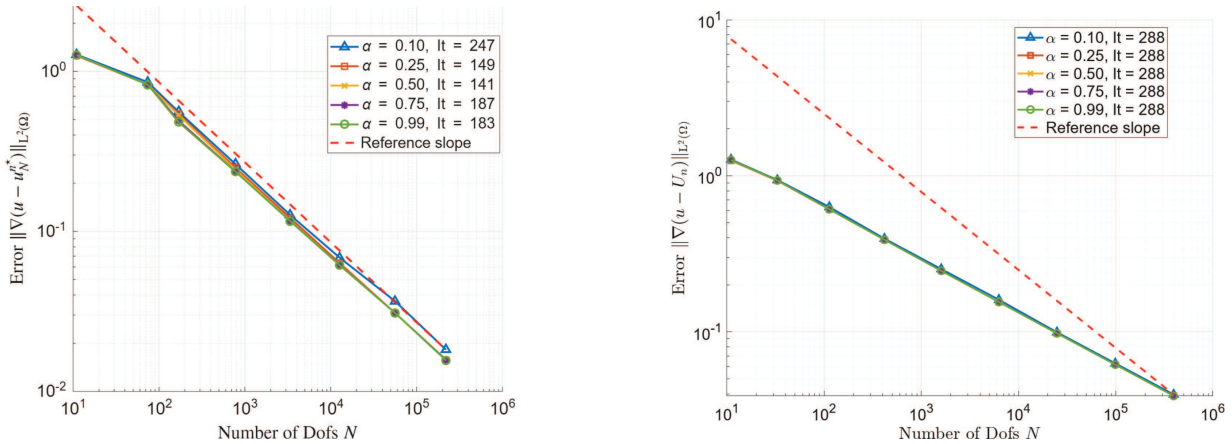


FIGURE 4. Experiment 2: Computational performance for several values of α and their corresponding total number of iterations used to reach the OCR with $\gamma = 4$ on graded meshes (left), and uniform meshes (right); the dashed line represents a reference slope of $-1/2$.

Similarly to the first experiment, we perform an optimization analysis in the parameter α . Figure 5 depicts the golden section method for $\gamma = 1$. More precisely, we illustrate the optimization process to reach the prescribed error of $\|\nabla e\|_{L^2(\Omega)} \leq 2 \cdot 10^{-2}$ on a graded mesh with roughly $3.9 \cdot 10^5$ degrees of freedom. We require 2 iterations with a corresponding optimal value of $\alpha_{\text{opt}} \approx 0.9152$, thereby underlining the effectiveness of the graded mesh approach in handling singularities at corners.

4.3. Experiment 3: Exponential nonlinearity and mixed boundary data

In this final example, we aim to test the effectiveness of graded meshes on a model problem with mixed boundary conditions and a (near) limit case of exponential nonlinearity in the sense of Definition 8. Specifically, on the L-shaped domain Ω from (40), we consider the boundary value problem

$$-\Delta u + e^{4|u|^{0.9}} u = f \quad \text{in } \Omega$$

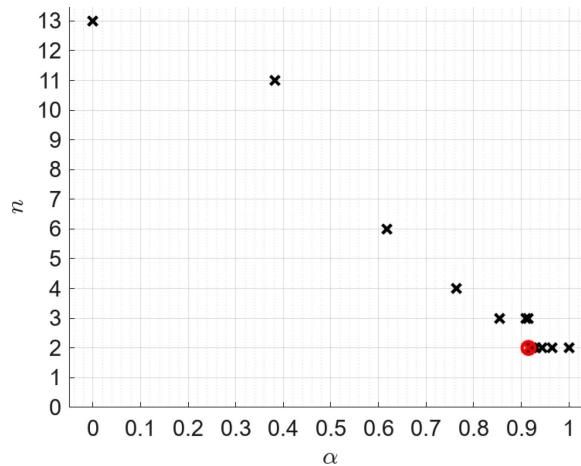


FIGURE 5. Experiment 2: Golden section optimization for the parameter α in the Picard iteration (with $\gamma = 1$). Identification of α_{opt} (with the optimal value circled).

$$\begin{aligned} u &= 0 && \text{on } \Gamma_{\mathcal{D}} \\ \partial_{\mathbf{n}} u &= 0 && \text{on } \Gamma_{\mathcal{N}}, \end{aligned}$$

where $\Gamma_{\mathcal{N}} = \{0\} \times (0, 1)$ and $\Gamma_{\mathcal{D}} = \partial\Omega \setminus \bar{\Gamma}_{\mathcal{N}}$. We choose f such that the exact solution is given by

$$u(x, y) = r^{-2/3} y(1 - x^2)(1 - y^2),$$

with a singularity at the corner $\mathbf{c}_1 = (0, 0)$; it holds $u \in H_{\beta}^2(\Omega)$ for the weight function $\Phi_{\beta} = r^{\beta}$, with $\beta > 2/3$, which corresponds to the mixed Dirichlet–Neumann case in (28). Accordingly, we propose to use graded meshes based on $\beta = 0.7$, for reasonable iteration parameters $\alpha = 0.5$ and $\gamma = 2$, and with the h -values chosen exactly as in Experiment 2. From Figure 6 we observe that the desired optimal rate is attained quickly, thereby demonstrating the benefit of the adjusted parameter settings in the current situation of mixed boundary conditions.

Furthermore, as in the previous experiments, Figure 7 depicts a computational analysis of the Picard scheme for various values of the iteration parameter α to reach the prescribed error of $\|\nabla e\|_{L^2(\Omega)} \leq 2 \cdot 10^{-2}$ on a mesh of approximately $4.3 \cdot 10^5$ degrees of freedom; for $\alpha_{\text{opt}} \approx 0.7416$, which results in 3 iterations, we perceive the high efficiency of the graded meshes to handle singularities, even with mixed boundary conditions. However, in contrast to Figures 2 and 5, we observe in Figure 7 a more unstable behavior in the required number of iterations, as well as a potential divergence of the scheme for larger values of α . In fact, this is made even clearer by Figure 8, where we increase the scaling parameter for the number of iterations \tilde{n} , cf. (41), from $\gamma = 1$ to $\gamma = 10$ to reach the OCR on the entire mesh sequence. We recognize three different stages in the behavior with respect to α : For intermediate values up to approximately $\alpha = 0.79$, a stable, oscillatory behavior is observed; beyond this range, we enter a small transitional phase where the number of iterations increases considerably; finally, from around $\alpha = 0.91$ onward, the Picard scheme does no longer converge within a sensible number of iterations.

We further investigate the failure of the Picard iteration to converge for values $\alpha \approx 1$. To this end, we employ two preset maximal iteration limits resulting from $\gamma = 1$ (i.e., $\sum \tilde{n} = 69$) and $\gamma = 10$ (i.e., $\sum \tilde{n} = 690$). Figure 9 shows three distinct behaviors for both $\gamma = 1$ and $\gamma = 10$: For $\gamma = 1$, we see that convergence begins to deteriorate from $\alpha = 0.85$ for larger numbers of degrees of freedom, yet, remains stable for $\gamma = 10$. For $\alpha = 0.95$, however, neither value of γ is able to retain convergence. In conclusion, for higher α -values (uniformly away

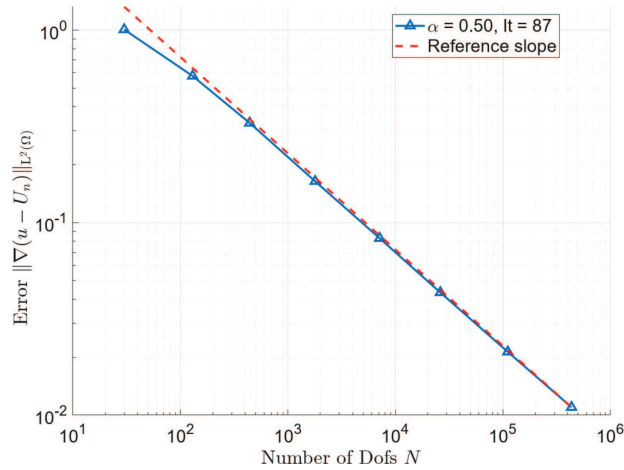


FIGURE 6. Experiment 3: Convergence behavior for a sequence of graded meshes with $\alpha = 0.5$, $\beta = 0.7$ and $\gamma = 2$.

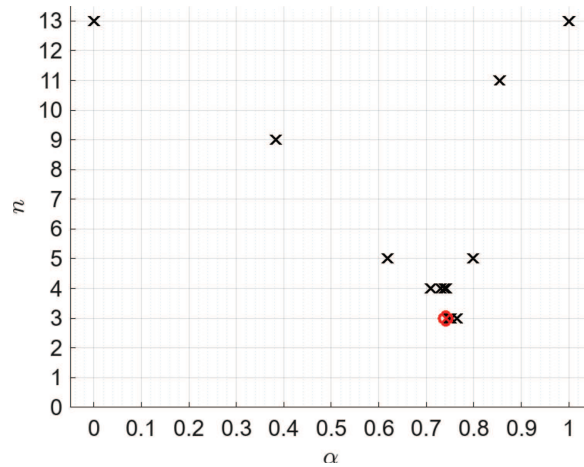


FIGURE 7. Experiment 3: Golden section optimization for the parameter α in the Picard iteration (with $\gamma = 1$). Identification of α_{opt} (with the optimal value circled).

from a certain upper limit < 1), OCR can still be retrieved by appropriately increasing the number of iterations. For α -values close to 1, however, the iteration seems to diverge completely, thereby indicating that these values may possibly exceed the threshold presented in Proposition 3.4.

In summary, our numerical study across the three experiments suggests that $\alpha \simeq 0.8$ is generally a good choice. This observation is reinforced by Figure 8, which shows that $\alpha \simeq 0.8$ remains effective given a sufficient number of iterations is being performed. This supports the application of the Picard iteration as an iterative nonlinear solver for the problem class under consideration. In particular, this method is stable and robust for a wide range of α values, requiring only a few iterations to reach optimal convergence rates, even for strong exponential nonlinearities.

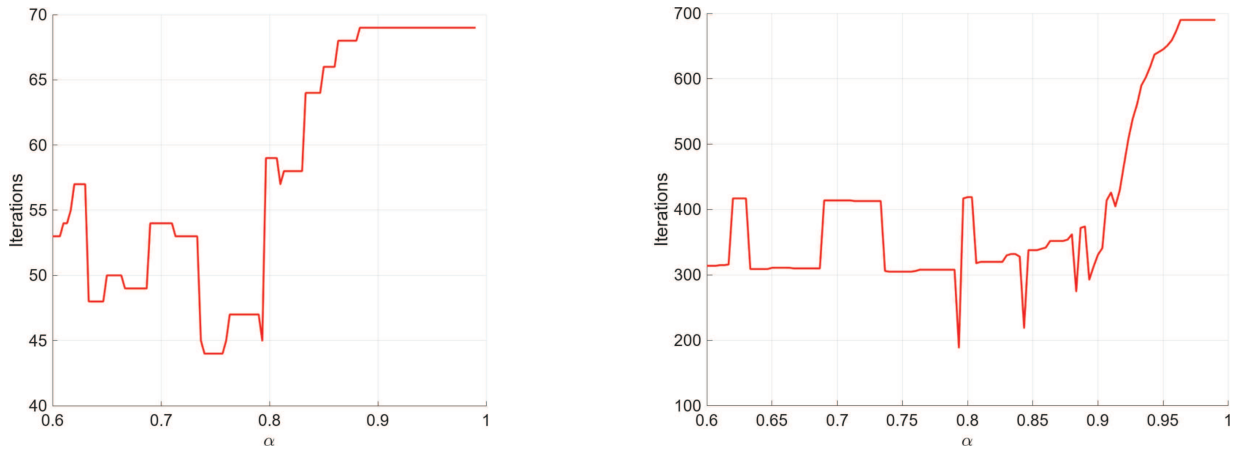


FIGURE 8. Experiment 3: Numbers of iterations for $\gamma = 1$ (left) and $\gamma = 10$ (right) needed across all meshes in terms of $\alpha \in [0.6, 0.99]$ for the nonlinearity $g(u) = \exp(4|u|^{0.9}u)$ and mixed boundary conditions. Distance between two α -samples: $\Delta\alpha = 1/30$.

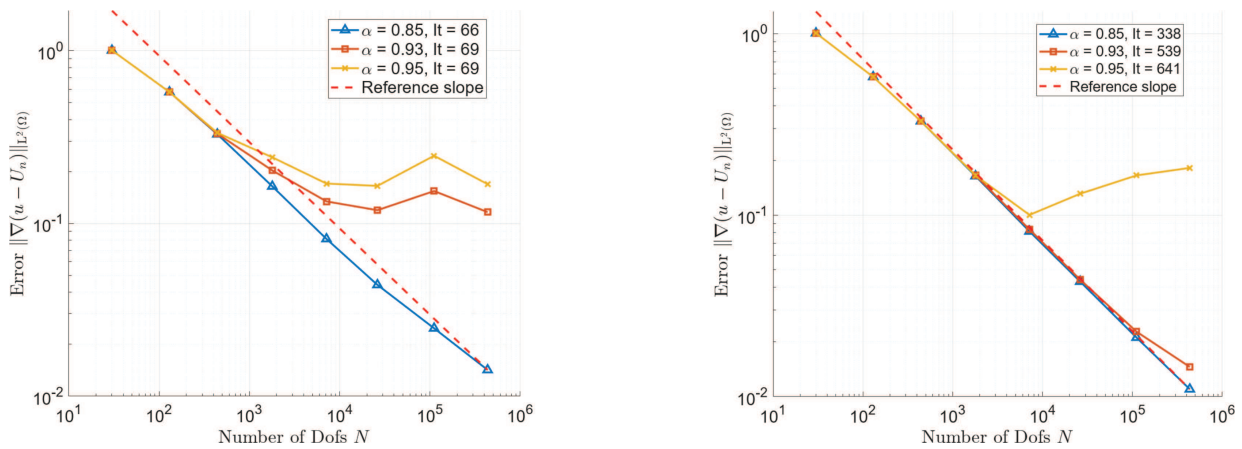


FIGURE 9. Experiment 3: Stages of divergence for various α values with $\gamma = 1$ (left) and $\gamma = 10$ (right).

4.4. Additional discussion: Scheme stability

Finally, in Figures 10 and 11, we explore how the strength of the nonlinearity $g(u)$ and the type of boundary conditions influence the convergence behavior of the Picard iteration. We fix $\gamma = 10$ in (41), and consider the range $\alpha \in [0.6, 0.99]$ for uniformly spaced samples of distance $\Delta\alpha = 1/30$ as before.

In Figure 10 we compare throughout the mesh sequence homogeneous Dirichlet and mixed boundary data for nearly critical nonlinearities. We see that the type of boundary conditions significantly impacts the convergence regime of the Picard iteration. In the former case (left plot), in spite of a larger constant in the exponent of the nonlinearity, a parabolic shape that resembles the plots in Figures 2, 5 and 7 is obtained. However, compared to the previous experiments, the iteration numbers are considerably larger. This is even more pronounced in the right plot, where mixed boundary conditions are investigated; in addition, step-like plateaus appear, where the

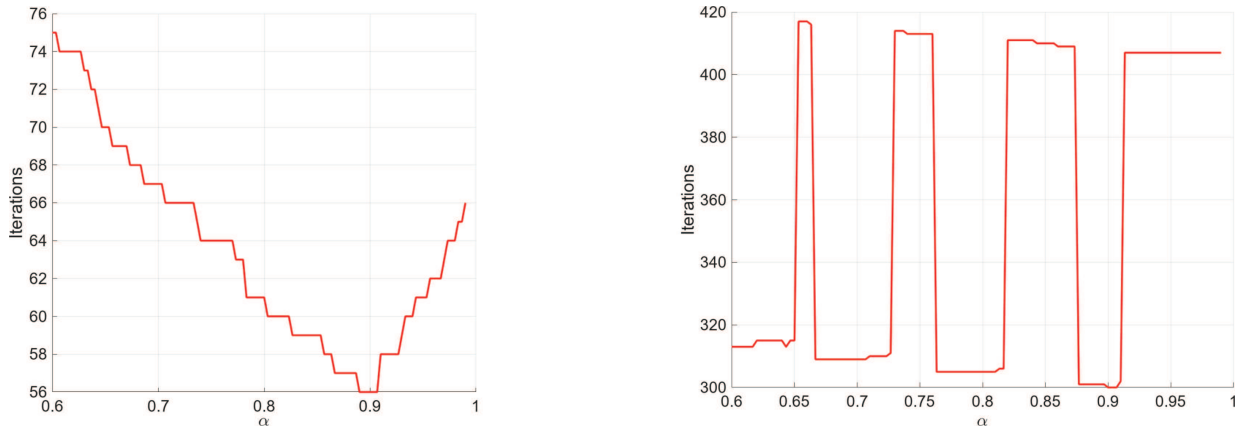


FIGURE 10. Numbers of iterations needed across all meshes in terms of α for $\gamma = 10$. Left: nonlinearity $g(u) = \exp(4|u|^{0.9}u)$ and *Dirichlet boundary conditions*. Right: nonlinearity $g(u) = \exp(|u|^{0.9}u)$ and *mixed boundary conditions*.

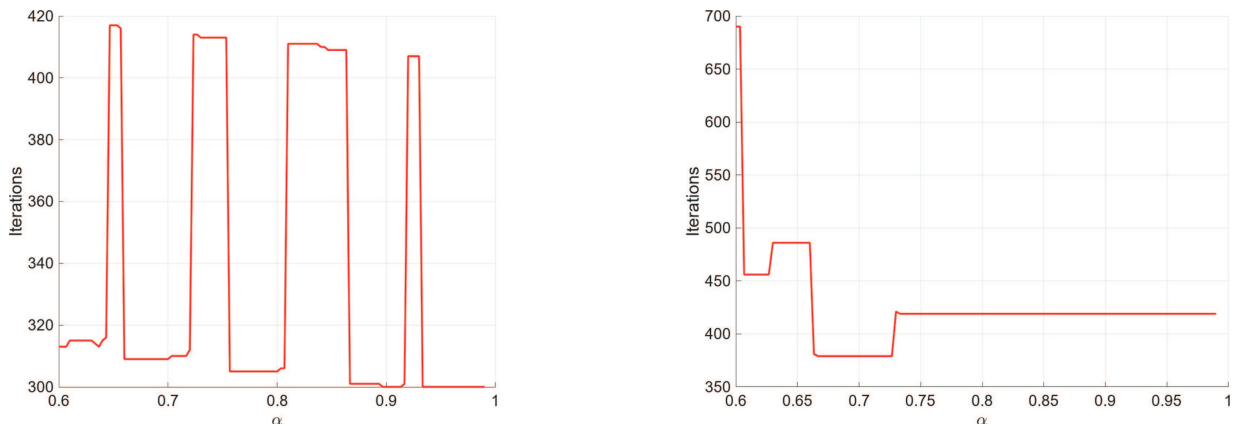


FIGURE 11. Numbers of iterations needed across all meshes in terms of α for $\gamma = 10$, with mixed boundary conditions. Left: nonlinearity $g(u) = \exp(u)$. Right: nonlinearity $g(u) = \exp(4u)$.

number of iterations are locally constant for increasing values of α . Evidently, this oscillatory behavior is quite unstable, as slight variations in α can cause a notable increase or decrease in the iteration numbers.

Finally, Figure 11 focuses solely on mixed boundary conditions. A significant change in shape and behavior is observed for two different exponential nonlinearities. Another remarkable observation lies in the even more drastic increase in the number of iterations, similar to Figure 10 (right), that is visible in both plots in Figure 11.

In conclusion, on a fairly speculative note, our experiments may indicate that the use of mixed boundary conditions plays a more important role with regard to the iteration numbers needed to reach the OCR by means of the Picard scheme than the strength of the (subcritical) nonlinearity.

5. CONCLUSIONS

In our work, we have developed the *a priori* error analysis of the iterative linearized Galerkin (ILG) framework for semilinear elliptic problems with mixed boundary conditions and exponentially or even subcritically grow-

ing nonlinear reactions. For such problems, we have demonstrated well-posedness under suitable monotonicity assumptions. This is accomplished through a new local analysis for a contractive Picard iteration, which, on the one hand, implies existence and uniqueness of a solution (on both the continuous and discrete level), and, on the other hand, motivates an iterative finite element scheme. A key ingredient is our modification of the Trudinger inequality that allows to remove the usual norm restriction, thereby enabling a search for weak solutions within the entire space $H^1(\Omega)$; in the pure Dirichlet case, based on the Moser–Trudinger inequality, our analysis extends to certain nonlinearities of critical growth. In addition, we have derived a quasi-optimality bound that, in a subsequent step, has led to optimal convergence for the iterative numerical approximation scheme on graded meshes. Furthermore, our numerical experiments validate the optimal convergence rate and thereby underscore the effectivity of the proposed approach in handling both challenging reaction nonlinearities and possible corner singularities.

FUNDING

The authors acknowledge the financial support of the Swiss National Science Foundation (SNSF), Grant No. 200021_212868.

DATA AVAILABILITY STATEMENT

The research data associated with this article are included in the article.

REFERENCES

- [1] R.A. Adams and J.J.F. Fournier, Sobolev Spaces. *Pure and Applied Mathematics*. Academic Press (2003).
- [2] M. Amrein and T.P. Wihler, An adaptive space-time Newton-Galerkin approach for semilinear singularly perturbed parabolic evolution equations. *IMA J. Numer. Anal.* **37** (2017) 2004–2019.
- [3] I. Babuška and B.Q. Guo, Regularity of the solution of elliptic problems with piecewise analytic data. Part I. Boundary value problems for linear elliptic equation of second order. *SIAM J. Math. Anal.* **19** (1988) 172–203.
- [4] I. Babuška and B.Q. Guo, Regularity of the solution of elliptic problems with piecewise analytic data. II: the trace spaces and application to the boundary value problems with nonhomogeneous boundary conditions. *SIAM J. Math. Anal.* **20** (1989) 763–781.
- [5] I. Babuška, R.B. Kellogg and J. Pitkäranta, Direct and inverse error estimates for finite elements with mesh refinements. *Numer. Math.* **33** (1979) 447–471.
- [6] R. Becker, M. Brunner, M. Innerberger, J.M. Melenk and D. Praetorius, Cost-optimal adaptive iterative linearized FEM for semilinear elliptic PDEs. *ESAIM Math. Model. Numer. Anal.* **57** (2023) 2193–2225.
- [7] C. Bernardi, J. Dakroub, G. Mansour and T. Sayah, A posteriori analysis of iterative algorithms for a nonlinear problem. *J. Sci. Comput.* **65** (2015) 672–697.
- [8] C. Bernardi, J. Dakroub, G. Mansour, F. Rafei and T. Sayah, Convergence analysis of two numerical schemes applied to a nonlinear elliptic problem. *J. Sci. Comput.* **71** (2017) 329–347.
- [9] M. Brunner, D. Praetorius and J. Streitberger, Cost-optimal adaptive FEM with linearization and algebraic solver for semilinear elliptic PDEs. *Numer. Math.* **157** (2025) 409–445.
- [10] P.G. Ciarlet, Linear and Nonlinear Functional Analysis with Applications. SIAM, Philadelphia, PA (2013).
- [11] S. Congreve and T.P. Wihler, Iterative Galerkin discretizations for strongly monotone problems. *J. Comput. Appl. Math.* **311** (2017) 457–472.
- [12] L. El Alaoui, A. Ern and M. Vohralík, Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems. *Comput. Methods Appl. Mech. Eng.* **200** (2011) 2782–2795.
- [13] A. Ern and M. Vohralík, Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs. *SIAM J. Sci. Comput.* **35** (2013) A1761–A1791.
- [14] G. Gantner, A. Haberl, D. Praetorius and B. Stiffler, Rate optimal adaptive FEM with inexact solver for nonlinear operators. *IMA J. Numer. Anal.* **38** (2018) 1797–1831.
- [15] G. Gantner, A. Haberl, D. Praetorius and S. Schimanko, Rate optimality of adaptive finite element methods with respect to the overall computational costs. *Math. Comput.* **90** (2021) 2011–2040.
- [16] Y. He, P. Houston, C. Schwab and T.P. Wihler, Exponential convergence of hp -ILGFEM for semilinear elliptic boundary value problems with monomial reaction. *IMA J. Numer. Anal.* **6** (2025) draf030.

- [17] P. Heid and T.P. Wihler, Adaptive iterative linearization Galerkin methods for nonlinear problems. *Math. Comput.* **89** (2018) 2707–2734.
- [18] P. Heid, D. Praetorius and T.P. Wihler, Energy contraction and optimal convergence of adaptive iterative linearized finite element methods. *Comput. Methods Appl. Math.* **21** (2021) 407–422.
- [19] D. Kinderlehrer and G. Stampacchia, An Introduction to Variational Inequalities and Their Applications. Vol. 31 of *Classics in Applied Mathematics*. SIAM, Philadelphia, PA (2000). Reprint of the 1980 original.
- [20] V.A. Kondrat'ev, Boundary problems for elliptic equations in domains with conical or angular points. *Trudy Moskov. Mat. Obšč.* **16** (1967) 209–292.
- [21] J. Moser, A sharp form of an inequality by N. Trudinger. *Indiana Univ. Math. J.* **20** (1971) 1077–1092.
- [22] C. Schwab, *p*- and *hp*-Finite Element Methods: Theory and Applications in Solid and Fluid Mechanics. *Numer. Math. Sci. Comput.* Clarendon Press (1998).
- [23] F. Tröltzsch, Optimal Control of Partial Differential Equations: Theory, Methods, and Applications. Vol. 112 of *Graduate Studies in Mathematics*. American Mathematical Society (2010).
- [24] N.S. Trudinger, On imbeddings into Orlicz spaces and some applications. *Indiana Univ. Math. J.* **17** (1967) 474–483.
- [25] B. Vexler, A priori error estimates for finite element discretization of semilinear elliptic equations with non-Lipschitz nonlinearities. *M2AN Math. Model. Numer. Anal.* **59** (2025) 1095–1112.
- [26] E.H. Zarantonello, Solving Functional Equations by Contractive Averaging. Mathematics Research Center, University of Wisconsin (1960).
- [27] E. Zeidler and L.F. Boron, Nonlinear Functional Analysis and its Applications: II/B: Nonlinear Monotone Operators. Springer (1990).



Please help to maintain this journal in open access!

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at <https://edpsciences.org/en/subscribe-to-open-s2o>.