

Exponential Convergence of hp-ILGFEM for semilinear elliptic boundary value problems with monomial reaction

Y. He and P. Houston and C. Schwab and T. Wihler

Research Report No. 2024-15
April 2024

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

EXPONENTIAL CONVERGENCE OF hp -ILGFEM FOR SEMILINEAR ELLIPTIC BOUNDARY VALUE PROBLEMS WITH MONOMIAL REACTION

YANCHEN HE, PAUL HOUSTON, CHRISTOPH SCHWAB, AND THOMAS P. WIHLE

ABSTRACT. We study the fully explicit numerical approximation of a semilinear elliptic boundary value model problem, which features a monomial reaction and analytic forcing, in a bounded polygon $\Omega \subset \mathbb{R}^2$ with a finite number of straight edges. In particular, we analyze the convergence of hp -type iterative linearized Galerkin (hp -ILG) solvers. Our convergence analysis is carried out for conforming hp -finite element (FE) Galerkin discretizations on sequences of regular, simplicial partitions of Ω , with geometric corner refinement, with polynomial degrees increasing in sync with the geometric mesh refinement towards the corners of Ω . For a sequence of discrete solutions generated by the ILG solver, with a stopping criterion that is consistent with the exponential convergence of the exact hp -FE Galerkin solution, we prove exponential convergence in $H^1(\Omega)$ to the unique weak solution of the boundary value problem. Numerical experiments illustrate the exponential convergence of the numerical approximations obtained from the proposed scheme in terms of the number of degrees of freedom as well as of the computational complexity involved.

1. INTRODUCTION

On an open bounded polygon $\Omega \subset \mathbb{R}^2$ with straight edges we aim to design and analyze a fully discrete numerical approximation scheme in terms of an exponentially convergent hp -version finite element method (hp -FEM) for the semilinear elliptic diffusion-reaction model equation

$$-\Delta u + \lambda u^{2q+1} = f \quad \text{in } \Omega, \tag{1a}$$

where $q \in \mathbb{N}_0$ is a fixed integer, $\lambda > 0$ is a positive constant, and $f \in L^{2(q+1)/(2q+1)}(\Omega)$ is a given source function (independent of u). In order to impose *Dirichlet* and *Neumann* boundary conditions, we signify the individual (straight) edges of the boundary $\Gamma := \partial\Omega$ by Γ_i , $i = 1, 2, \dots, m$; here, we suppose that each edge Γ_i is the line connecting two consecutive corners \mathbf{c}_i and \mathbf{c}_{i+1} of Ω (where we let $\mathbf{c}_{m+1} = \mathbf{c}_1$). For a *nonempty* index set $\mathcal{D} \subseteq \{1, 2, \dots, m\}$, which will be used to identify *Dirichlet* boundary edges, we define the remaining indices $\mathcal{N} := \{1, 2, \dots, m\} \setminus \mathcal{D}$ to specify the *Neumann* boundary part. In accordance with the (disjoint) decomposition $\mathcal{D} \cup \mathcal{N} = \{1, \dots, m\}$, we introduce the boundary conditions

$$u = 0 \quad \text{on } \Gamma_{\mathcal{D}} := \bigcup_{i \in \mathcal{D}} \Gamma_i, \tag{1b}$$

and

$$\partial_{\mathbf{n}} u = 0 \quad \text{on } \Gamma_{\mathcal{N}} := \bigcup_{i \in \mathcal{N}} \Gamma_i, \tag{1c}$$

where $\partial_{\mathbf{n}} u$ is the outward normal derivative along $\Gamma_{\mathcal{N}}$.

SEMINAR FOR APPLIED MATHEMATICS, ETH ZÜRICH, CH-8092 ZÜRICH, SWITZERLAND
SCHOOL OF MATHEMATICAL SCIENCES, UNIVERSITY OF NOTTINGHAM, UNIVERSITY PARK, NOTTINGHAM NG7 2RD, UK

SEMINAR FOR APPLIED MATHEMATICS, ETH ZÜRICH, CH-8092 ZÜRICH, SWITZERLAND
MATHEMATICS INSTITUTE, UNIVERSITY OF BERN, CH-3012 BERN, SWITZERLAND

Date: Draft—April 30, 2024.

TPW acknowledges the support of the Swiss National Science Foundation, Grant No. 200021_212868.

Admitting a variational formulation, see (2) below, the boundary value problem (1) will be naturally discretized by applying conforming Galerkin projections on a sequence of subspaces $\mathbb{W}_N \subset H^1_{\mathcal{D}}(\Omega)$ of finite dimension $N \sim \dim(\mathbb{W}_N)$, where

$$H^1_{\mathcal{D}}(\Omega) := \{v \in H^1(\Omega) : v|_{\Gamma_{\mathcal{D}}} = 0\}$$

is the usual Sobolev space of all weakly differentiable functions in the Lebesgue space $L^2(\Omega)$, with first-order partial derivatives in $L^2(\Omega)$, and zero boundary values along $\Gamma_{\mathcal{D}}$ in the sense of traces. Owing to the monotone structure of the nonlinear reaction term in (1), the associated standard weak formulation is uniquely solvable, both in the continuous and in the discrete setting, and a quasi-optimality property can be derived (see Propositions 1, 3, and 4 below).

The Galerkin discretization of (1) results in a nonlinear algebraic system of $\dim(\mathbb{W}_N)$ equations, whose structure depends on the subspace \mathbb{W}_N as well as on the choice of basis functions for \mathbb{W}_N , that must be solved by an appropriate numerical approximation procedure. The purpose of the present paper is to propose an iterative solver for these nonlinear equations, which is based on *linear* Galerkin discretizations in the individual steps, and to prove its global convergence. In particular, in the case where \mathbb{W}_N is realized by a sequence of *hp*-finite element subspaces, for analytic forcing f in $\bar{\Omega}$ in (1a), we show that it is possible to terminate the iterative procedure after finitely many steps in a way that the resulting numerical approximations exhibit *exponential rates of convergence* in $H^1(\Omega)$ as expressed in terms of the overall computational work (see the main result, Theorem 3, of this paper).

Previous work. Earlier results on the numerical analysis of iterative Galerkin methods for nonlinear elliptic partial differential equations (PDEs), which are closely related to the approach developed here, can be found, e.g., in (8; 9), where low-order (so-called “*h*-version”) FEM for the solution of (1) have been proposed and analyzed. In the context of *hp*-version FEM, we refer to the article (11) on strongly monotone elliptic problems. More generally, the so-called *iterative linearized Galerkin (ILG)* methodology for nonlinear PDEs, which is based on *a-posteriori* residual estimators for the Galerkin discretization error and the iterative solution of the associated nonlinear algebraic equations, has been developed in the series of papers (4; 17; 18) and the references cited therein. In addition, we refer to the works (12; 13), where the use of inexact linear solvers was also taken into account. Finally, for iterative linearized *hp*-adaptive (discontinuous) Galerkin discretizations for semilinear PDEs, we mention the paper (19).

Contribution. If the source term f in the PDE (1) is analytic in $\bar{\Omega}$, then the (unique) weak solution of (1) belongs to a class of functions that are analytic at each point in (the open domain) Ω , with quantitative control on the loss of analyticity of these functions in a vicinity of the corner points; see the recent work (16). More precisely, the weak solution belongs to a class of functions that satisfy analytic estimates in a scale of corner-weighted Hilbertian Sobolev spaces of Kondrat’ev type. As it is well-known, see, e.g., (14; 21), membership of a function in such classes allows for the establishment of *exponential approximability* in suitable discrete spaces of continuous piecewise polynomial functions in Ω on regular, simplicial partitions that are geometrically refined towards the corners of Ω . For the semilinear elliptic boundary value problem (1) under consideration, we formulate the corresponding exponential approximability result in Theorem 2 below.

The Galerkin discretization of a suitable variational formulation of (1) amounts to a possibly large nonlinear algebraic system of equations for the coefficients in the representation of the Galerkin approximation. We present an iterative scheme for the efficient numerical solution of these equations, and establish geometric convergence of the iterates in the $H^1(\Omega)$ -norm in Theorem 1 below. The proof is based on energy type arguments, and thereby, constants in the contraction rate bounds do not explicitly depend on the dimension d_N of the underlying *hp*-Galerkin approximation space.

The main achievement of our paper is a fully explicit, iterative *hp*-version finite element scheme for the numerical solution of (1), with a rigorous theoretical analysis of its exponential approximation properties and computational complexity. The key idea is to terminate the iteration in tandem with the (asymptotically) exponential convergence rates in the *hp*-Galerkin discretization, and thereby, to reach a prescribed target accuracy $0 < \varepsilon \ll 1$ in the $H^1(\Omega)$ -norm. We prove

that this procedure requires at most $\mathcal{O}(|\log(\varepsilon)|^7)$ floating point operations, which is confirmed computationally in §5 of this paper, where we report on several numerical experiments for the proposed *hp*-ILG method that are in complete agreement with the theoretical error bounds.

Notation. For a summability index $r \in (0, \infty]$, we denote by $L^r(\Omega)$ the usual spaces of Lebesgue measurable, r -integrable functions in Ω . For a nonnegative integer k , we write $H^k(\Omega)$ to be the Hilbertian Sobolev space of k -times weakly differentiable functions in Ω whose weak derivatives of order k belong to $L^2(\Omega)$, with the convention $H^0(\Omega) = L^2(\Omega)$. Furthermore, in any statements about “error-vs.-work”, the notion of “work” refers to an integer number of floating-point operations.

Outline. In §2 we discuss the weak formulation of the boundary value problem (1). In particular, in §2.3, we introduce corner-weighted Sobolev spaces of Kondrat’ev type, and revisit an analytic regularity result for the weak solution from (16). Then, §3 deals with the design and analysis of the ILG approach in abstract Galerkin spaces. Furthermore, in the context of *hp*-FE spaces, we establish the geometric rate of convergence, i.e. reaching a target tolerance $\varepsilon > 0$ in $\mathcal{O}(|\log(\varepsilon)|)$ many iteration steps. We develop an error vs. complexity analysis of the *hp*-discretization, including the nonlinear iteration with finite termination, in §4. In §5 we present some numerical examples that confirm the theoretical results within a computational context. Finally, in §6 we summarize the principal findings of this article.

2. WEAK SOLUTION

We introduce a weak formulation of (1), and establish the existence and uniqueness of weak solutions. In addition, from (16), we recapitulate the analytic regularity of the weak solution u of (1) in Ω , i.e., we discuss *a priori* estimates for u in corner-weighted Sobolev spaces of arbitrary order, for given forcing f that is analytic in Ω . Crucially, this corner-weighted, analytic regularity is known to imply exponential approximability of so-called *hp*-FE approximations which will be discussed in §3.3 below.

2.1. Weak formulation. The weak formulation of (1) is to find $u \in H_{\mathcal{D}}^1(\Omega)$ such that

$$a(u, v) + \lambda b(u; v) = \int_{\Omega} f v \, \mathbf{d}\mathbf{x} \quad \forall v \in H_{\mathcal{D}}^1(\Omega), \quad (2)$$

where we define the standard bilinear form via

$$a(v, w) := \int_{\Omega} \nabla v \cdot \nabla w \, \mathbf{d}\mathbf{x}, \quad v, w \in H_{\mathcal{D}}^1(\Omega), \quad (3)$$

and, for any given $u \in H_{\mathcal{D}}^1(\Omega)$, the linear form

$$v \mapsto b(u; v) := \int_{\Omega} u^{2q+1} v \, \mathbf{d}\mathbf{x}, \quad v \in H_{\mathcal{D}}^1(\Omega). \quad (4)$$

Remark 1. Exploiting the continuous Sobolev embedding bound

$$\|v\|_{L^{2(q+1)}(\Omega)} \leq C_q(\Omega) \|\nabla v\|_{L^2(\Omega)} \quad \forall v \in H_{\mathcal{D}}^1(\Omega), \quad (5)$$

where $C_q(\Omega) > 0$ is a constant depending only on Ω , \mathcal{D} and \mathcal{N} , and $q \in [0, \infty)$, for any fixed $u \in H_{\mathcal{D}}^1(\Omega)$, we note that the mapping $v \mapsto b(u; v)$ is a bounded linear functional on $H_{\mathcal{D}}^1(\Omega)$. Indeed, this follows immediately from Hölder’s inequality, which implies that

$$|b(u; v)| \leq \|u\|_{L^{2(q+1)}(\Omega)}^{2q+1} \|v\|_{L^{2(q+1)}(\Omega)} \leq C_q(\Omega)^{2(q+1)} \|\nabla u\|_{L^2(\Omega)}^{2q+1} \|\nabla v\|_{L^2(\Omega)},$$

for any $v \in H_{\mathcal{D}}^1(\Omega)$.

The following technical result is instrumental for our ensuing convergence analysis of the *hp*-ILG discretization of (1).

Lemma 1. For any $q \in \mathbb{N}_0$ the nonlinear form from (4) satisfies the monotonicity property

$$b(u; u - v) - b(v; u - v) \geq 0 \quad \forall u, v \in \mathbb{H}_D^1(\Omega). \quad (6)$$

Furthermore, the (local) Lipschitz continuity bound holds

$$\begin{aligned} & |b(u; v) - b(w; v)| \\ & \leq (2q + 1)C_q(\Omega)^{2(q+1)} (\|\nabla u\|_{L^2(\Omega)} + \|\nabla(u - w)\|_{L^2(\Omega)})^{2q} \|\nabla(u - w)\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)}, \end{aligned} \quad (7)$$

for any $u, v, w \in \mathbb{H}_D^1(\Omega)$, with $C_q(\Omega)$ the constant from (5).

Proof. The estimate (6) follows immediately by noticing that the nonlinear reaction term $g(t) := t^{2q+1}$, $t \in \mathbb{R}$, occurring in (1) satisfies the monotonicity property

$$(g(t_1) - g(t_2))(t_1 - t_2) \geq 0 \quad \forall t_1, t_2 \in \mathbb{R};$$

then, for any $u, v \in \mathbb{H}_D^1(\Omega)$, we infer that

$$b(u; u - v) - b(v; u - v) = \int_{\Omega} (u^{2q+1} - v^{2q+1})(u - v) \, \mathbf{d}\mathbf{x} \geq 0. \quad (8)$$

In order to derive the bound (7), for any $x, y \in \mathbb{R}$, we observe the binomial formula

$$x^{2q+1} = \sum_{k=0}^{2q+1} \binom{2q+1}{k} (x - y)^{2q+1-k} y^k,$$

or equivalently,

$$x^{2q+1} - y^{2q+1} = \sum_{k=0}^{2q} \binom{2q+1}{k} (x - y)^{2q+1-k} y^k.$$

Therefore, for any $u, v, w \in \mathbb{H}_D^1(\Omega)$, it follows that

$$|b(w; v) - b(u; v)| \leq \sum_{k=0}^{2q} \binom{2q+1}{k} \int_{\Omega} |u - w|^{2q+1-k} |u|^k |v| \, \mathbf{d}\mathbf{x}.$$

Applying a (triple) Hölder inequality, for $1 \leq k \leq 2q$, we note that

$$\begin{aligned} & \int_{\Omega} |u - w|^{2q+1-k} |u|^k |v| \, \mathbf{d}\mathbf{x} \\ & \leq \left(\int_{\Omega} |u - w|^{2(q+1)} \right)^{(2q+1-k)/2(q+1)} \left(\int_{\Omega} |u|^{2(q+1)} \right)^{k/2(q+1)} \left(\int_{\Omega} |v|^{2(q+1)} \right)^{1/2(q+1)}; \end{aligned}$$

it is straightforward to see that this bound extends to the case $k = 0$. Hence, we obtain

$$\begin{aligned} |b(u; v) - b(w; v)| & \leq \|v\|_{L^{2(q+1)}(\Omega)} \sum_{k=0}^{2q} \binom{2q+1}{k} \|u - w\|_{L^{2(q+1)}(\Omega)}^{2q+1-k} \|u\|_{L^{2(q+1)}(\Omega)}^k \\ & = \|u - w\|_{L^{2(q+1)}(\Omega)} \|v\|_{L^{2(q+1)}(\Omega)} \sum_{k=0}^{2q} \binom{2q+1}{k} \|u - w\|_{L^{2(q+1)}(\Omega)}^{2q-k} \|u\|_{L^{2(q+1)}(\Omega)}^k. \end{aligned}$$

Noticing that

$$\binom{2q+1}{k} \leq (2q+1) \binom{2q}{k}, \quad 0 \leq k \leq 2q,$$

and recalling the Sobolev embedding bound (5), completes the proof. \square

2.2. Well-posedness of weak formulation. The semilinear boundary value problem (1) admits a unique weak solution in $H_{\mathcal{D}}^1(\Omega)$.

Proposition 1 (Existence, uniqueness, and stability). *For the weak formulation (2) there exists exactly one solution $u \in H_{\mathcal{D}}^1(\Omega)$. This solution is stable in the sense that*

$$\|\nabla u\|_{L^2(\Omega)} \leq \rho, \quad (9a)$$

where

$$\rho := C_q(\Omega) \|f\|_{L^{2(q+1)/(2q+1)}(\Omega)}, \quad (9b)$$

with $C_q(\Omega) > 0$ the Sobolev embedding constant from (5).

Proof. We aim to apply the main theorem on monotone potential operators. To this end, we define the functional

$$F(v) := \frac{1}{2} \int_{\Omega} |\nabla v|^2 \, d\mathbf{x} + \frac{\lambda}{2(q+1)} \int_{\Omega} v^{2(q+1)} \, d\mathbf{x} - \int_{\Omega} f v \, d\mathbf{x}, \quad v \in H_{\mathcal{D}}^1(\Omega),$$

which is well-defined owing to the Sobolev embedding bound (5). For any $v \in H_{\mathcal{D}}^1(\Omega)$, employing Hölder's inequality, and exploiting (5), we note that

$$\left| \int_{\Omega} f v \, d\mathbf{x} \right| \leq \|f\|_{L^{2(q+1)/(2q+1)}(\Omega)} \|v\|_{L^{2(q+1)}(\Omega)} \leq C_q(\Omega) \|f\|_{L^{2(q+1)/(2q+1)}(\Omega)} \|\nabla v\|_{L^2(\Omega)} \quad \forall v \in H_{\mathcal{D}}^1(\Omega). \quad (10)$$

Hence, we deduce that

$$F(v) \geq \frac{1}{2} \|\nabla v\|_{L^2(\Omega)}^2 - C_q(\Omega) \|f\|_{L^{2(q+1)/(2q+1)}(\Omega)} \|\nabla v\|_{L^2(\Omega)} \rightarrow +\infty,$$

whenever $\|\nabla v\|_{L^2(\Omega)} \rightarrow \infty$, i.e., F is *weakly coercive*. For the Gâteaux derivative of F we have that

$$\langle F'(u), v \rangle = \int_{\Omega} \nabla u \cdot \nabla v \, d\mathbf{x} + \lambda \int_{\Omega} u^{2q+1} v \, d\mathbf{x} - \int_{\Omega} f v \, d\mathbf{x} \quad \forall u, v \in H_{\mathcal{D}}^1(\Omega),$$

i.e., the weak formulation (2) is equivalent to

$$\langle F'(u), v \rangle = 0 \quad \forall v \in H_{\mathcal{D}}^1(\Omega), \quad (11)$$

where $\langle \cdot, \cdot \rangle$ signifies the dual product. Furthermore, using (6), we observe that

$$\langle F'(u) - F'(v), u - v \rangle = \|\nabla(u - v)\|_{L^2(\Omega)}^2 + \lambda(b(u; u - v) - b(v; u - v)) \geq \|\nabla(u - v)\|_{L^2(\Omega)}^2 > 0$$

for all $u \neq v$ in $H_{\mathcal{D}}^1(\Omega)$, i.e., F' is strictly monotone. Then, applying (24, Thm. 25.F) guarantees that there exists a unique solution of the weak formulation (11). Finally, for $v = u$ in (2), applying (10), and invoking (9b), we conclude that

$$\|\nabla u\|_{L^2(\Omega)}^2 \leq a(u, u) + \lambda b(u; u) = \int_{\Omega} f u \, d\mathbf{x} \leq C_q(\Omega) \|f\|_{L^{2(q+1)/(2q+1)}(\Omega)} \|\nabla u\|_{L^2(\Omega)} \leq \rho \|\nabla u\|_{L^2(\Omega)},$$

which completes the argument. \square

2.3. Analytic regularity in scales of Hilbertian, corner-weighted Sobolev spaces. If the underlying domain Ω for the boundary value problem (1) exhibits corners then it is well-known that the inverse Dirichlet Laplace operator does not provide full elliptic regularity. Indeed, for $f \in L^2(\Omega)$, the solution belongs to $H^2(\Omega^\circ)$ in any open, interior subset Ω° , with $\overline{\Omega^\circ} \subset \Omega$. The H^2 -regularity does in general, however, not hold in a vicinity of the corners. To describe this behaviour more precisely, the scale of Hilbertian, corner-weighted Sobolev spaces $H_{\beta}^{k,l}(\Omega)$, for $k \geq l \in \mathbb{N}$, has been introduced, e.g., in (6); here, $\beta = (\beta_1, \dots, \beta_m)$ is a vector of (scalar) corner weight exponents $\beta_i \in [0, 1)$, $i = 1, \dots, m$, that are associated to the m corner points $\mathbf{c}_1, \dots, \mathbf{c}_m$ of Ω . Then, based on the corner-weight function $\Phi_{\beta}(\mathbf{x}) := \prod_{i=1}^m \text{dist}(\mathbf{x}, \mathbf{c}_i)^{\beta_i}$, for integers k and ℓ with $k \geq \ell \geq 0$, we introduce the corner-weighted Sobolev norms $\|\cdot\|_{H_{\beta}^{k,\ell}}$ in Ω via

$$\|v\|_{H_{\beta}^{k,\ell}(\Omega)}^2 := \|v\|_{H^{\ell-1}(\Omega)}^2 + \sum_{j=\ell}^k \sum_{\alpha_1+\alpha_2=j} \|\Phi_{\beta+j-\ell} \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} v\|_{L^2(\Omega)}^2, \quad (12)$$

and define corner-weighted Sobolev spaces

$$\mathbf{H}_{\beta}^{k,\ell}(\Omega) = \left\{ v \in \mathbf{H}^1(\Omega) : \|v\|_{\mathbf{H}_{\beta}^{k,\ell}(\Omega)} < \infty \right\}.$$

In (12), for $\ell = 0$, the term $\|v\|_{\mathbf{H}^{\ell-1}(\Omega)}^2$ is dropped. Referring to, e.g., (21) and the references cited therein, it also holds that

$$\mathbf{H}_{\beta}^{2,2}(\Omega) \subset C(\bar{\Omega}) \quad \text{with continuous embedding,}$$

and

$$\mathbf{H}_{\beta}^{2,2}(\Omega) \subset \mathbf{H}^1(\Omega) \quad \text{with compact embedding.}$$

Based on the corner-weighted Sobolev spaces $\mathbf{H}_{\beta}^{k,\ell}(\Omega)$ of finite order k , we introduce *corner-weighted, analytic classes* $\mathbf{B}_{\beta}^{\ell}(\Omega)$.

Definition 1 (Weighted analytic class $\mathbf{B}_{\beta}^{\ell}(\Omega)$). *Let $\ell \geq 0$ be an integer. A function v belongs to the class $\mathbf{B}_{\beta}^{\ell}(\Omega)$ if*

- (1) *v belongs to $\mathbf{H}_{\beta}^{k,\ell}(\Omega)$ for all integer $k \geq \ell$, and*
- (2) *it holds that there exist constants $C_v, d_v > 0$ such that*

$$\forall k \geq \ell : \quad \|\Phi_{\beta+k-\ell} |D^k v|\|_{L^2(\Omega)} \leq C_v d_v^{k-\ell} (k-\ell)!,$$

with the notation $|D^k v| = \sum_{\alpha_1+\alpha_2=k} |\partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} v|$.

For $1 \leq i \leq m$, we denote by $\omega_i \in (0, 2\pi)$ the interior angle of Ω at the corner \mathbf{c}_i . The following regularity result was established in (16).

Proposition 2 (Regularity of weak solution). *Let $\beta \in [0, 1)^m$ such that, for any $i \in \{1, 2, \dots, m\}$, it holds*

$$\begin{cases} \beta_i > 1 - \pi/\omega_i & \text{if } \{i-1, i\} \subset \mathcal{D} \text{ or } \{i-1, i\} \subset \mathcal{N}, \\ \beta_i > 1 - \pi/2\omega_i & \text{otherwise.} \end{cases} \quad (13)$$

Then for $f \in \mathbf{B}_{\beta}^0(\Omega)$ in (1), the solution $u \in \mathbf{H}_{\mathcal{D}}^1(\Omega)$ of (1) belongs to $\mathbf{B}_{\beta}^2(\Omega)$.

Remark 2. We note that the first bound in (13) refers to a corner point \mathbf{c}_i that connects two adjacent ‘‘Dirichlet edges’’ or two ‘‘Neumann edges’’, while the second bound applies to a corner point \mathbf{c}_i where the type of boundary condition changes.

3. ITERATIVE LINEARIZED GALERKIN (ILG) DISCRETIZATION

It is well-known that certain Galerkin discretizations of elliptic equations, which are based on sequences $\{\mathbb{W}_N\}_{N \geq 1}$ of subspaces $\mathbb{W}_N \subset \mathbf{H}_{\mathcal{D}}^1(\Omega)$ of dimension at most N , are able to achieve *exponential rates of convergence* if the underlying weak solution exhibits $\mathbf{B}_{\beta}^2(\Omega)$ -regularity as provided in Proposition 2 above; see, e.g., (21) and the reference therein. We point out, however, that such results are based on *assuming* that the corresponding nonlinear Galerkin equations are solved exactly. Evidently, this is an unrealistic hypothesis in the context of practical numerical solution methods for *nonlinear* problems; indeed, in the absence of an exact nonlinear solver, the corresponding system of $\mathcal{O}(N)$ nonlinear algebraic equations for the unknowns in the Galerkin solution $u_N \in \mathbb{W}_N$ need to be solved approximately by some iterative process that must be run to an accuracy of the order of the best approximation error. This approach will be addressed in this section.

We begin by providing an abstract error analysis of approximate solutions in generic, closed Galerkin subspaces of $\mathbf{H}_{\mathcal{D}}^1(\Omega)$. Subsequently, we consider the specific context of *hp*-FE discretizations in §4.

3.1. Finite-dimensional Galerkin approximations. In the sequel, let $\mathbb{W}_N \subset H_D^1(\Omega)$ denote any *finite-dimensional* (and thus closed) subspace, and consider the Galerkin discretization of the boundary value problem (1) on \mathbb{W}_N in weak form: find $U \in \mathbb{W}_N$ such that

$$a(U, v) + \lambda b(U; v) = \int_{\Omega} f v \, d\mathbf{x} \quad \forall v \in \mathbb{W}_N. \quad (14)$$

The following result follows verbatim as in the proof of Proposition 1.

Proposition 3 (Well-posedness of Galerkin discretization). *There exists a unique solution $U \in \mathbb{W}_N$ of the discrete weak formulation (14), and the stability bound $\|\nabla U\|_{L^2(\Omega)} \leq \rho$ holds, with ρ from (9b).*

The Galerkin approximation U is a quasi-optimal approximation of u .

Proposition 4 (Quasi-optimality). *For the error between the solution $u \in H_D^1(\Omega)$ of (2) and its Galerkin approximation $U \in \mathbb{W}_N$ of (14), the following bound holds*

$$\|\nabla(u - U)\|_{L^2(\Omega)} \leq C \inf_{w \in \mathbb{W}_N} \|\nabla(u - w)\|_{L^2(\Omega)},$$

for a constant $C > 0$ depending only on λ , q , Ω , and f in (1).

Proof. Introducing the error $\mathbf{e}_{\mathbb{W}_N} := u - U$, we observe the *Galerkin orthogonality* of $\mathbf{e}_{\mathbb{W}_N}$:

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

Thereby, for any $w \in \mathbb{W}_N$, the following identity holds

$$\begin{aligned} \|\nabla \mathbf{e}_{\mathbb{W}_N}\|_{L^2(\Omega)}^2 &= a(\mathbf{e}_{\mathbb{W}_N}, u - w) + a(\mathbf{e}_{\mathbb{W}_N}, w - U) \\ &= a(\mathbf{e}_{\mathbb{W}_N}, u - w) - \lambda b(u; w - U) + \lambda b(U; w - U) \\ &= a(\mathbf{e}_{\mathbb{W}_N}, u - w) - \lambda b(u; u - U) + \lambda b(U; u - U) - \lambda b(u; w - u) + \lambda b(U; w - u), \end{aligned}$$

for the exact Galerkin approximation $U \in \mathbb{W}_N$. Exploiting the monotonicity property stated in (6), we observe that

$$\|\nabla \mathbf{e}_{\mathbb{W}_N}\|_{L^2(\Omega)}^2 \leq \|\nabla \mathbf{e}_{\mathbb{W}_N}\|_{L^2(\Omega)} \|\nabla(u - w)\|_{L^2(\Omega)} - \lambda b(u; w - u) + \lambda b(U; w - u).$$

Recalling the stability estimate (9a), which holds for both u and U , cf. Proposition 3, and applying (7), we deduce that

$$\|\nabla \mathbf{e}_{\mathbb{W}_N}\|_{L^2(\Omega)} \leq \|\nabla(u - w)\|_{L^2(\Omega)} + \lambda(2q + 1)C_q(\Omega)^{2(q+1)}(3\rho)^{2q} \|\nabla(u - w)\|_{L^2(\Omega)},$$

with ρ from (9b). Since $w \in \mathbb{W}_N$ was chosen arbitrarily, the proof is complete. \square

3.2. Iterative solution. We consider an iterative linearization for the solution of the discrete Galerkin formulation (14) in terms of a *Picard scheme (also called Zarantanello iteration)*, cf. (11; 18): starting from any initial guess $U_0 \in H_D^1(\Omega)$, for a (fixed) parameter $0 < \alpha \leq 1$, consider a sequence $\{U_n\}_{n \geq 1} \subset \mathbb{W}_N$ that is generated by the iteration

$$a(U_{n+1}, v) = (1 - \alpha)a(U_n, v) + \alpha \left(\int_{\Omega} f v \, d\mathbf{x} - \lambda b(U_n; v) \right) \quad \forall v \in \mathbb{W}_N, \quad (15)$$

for $n \geq 0$.

Proposition 5 (Contraction of ILG procedure). *For any initial guess $U_0 \in H_D^1(\Omega)$, the iteration (15), with*

$$\alpha \in (0, 1] \cap (0, 2/(L^2 + 1)), \quad (16)$$

where

$$L := \lambda(2q + 1)C_q(\Omega)^{2(q+1)}(2\rho + \|\nabla U_0\|_{L^2(\Omega)})^{2q}, \quad (17)$$

satisfies the bound

$$\|\nabla(U - U_{n+1})\|_{L^2(\Omega)} \leq r_{\alpha}^{n+1} \|\nabla(U - U_0)\|_{L^2(\Omega)} \quad \forall n \geq 0,$$

with the contraction constant

$$r_{\alpha} := \sqrt{(1 - \alpha)^2 + \alpha^2 L^2} < 1. \quad (18)$$

Here, $U \in \mathbb{W}_N$ is the solution of the Galerkin formulation (14), ρ is defined in (9b), and $C_q(\Omega)$ is the Sobolev embedding constant from (5). In particular, the iteration (15) converges strongly in $H_{\mathcal{D}}^1(\Omega)$, i.e., $\lim_{n \rightarrow \infty} \|\nabla(U - U_n)\|_{L^2(\Omega)} = 0$.

Remark 3. For the value $\alpha_\star := (1 + L^2)^{-1} < 1$, with L from (17), we note that $r_{\min} := r_{\alpha_\star} = L(1 + L^2)^{-1/2} < 1$ is the minimal value of r_α in (18).

Proof of Proposition 5. We proceed by induction with respect to $n \geq 0$. To this end, we follow along the lines of the classical theory for monotone operators, see, e.g., (20, §3.3), whereby we work with a *local* Lipschitz continuity property similar to the analysis in (7). Let $R := \|\nabla(U - U_0)\|_{L^2(\Omega)}$, and suppose that

$$\|\nabla(U - U_k)\|_{L^2(\Omega)} \leq Rr_\alpha^k, \quad \text{for } k = 0, \dots, n, \quad (19)$$

which is obviously true for $n = 0$. Then, by virtue of Remark 1, we are able to define a unique (Riesz representer) $\xi_n \in \mathbb{W}_N$ by the weak formulation

$$a(\xi_n, v) = b(U; v) - b(U_n; v) \quad \forall v \in \mathbb{W}_N. \quad (20)$$

In light of (7) we note that

$$\begin{aligned} & |b(U; v) - b(U_n; v)| \\ & \leq (2q + 1)C_q(\Omega)^{2(q+1)} (\|\nabla U\|_{L^2(\Omega)} + \|\nabla(U - U_n)\|_{L^2(\Omega)})^{2q} \|\nabla(U - U_n)\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)}, \end{aligned}$$

for any $v \in H_{\mathcal{D}}^1(\Omega)$. Furthermore, by virtue of Proposition 3 and due to the induction assumption (19), it follows that

$$|b(U; v) - b(U_n; v)| \leq (2q + 1)C_q(\Omega)^{2(q+1)} (\rho + R)^{2q} \|\nabla(U - U_n)\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)}, \quad (21)$$

for any $v \in H_{\mathcal{D}}^1(\Omega)$, i.e., the right-hand side of (20) is a bounded linear functional with respect to v , and thus, $\xi_n \in \mathbb{W}_N$ is well-defined. In addition, letting $v = \xi_n$ in (20) and (21), we infer that

$$\|\nabla \xi_n\|_{L^2(\Omega)} \leq (2q + 1)C_q(\Omega)^{2(q+1)} (\rho + R)^{2q} \|\nabla(U - U_n)\|_{L^2(\Omega)}. \quad (22)$$

Moreover, upon subtracting the weak formulations (14) and (15), we observe that

$$\begin{aligned} a(U - U_{n+1}, v) &= (1 - \alpha)a(U - U_n, v) + \alpha\lambda(b(U_n; v) - b(U; v)) \\ &= (1 - \alpha)a(U - U_n, v) - \alpha\lambda a(\xi_n, v), \end{aligned}$$

for any $v \in \mathbb{W}_N$, which implies the identity

$$U - U_{n+1} = (1 - \alpha)(U - U_n) - \alpha\lambda\xi_n.$$

Thus,

$$\begin{aligned} & \|\nabla(U - U_{n+1})\|_{L^2(\Omega)}^2 \\ &= (1 - \alpha)^2 \|\nabla(U - U_n)\|_{L^2(\Omega)}^2 - 2\alpha(1 - \alpha)\lambda a(\xi_n, U - U_n) + \alpha^2\lambda^2 \|\nabla \xi_n\|_{L^2(\Omega)}^2. \end{aligned}$$

Inserting $v = U - U_n$ in (20), and using (6), we notice that $a(\xi_n, U - U_n) \geq 0$. Therefore, recalling (22), we deduce that

$$\|\nabla(U - U_{n+1})\|_{L^2(\Omega)}^2 \leq \left((1 - \alpha)^2 + \alpha^2\lambda^2(2q + 1)^2 C_q(\Omega)^{4(q+1)} (\rho + R)^{4q} \right) \|\nabla(U - U_n)\|_{L^2(\Omega)}^2.$$

Employing Proposition 3, gives

$$R \leq \|\nabla U\|_{L^2(\Omega)} + \|\nabla U_0\|_{L^2(\Omega)} \leq \rho + \|\nabla U_0\|_{L^2(\Omega)},$$

and thereby the bound $\|\nabla(U - U_{n+1})\|_{L^2(\Omega)} \leq Rr_\alpha^{n+1}$ holds, which concludes the induction argument. \square

3.3. Exponential convergence of the ILG iteration. We consider sequences $\{\mathbb{W}_N\}_{N \geq 1} \subset \mathbf{H}_{\mathcal{D}}^1(\Omega)$ of subspaces \mathbb{W}_N of finite dimension $\dim(\mathbb{W}_N) \sim N$ which allow for exponentially convergent approximations of functions $u \in \mathbf{B}_{\beta'}^2(\Omega)$, where we set $\beta' := (\beta'_1, \dots, \beta'_m) \in [0, 1]^m$ with

$$\beta'_i := \frac{1}{2} (1 + \max(\beta_i, 1 - \pi/2\omega_i)), \quad 1 \leq i \leq m. \quad (23)$$

More precisely, we suppose that there exist constants $K, b, \gamma > 0$ independent of N with

$$\inf_{w \in \mathbb{W}_N} \|\nabla(u - w)\|_{\mathbf{L}^2(\Omega)} \leq K \exp(-bd_N^\gamma), \quad \text{where } d_N \geq \dim(\mathbb{W}_N), N \geq 0. \quad (24)$$

Specific choices of $\{\mathbb{W}_N\}_{N \geq 1}$ that realize (24) with $\gamma = 1/3$ will be given in §4 below.

Theorem 1 (Convergence of Picard iteration). *Suppose that the right-hand side in (1) satisfies $f \in \mathbf{B}_{\beta}^0(\Omega)$. Furthermore, consider a family of Galerkin spaces $\{\mathbb{W}_N\}_{N \geq 1}$ that features the exponential approximation property (24). Then, for any fixed $N \geq 1$ and any initial guess $U_0 \in \mathbf{H}_{\mathcal{D}}^1(\Omega)$, performing $n = \mathcal{O}(d_N^\gamma)$ steps of the Picard iteration scheme (15), with suitable $\alpha \in (0, 1]$ cf. Proposition 5, leads to the error estimate*

$$\|\nabla(u - U_n)\|_{\mathbf{L}^2(\Omega)} \leq C \exp(-\kappa d_N^\gamma), \quad (25)$$

where $u \in \mathbf{H}_{\mathcal{D}}^1(\Omega)$ is the weak solution of (1), with constants $C, \kappa > 0$ that are independent of N , and with $\gamma > 0$ from (24).

Proof. If $\beta' \in (0, 1)^m$ is defined as in (23) then $f \in \mathbf{B}_{\beta'}^0(\Omega)$. Owing to Proposition 2, we note that the solution of (2) satisfies $u \in \mathbf{B}_{\beta'}^2(\Omega)$. Then, for any subspace \mathbb{W}_N and any $n \geq 0$, using the triangle inequality, we have

$$\|\nabla(u - U_n)\|_{\mathbf{L}^2(\Omega)} \leq \|\nabla(u - U)\|_{\mathbf{L}^2(\Omega)} + \|\nabla(U - U_n)\|_{\mathbf{L}^2(\Omega)},$$

where $U \in \mathbb{W}_N$ is the solution of (14). The first term on the right-hand side of the above inequality can be estimated by combining the quasi-optimality property from Proposition 4 and the exponential approximation bound (24), thereby yielding

$$\|\nabla(u - U)\|_{\mathbf{L}^2(\Omega)} \leq C_1 \exp(-bd_N^\gamma), \quad (26)$$

for a constant $C_1 > 0$. Furthermore, the second term is bounded by employing Proposition 5, i.e., we have that

$$\|\nabla(U - U_n)\|_{\mathbf{L}^2(\Omega)} \leq C_2 r_\alpha^n, \quad (27)$$

with constants $0 < r_\alpha < 1$ and $C_2 > 0$. Choosing $n = \mathcal{O}(d_N^\gamma)$ shows that

$$r_\alpha^n \leq \exp(-b'd_N^\gamma), \quad (28)$$

for some $b' > 0$ independent of n and N , which completes the proof. \square

4. EXPONENTIAL CONVERGENCE AND ε -COMPLEXITY OF hp-ILGFEM

Based on the convergence result for the Picard iteration (15) from Theorem 1, we are ready to formulate a fully discrete ILG scheme, and to establish its exponential convergence, respectively, polylogarithmic ε -complexity for the numerical approximation of the semilinear boundary value problem (1) in a polygon Ω with analytic data $f \in \mathbf{B}_{\beta}^0(\Omega)$. We first discuss the design of (sequences of) so-called hp-FE spaces \mathbb{W}_N , which, together with the analytic regularity $u \in \mathbf{B}_{\beta'}^2(\Omega)$ stated in Proposition 2, affords the exponential convergence rate (24) with $\gamma = 1/3$ for the corresponding exact Galerkin solutions $U \in \mathbb{W}_N$, cf. (14). We then address the realization of the iteration (15), in particular, the solver complexity of the linear part based on the hp-FE spaces \mathbb{W}_N . Here, we emphasize that the Laplace operator and the monomial nonlinearity in (1) allow for an *exact computation* (assuming absence of rounding errors) of all bilinear and nonlinear forms (in terms of $a(\cdot, \cdot)$ and $b(\cdot; \cdot)$, respectively) in the ILG iteration (15), and thus, to dispense with a quadrature error analysis.

4.1. hp -Approximations on geometric corner meshes. Under the analytic $B_{\beta}^2(\Omega)$ -regularity in scales of corner-weighted Sobolev spaces of the weak solution $u \in H_{\mathcal{D}}^1(\Omega)$ of (1), which has been shown in (16), see also §2.3, subspace sequences $\{\mathbb{W}_N\}_{N \geq 1}$ in (14) that satisfy (24) with $\gamma = 1/3$ have been constructed in (14; 21) and the references cited therein. We briefly recapitulate their construction here.

For integers $k \geq 1$, we denote by \mathcal{T}_k a regular, simplicial partition of the polygon Ω into at most $\mathcal{O}(k)$ open triangles T . We assume that the partitions $\{\mathcal{T}_k\}_{k \geq 1}$ are nested, uniformly shape regular, and *geometrically refined* to the corners $\mathcal{C} = \{\mathbf{c}_i\}_{i=1}^m \subset \partial\Omega$ of the polygon Ω ; the last property refers to the existence of constants $C > 0$ and $\sigma \in (0, 1)$ such that it holds

$$\forall k \ \forall T \in \mathcal{T}_k : \bar{T} \cap \mathcal{C} = \emptyset : \quad 0 < \sigma \leq \text{diam}(T)/\text{dist}(T, \mathcal{C}) \leq \sigma^{-1}, \quad (29a)$$

and

$$\forall k \ \forall T \in \mathcal{T}_k : \bar{T} \cap \mathcal{C} \neq \emptyset : \quad \text{diam}(T) \leq C\sigma^k. \quad (29b)$$

Geometric partitions satisfying the above conditions can be constructed in any polygon Ω by recursive bisection refinement departing from a regular, admissible initial triangulation \mathcal{T}_1 of Ω . We shall refer to triangulations generated in this way as *geometric corner meshes* in Ω . The aforementioned recursive bisection refinement constructions yield the existence of a constant $C > 0$ (depending on Ω and on \mathcal{T}_1) such that

$$\forall k \in \mathbb{N} : \quad \#(\mathcal{T}_k) \leq Ck.$$

With the geometric corner meshes in hand, the spaces \mathbb{W}_N of dimension $d_N = \mathcal{O}(N)$ consist of continuous functions in $\bar{\Omega}$ whose restriction to each element $T \in \mathcal{T}_k$, $k \geq 1$, are polynomials of uniform total polynomial degree $p_k \geq 1$, and obey the homogenous Dirichlet boundary condition (1b). The spaces \mathbb{W}_N are built on the sequences $\{\mathcal{T}_k\}_{k \geq 1}$ of geometric corner meshes

$$\mathbb{W}_N = \mathbb{S}_{\mathcal{D}}^{p_k}(\Omega; \mathcal{T}_k) = \{v \in H_{\mathcal{D}}^1(\Omega) \mid \forall T \in \mathcal{T}_k : v|_T \in \mathbb{P}^{p_k}(T)\}, \quad (30)$$

with $\mathbb{P}^{p_k}(T)$ signifying the space of all polynomials of total degree at most p_k over T , $T \in \mathcal{T}_k$, $k \geq 1$, whereby we choose $p_k \simeq ck$, for some fixed sufficiently large constant $c > 0$. Then it follows that there exists a constant $C > 0$ such that

$$d_N = \dim(\mathbb{S}_{\mathcal{D}}^{p_k}(\Omega; \mathcal{T}_k)) \leq Ck^3 \simeq p_k^3,$$

for all N . Based on the above construction (30) of the hp -FE spaces \mathbb{W}_N , the corresponding Galerkin discretizations of (1), cf. (14), exhibit the exponential approximability (24) with $\gamma = 1/3$.

Theorem 2 (Exponential convergence of hp -Galerkin approximations). *For $\beta \in [0, 1)^m$, consider the semilinear boundary value problem (1) with data $f \in B_{\beta}^0(\Omega)$. Then, for the Galerkin projection of the weak solution $u \in H_{\mathcal{D}}^1(\Omega)$ of (1) in the hp -FE spaces \mathbb{W}_N of dimension $d_N = \mathcal{O}(N)$ from (30), there exists constants $b, C > 0$ (generally depending on u) such that the following exponential approximability bound holds*

$$\inf_{w_N \in \mathbb{W}_N} \|\nabla(u - w_N)\|_{L^2(\Omega)} \leq C \exp(-bN^{1/3}). \quad (31)$$

In particular the exact Galerkin solutions $U \in \mathbb{W}_N$ in (14), whose existence and uniqueness was ensured in Proposition 3, satisfy the exponential convergence bound (24) with $\gamma = 1/3$.

Proof. Based on the assumed analytic regularity $f \in B_{\beta}^0(\Omega)$ on the data, by Proposition 2 and (16), the unique weak solution $u \in H_{\mathcal{D}}^1(\Omega)$ of (1) belongs to the corner-weighted analytic class $B_{\beta}^2(\Omega)$. The exponential approximability (31) is then a direct consequence of well-known approximation properties of the hp -FE spaces \mathbb{W}_N introduced above, see e.g. (14; 21) and the references cited therein. \square

Remark 4. For subspaces $\mathbb{W}_N \subset H_{\mathcal{D}}^1(\Omega)$ spanned by continuous, piecewise polynomial functions on regular, simplicial partitions \mathcal{T}_k , $k \geq 1$, of the polygon Ω , under the present assumption of the polynomial nonlinearity in the reaction term in (1), we emphasize that the stiffness matrix corresponding to the Galerkin discretization of the bilinear form $a(\cdot, \cdot)$ from (3), and the discrete

representation of the nonlinear form $b(\cdot; \cdot)$ from (4) can be evaluated exactly, e.g., by Gaussian quadrature rules of sufficiently high order applied on each element $T \in \mathcal{T}_k$, $k \geq 1$.

4.2. Solution of the nonlinear algebraic system. On any given finite-dimensional subspace $\mathbb{W}_N \subset H_D^1(\Omega)$, we emphasize that the solution of the *linear* algebraic system resulting from the ILG scheme (15) features the *same* stiffness matrix (associated to the Laplace operator) in each iteration step. On a related note, for $n \geq 0$, exploiting Remark 1, we can introduce a unique (Riesz representative) $\eta_n \in \mathbb{W}_N$ by

$$a(\eta_n, v) = \int_{\Omega} f v \, d\mathbf{x} - \lambda b(U_n; v) \quad \forall v \in \mathbb{W}_N, \quad (32)$$

which allows us to write the iterative scheme (15) in the strong form

$$U_{n+1} = (1 - \alpha)U_n + \alpha\eta_n, \quad n \geq 0. \quad (33)$$

We will now estimate the operation count incurred in the implementation of iteration (33). We begin our discussion with the observation that, on a given Galerkin space \mathbb{W}_N , the Cholesky decomposition of the stiffness matrix corresponding to the bilinear form $a(\cdot, \cdot) : \mathbb{W}_N \times \mathbb{W}_N \rightarrow \mathbb{R}$ from (3) can be performed once and for all, so that the computational work is limited to one backsolve only in each step of (33). Without any further assumption, the Cholesky decomposition of the matrix corresponding to the bilinear form $a(\cdot, \cdot)$ incurs work of $\mathcal{O}(d_N^3)$, where $d_N = \dim(\mathbb{W}_N) = \mathcal{O}(N)$. From Theorem 1, the number n of iterations of (33) required to reach the approximation error bound (26) is given by $\mathcal{O}(d_N')$.

4.2.1. Evaluation cost of the nonlinear term. In the sequel, for ease of notation, for meshes \mathcal{T}_k with an associated polynomial degree p_k , cf. §4.1, we will simply write \mathcal{T}_p and p , instead, respectively. Recalling that $\Omega \subset \mathbb{R}^2$ is a polygon, and that \mathbb{W}_N in (30) is a space of continuous, piecewise polynomial functions on a regular triangulation \mathcal{T}_p of Ω with a corresponding (uniform throughout \mathcal{T}_p) total degree $p \geq 1$, we require $\mathcal{O}(p^2)$ many Gauss points in each triangle $T \in \mathcal{T}_p$ for an *exact integration* of the (polynomial) nonlinear form $\int_T w^{2q+1} v \, d\mathbf{x}$, for $w, v \in \mathbb{P}^p(T)$; here, our focus is on high-order hp -approximations where $p \gg q$, and thus

$$\deg(w^{2q+1} v) \leq 2p(q+1) = \mathcal{O}(p).$$

Therefore, in total, the evaluation of the nonlinear form

$$b(w; v) = \sum_{T \in \mathcal{T}_p} b_T(w; v) := \sum_{T \in \mathcal{T}_p} \int_T w^{2q+1} v \, d\mathbf{x}, \quad v, w \in \mathbb{W}_N,$$

occurring on the right-hand side of (32), amounts to a work of $\mathcal{O}(\#\mathcal{T}_p)p^2$. In addition, if the right-hand side source function f in (1) is a polynomial (which we henceforth assume) then we note that the evaluation of the linear form $\int_{\Omega} f v \, d\mathbf{x}$ amounts to the same cost.

4.2.2. Error vs. work for dense Cholesky decomposition. Assuming that we have at hand the Cholesky factors of the global stiffness matrix associated to the bilinear form $a(\cdot, \cdot)$, which requires $\mathcal{O}(N^3)$ operations *once only*, each iteration step in (32) necessitates only one backsolve of $\mathcal{O}(N^2)$ work as well as one evaluation of the nonlinearity. Then, in view of our considerations in §4.2.1, the total work for the approximate solution of the nonlinear algebraic system using n iteration steps amounts to a computational work of

$$\mathcal{O}(N^3) + n(\mathcal{O}(N^2) + \mathcal{O}(\#\mathcal{T}_p)p^2).$$

In case of hp -FE approximations with $\#\mathcal{T}_p = \mathcal{O}(p)$ many triangles and $d_N = \mathcal{O}(N) = \mathcal{O}(p^3)$, $\gamma = 1/3$, cf. Theorem 2, this transforms to $\mathcal{O}(p^9) + n(\mathcal{O}(p^6) + \mathcal{O}(p^3))$. Therefore, in light of Theorem 1, applying $n = \mathcal{O}(d_N') = \mathcal{O}(N^{1/3}) = \mathcal{O}(p)$ iterations in (32) to obtain consistency with the Galerkin discretization error, a total work of

$$\mathcal{O}(p^9) + p(\mathcal{O}(p^6) + \mathcal{O}(p^3)) = \mathcal{O}(p^9) + \mathcal{O}(p^7) \quad (34)$$

many operations are required. In particular, if we do not exploit any sparsity structure in the hp -FE matrices of the linear principal part of (1), the total work for the computation of the approximate

Galerkin solution to the accuracy of the hp -FE discretization error (with exact solution of the Galerkin equations) is still dominated by the work for the linear solve $\mathcal{O}(N^3)$.

4.2.3. Static condensation Cholesky decomposition. We discuss how to reduce the above operation count by exploiting a suitable separation of the polynomial basis functions on each triangle $T \in \mathcal{T}_p$, $p \geq 1$, into 3 *nodal*, $\mathcal{O}(p)$ *face* and $\mathcal{O}(p^2)$ *internal modes*; see (10; 15; 21; 23) for details. Using static condensation on each of the $\mathcal{O}(p)$ elements in the geometric mesh \mathcal{T}_p , $p \geq 1$, and noting that

$$\#\text{dof}/\#\text{elements} = \mathcal{O}(p^3)/\mathcal{O}(p) = \mathcal{O}(p^2),$$

costs $\mathcal{O}(p^6)$ flops per element $T \in \mathcal{T}_p$, $p \geq 1$. Hence, multiplying by the number of elements, results in the following result.

Proposition 6. *The work for a global Cholesky decomposition of the stiffness matrix (corresponding to the bilinear form $a(\cdot, \cdot)$ from (3)) on the hp -FE space $\mathbb{S}_D^p(\Omega; \mathcal{T}_p)$, $p \geq 1$, with prior elementwise static condensation requires asymptotically, as $p \rightarrow \infty$, $\mathcal{O}(p^7)$ flops.*

Proof. We apply elementwise static condensation which yields only $\mathcal{O}(p)$ “external” unknowns on $\mathcal{O}(p)$ inter-element edges in \mathcal{T}_p . Consequently, upon condensation, a global dense linear system of size $\mathcal{O}(p^2)$ is obtained. The work of a Cholesky decomposition for this condensed system scales as $\mathcal{O}((p^2)^3) = \mathcal{O}(p^6)$ flops. This, in turn, implies that work for a nested dissection version of the initial Cholesky decomposition of the bilinear form $a(\cdot, \cdot) : \mathbb{W}_N \times \mathbb{W}_N \rightarrow \mathbb{R}$ can be reduced from $\mathcal{O}(d_N^3) = \mathcal{O}(p^9)$ to $\mathcal{O}(p^7)$ flops. \square

Remark 5. In particular, the above Proposition 6 shows that the work for the initial Cholesky decomposition of the condensed stiffness matrix of the linear part $a(\cdot, \cdot) : \mathbb{W}_N \times \mathbb{W}_N \rightarrow \mathbb{R}$ is asymptotically (as $d_N \rightarrow \infty$) of the same order as the work required to run the iteration (32) on each Galerkin space $\mathbb{W}_N = \mathbb{S}_D^p(\Omega; \mathcal{T}_p)$ in $n(\mathcal{O}(N^2) + \mathcal{O}(\#\mathcal{T}_p)p^2)) = \mathcal{O}(p^7)$ operations, cf. (34), to termination at “ hp approximation accuracy” (31).

Remark 6 (Static condensation Cholesky decomposition exploiting affine equivalence). The argument in Proposition 6 is not sharp in the sense that it does not exploit the constant-coefficient structure of the linear principal part of the operator in (1). Indeed, in this case, in the element stiffness matrix generation there is no necessity to employ numerical integration. In contrast, for a linear principal part of the form $-\text{div}(\mathbf{A}(\mathbf{x})\nabla(\cdot))$ with *non-constant coefficients*, i.e., for a diffusion matrix $\mathbf{A} : \Omega \rightarrow \mathbb{R}_{\text{sym}}^{2 \times 2}$ with variable coefficient functions a_{ij} that are analytic in $\bar{\Omega}$, numerical quadrature in the element stiffness matrices is necessary. The work needed to build the global stiffness matrix then scales by a constant factor compared to the work required to compute the stiffness matrix in the constant coefficient case. The use of quadrature will cause a further consistency error that is, however, of the same order as the discretization error bounds. We therefore expect the present error vs. work bound to hold also in this more general setting.

In addition, assuming that \mathcal{T}_p , $p \geq 1$, is a regular partition of Ω into affine-equivalent triangles, all element stiffness matrices are affine equivalent to the reference element stiffness matrix. Therefore, element stiffness matrix Cholesky factorization and Schur complement formation could be done once and for all on the reference element. This would reduce the cost bound of nested dissection Cholesky factorization of the stiffness matrix for the linear principal part to $\mathcal{O}(p^6)$. We refer to (10) for further exploitation of reference element stiffness and mass matrix sparsity afforded by the choice of a particular set of shape functions. As the total work for the evaluation of the nonlinearity during the ILG iteration is already $\mathcal{O}(p^7)$, we do not account for this lower complexity in the present article.

4.3. Exponential convergence of fully discrete iterative linearized hp -FE approximation. We will now investigate the complexity of the fully discrete iteration (15). To this end, in order to deal with potential quadrature errors of integrals involving the right-hand side function f in (1a), we make the following assumption:

Assumption 1. For a sequence $\{\mathcal{T}_p\}_{p \geq 1}$ of regular partitions of Ω into triangles, with geometric corner refinement and associated polynomial degrees $p = 1, 2, \dots$, cf. §4.1–4.2, we suppose that the right-hand side function $f \in L^2(\Omega) \cap B_{\mathcal{B}}^0(\Omega)$ in (1a) can be approximated by a corresponding

sequence of functions $\Pi_{\mathbf{p}}(f)$ with $\Pi_{\mathbf{p}}(f)|_T \in \mathbb{P}^{\mathbf{p}}(T)$ for all $T \in \mathcal{T}_p$, $p = 1, 2, \dots$, where $\mathbf{p} = \mathcal{O}(p)$, such that

$$\|f - \Pi_{\mathbf{p}}(f)\|_{\mathbf{L}^2(\Omega)} \leq C \exp(-cN^{1/3}), \quad (35)$$

with $N = N(p) = \dim(\mathbb{S}_{\mathcal{D}}^p(\Omega; \mathcal{T}_p)) = \mathcal{O}(p^3)$, as in (31), and constants $c, C > 0$.

Remark 7. This assumption is satisfied, evidently, if $f \in \mathbf{L}^2(\Omega) \cap \mathbf{B}_{\beta}^0(\Omega)$ is a polynomial function in Ω , and, in particular, for all functions that are real-analytic in $\bar{\Omega}$. More general conditions and fully discrete approximation methods, which are based on point-queries of $\mathbf{A}(\mathbf{x})$, cf. Remark 6, and $f(\mathbf{x})$ in Ω , will be studied in a forthcoming article.

Instead of the discrete iteration (15), which relies on exact integration of f , we now consider the *fully discrete hp-ILG scheme* given by

$$a(\mathcal{U}_{n+1}, v) = (1 - \alpha)a(\mathcal{U}_n, v) - \alpha \lambda b(\mathcal{U}_n; v) + \alpha \sum_{T \in \mathcal{T}_p} Q_{T, \mathbf{p}}(\Pi_{\mathbf{p}}(f)v) \quad \forall v \in \mathbb{W}_N, \quad (36)$$

where $Q_{T, \mathbf{p}}$ is an elementwise quadrature rule on each triangle $T \in \mathcal{T}_p$, which is based on $\mathcal{O}(p^2)$ (e.g. Gauss-type) quadrature points, and integrates any polynomial in $\mathbb{P}^{\mathbf{p}+p}(T)$ *exactly*. Similar to our previous discussion in §4.2.1, we note that the work for the evaluation of the quadrature in (36) amounts to $\mathcal{O}(\#\mathcal{T}_p)p^2$.

Lemma 2. *Let α satisfy (16), and suppose that the iterations (15) and (36) are initiated with the same starting guess $U_0 \in \mathbb{W}_N$. Then, under Assumption 1, for the difference between the corresponding iterative solutions U_n and \mathcal{U}_n the following bound holds*

$$\|\nabla(U_n - \mathcal{U}_n)\|_{\mathbf{L}^2(\Omega)} \leq C \exp(-\kappa N^{1/3}), \quad n \geq 0,$$

with constants $C, \kappa > 0$ independent of N .

Proof. We consider the auxiliary problem of finding $\tilde{U} \in \mathbb{W}_N$ such that

$$a(\tilde{U}, v) + \lambda b(\tilde{U}; v) = \sum_{T \in \mathcal{T}_p} Q_{T, \mathbf{p}}(\Pi_{\mathbf{p}}(f)v) \quad \forall v \in \mathbb{W}_N.$$

By Assumption 1, we observe that

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

This shows that Proposition 5 can be applied to the perturbed iteration (36), with the *same* contraction constant r_{α} from (18). In particular, we have the bound

$$\|\nabla(\tilde{U} - \mathcal{U}_{n+1})\|_{\mathbf{L}^2(\Omega)} \leq r_{\alpha}^{n+1} \|\nabla(\tilde{U} - U_0)\|_{\mathbf{L}^2(\Omega)} \quad \forall n \geq 0.$$

In addition, for the difference of \tilde{U} and the solution U of (14) it holds

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

Testing with $v = U - \tilde{U} \in \mathbb{W}_N$, and applying (8), leads to

$$\|\nabla(U - \tilde{U})\|_{\mathbf{L}^2(\Omega)}^2 \leq \|f - \Pi_{\mathbf{p}}(f)\|_{\mathbf{L}^2(\Omega)} \|U - \tilde{U}\|_{\mathbf{L}^2(\Omega)},$$

which, by means of the Poincaré inequality, results in

$$\|\nabla(U - \tilde{U})\|_{\mathbf{L}^2(\Omega)} \leq C_P \|f - \Pi_{\mathbf{p}}(f)\|_{\mathbf{L}^2(\Omega)},$$

for a constant $C_P > 0$ only depending on Ω . Hence, owing to the triangle inequality, we deduce that

$$\begin{aligned} \|\nabla(U - \mathcal{U}_{n+1})\|_{\mathbf{L}^2(\Omega)} &\leq \|\nabla(U - \tilde{U})\|_{\mathbf{L}^2(\Omega)} + \|\nabla(\tilde{U} - \mathcal{U}_{n+1})\|_{\mathbf{L}^2(\Omega)} \\ &\leq C_P \|f - \Pi_{\mathbf{p}}(f)\|_{\mathbf{L}^2(\Omega)} + r_{\alpha}^{n+1} \|\nabla(\tilde{U} - U_0)\|_{\mathbf{L}^2(\Omega)}. \end{aligned}$$

The first term on the right-hand side of the above bound can be estimated with (35), whilst the second term can be dealt with as in (27)–(28) with $d_N = \mathcal{O}(N)$ and $\gamma = 1/3$, see §4.1. \square

We are now prepared to state and prove our main result.

Theorem 3 (Polylogarithmic ε -complexity). *Suppose that the semilinear boundary value problem (1) features a corner-weighted analytic source term $f \in L^2(\Omega) \cap B_{\beta}^0(\Omega)$ for which Assumption 1 can be satisfied. Consider the hp -ILG discretization based on a sequence $\{\mathcal{T}_p\}_{p \geq 1}$ of regular partitions of Ω into triangles, with geometric corner refinement, cf. (29), and on Galerkin projections onto the associated discrete spaces $\mathbb{W}_{N(p)} = \mathbb{S}_{\mathcal{D}}^p(\Omega; \mathcal{T}_p)$, $p \geq 1$, from (30), with $N = N(p) = \dim(\mathbb{S}_{\mathcal{D}}^p(\Omega; \mathcal{T}_p)) = \mathcal{O}(p^3)$, and with $n(p) = \mathcal{O}(p)$ steps of the linearized iteration (15) on $\mathbb{W}_{N(p)}$. Then, the following hold true:*

- (a) *For every $0 < \varepsilon \leq 1$, there exists a polynomial degree $p = \mathcal{O}(|\log(\varepsilon)|)$ such that upon applying $n(p)$ steps of the fully discrete hp -ILG iteration procedure (36) the approximate Galerkin solutions $\mathcal{U}_{n(p)} \equiv \mathcal{U}_n \in \mathbb{W}_{N(p)}$ satisfy the error bound $\|\nabla(u - \mathcal{U}_{n(p)})\|_{L^2(\Omega)} \leq \varepsilon$.*
- (b) *The total computational cost $\text{work}(\varepsilon) \in \mathbb{N}$, measured in terms of float point operations necessary to compute the hp -ILG approximations $\mathcal{U}_{n(p)} \in \mathbb{W}_{N(p)}$ to accuracy $\varepsilon > 0$, is bounded by*

$$\text{work}(\varepsilon) \leq C(1 + |\log(\varepsilon)|)^7. \quad (37)$$

- (c) *In terms of the number $N(p)$ of unknowns, or in terms of $\text{work}(\varepsilon)$ required for the computation of the hp -ILG discretization of (1), there are constants $b, C > 0$ such that it holds*

$$\|\nabla(u - \mathcal{U}_{n(p)})\|_{L^2(\Omega)} \leq C \exp(-b(N(p))^{1/3}) \leq C \exp(-b \text{work}(\varepsilon)^{1/7}),$$

for $p = 1, 2, \dots$

Proof. In accordance with (25) and Lemma 2, the error $\|\nabla(u - \mathcal{U}_{n(p)})\|_{L^2(\Omega)}$ for the fully discrete hp -ILG scheme (36) after employing $n(p) = \mathcal{O}(p)$ iteration steps is of order $C \exp(-bp)$ for suitable constants $b, C > 0$. In particular, the hp -ILG iteration error is of the order of the hp -discretization error, cf. (31). Furthermore, coupling the parameter p (i.e., the polynomial degree, viz. the number of geometric mesh layers) to the prescribed target accuracy $0 < \varepsilon \leq 1$ as in the statement of the theorem via $p \geq C|\log(\varepsilon)|$ for a sufficiently large uniform constant $C > 0$, shows (a). The work estimate in (b) follows from Proposition 6 and from the discussion in §4.2.1 on the complexity for one evaluation of the hp -discretization of the nonlinearity. Finally, (c) is a direct consequence of (a) and (b). \square

Remark 8. We remark that the error vs. work bound (37) is analogous in asymptotic order to the (exponential) error vs. work bound for the hp -FE Galerkin solution of the corresponding linear problem in Ω .

5. NUMERICAL EXPERIMENTS

In this section we present a series of numerical experiments to computationally verify the theoretical error vs. work bounds derived in Theorem 3 for the hp -ILG discretization of the semilinear boundary value problem (1). To this end, we consider the semilinear boundary value problem (1), for $\lambda = q = 1$ and the constant function $f \equiv 1$, on two different domains:

- **Example 1:** We let Ω be the unit square $(0, 1)^2$.
- **Example 2:** Here, Ω is the “L-shaped domain”, $(-1, 1)^2 \setminus [0, 1] \times (-1, 0] \subset \mathbb{R}^2$.

We remark that in both settings, the analytical solution to (1) is unknown and hence a suitably fine hp -mesh approximation is computed for the purposes of evaluating the corresponding $L^2(\Omega)$ -norm of the (gradient of the) error in the hp -ILG approximation defined by (36). Given that $f = 1$ throughout this section, the integral involving f can be computed exactly, and hence (15) and (36) are equivalent. The sequence of FE spaces $\mathbb{W}_{N(p)} := \mathbb{S}_{\mathcal{D}}^p(\Omega; \mathcal{T}_p)$, $p \geq 1$, are constructed as outlined in §4.1; namely, given an initial uniform triangular mesh \mathcal{T}_1 , the space $\mathbb{S}_{\mathcal{D}}^1(\Omega; \mathcal{T}_1)$ is constructed based on employing (uniform order) continuous piecewise linear polynomials. Subsequently we undertake geometric refinement of each mesh \mathcal{T}_p , $p \geq 1$, whereby elements in the vicinity of the four, respectively, six corners of the domain Ω are subdivided, while simultaneously (uniformly) increasing the polynomial order p . For the construction of the underlying hp -FE polynomial

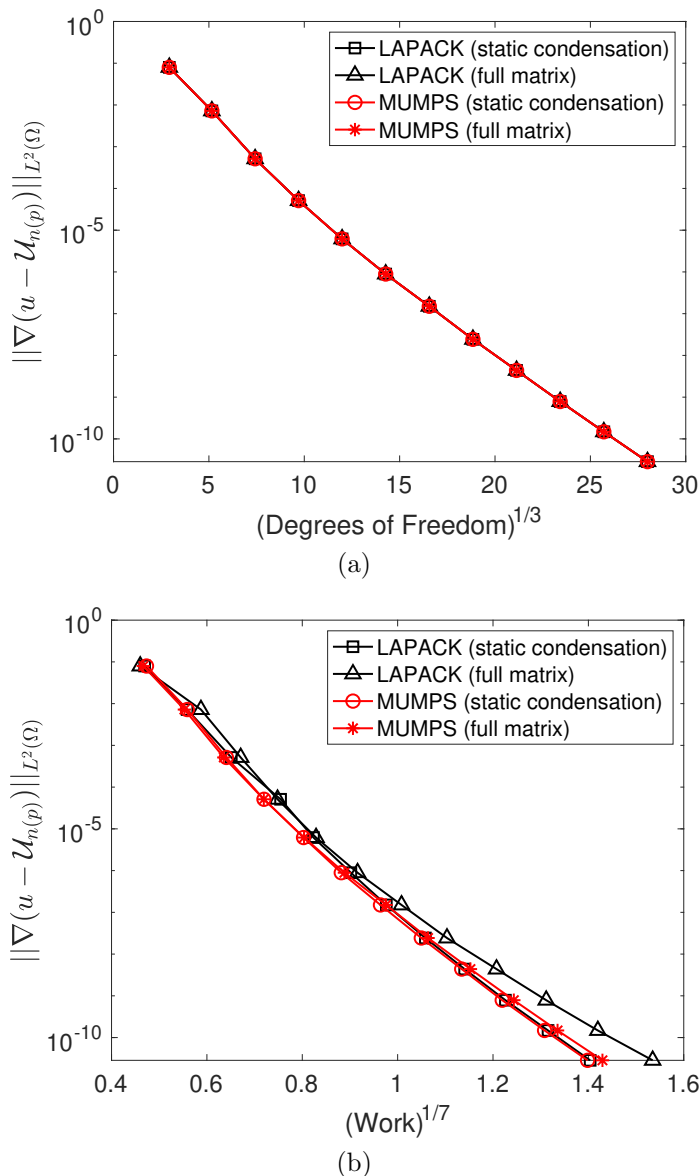


FIGURE 1. Example 1 (square domain): (a) Comparison of the error with respect to the third root of the number of degrees of freedom; (b) Comparison of the error with respect to the seventh root of the work (CPU time) measured in seconds.

spaces, we exploit the hierarchical basis defined in (22). Additionally, the sequences of corner-refined, geometric meshes used in the numerical experiments are built based on employing the standard red-green refinement strategy whereby (temporary) green refinement is undertaken to remove hanging nodes in the underlying computational mesh. The resulting sequence of meshes are generally not nested; however, we observe in the ensuing numerical experiments that nestedness of partitions is not necessary to ensure that the theoretical bounds in Theorem 3 still hold.

For the purposes of comparison, we consider the implementation of two different linear solvers for the factorization of the underlying matrix arising from the bilinear form $a(\cdot, \cdot)$ from (3). In addition to the Cholesky factorization DPOTRF implemented in LAPACK for dense matrices, cf. (5), we also consider the application of the MUltifrontal Massively Parallel Solver (MUMPS), see (1; 2; 3), which utilizes sparse (symmetric) storage of the underlying matrix problem. In

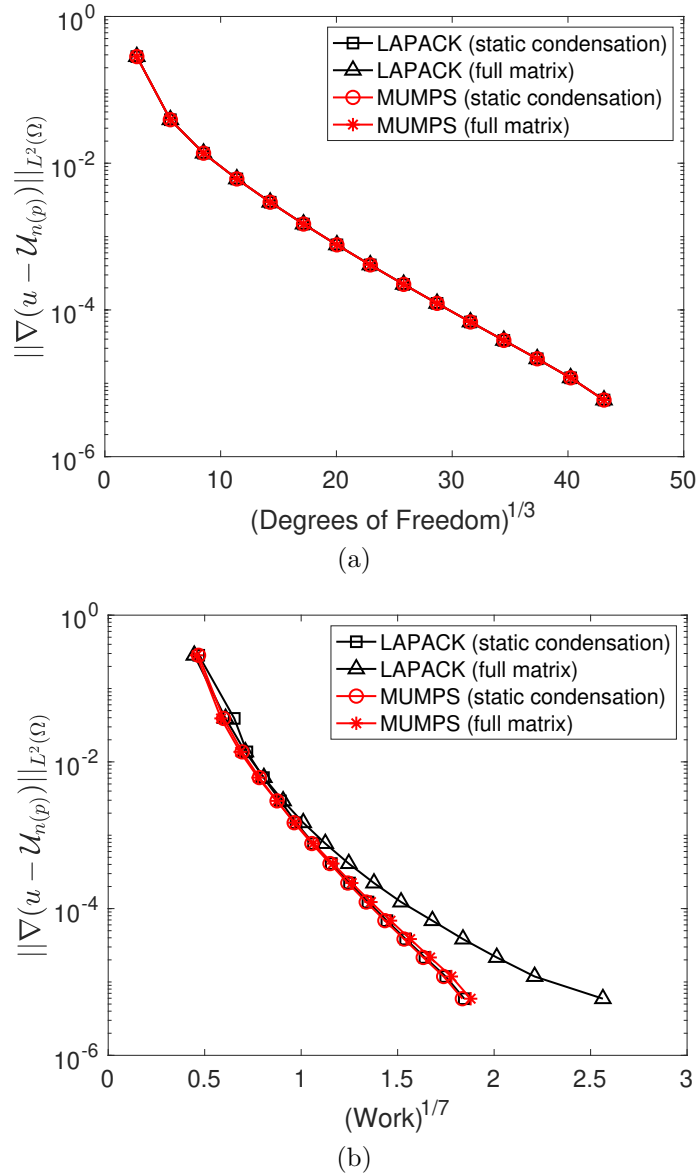


FIGURE 2. Example 2 (L-shaped domain): (a) Comparison of the error with respect to the third root of the number of degrees of freedom; (b) Comparison of the error with respect to the seventh root of the work (CPU time) measured in seconds.

Figures 1 & 2 we present a comparison of the norm of the error in the computed numerical solution $\mathcal{U}_{n(p)}$ for $p = 1, 2, \dots, 12$, respectively, $p = 1, 2, \dots, 15$, based on employing a starting mesh \mathcal{T}_1 comprising of 32, respectively 24, uniform triangular elements. Once the hp -ILG solution $\mathcal{U}_{n(p)} \in \mathbb{S}_{\mathcal{D}}^p(\Omega; \mathcal{T}_p)$ has been computed on a given mesh \mathcal{T}_p , for a particular polynomial degree p , this solution is then projected onto the refined finite element space $\mathbb{S}_{\mathcal{D}}^{p+1}(\Omega; \mathcal{T}_{p+1})$ to serve as the initial guess in the Picard fixed point scheme (36); for $p = 1$ the initial guess is the zero solution. The iteration defined in (36) is terminated once the ℓ_2 -norm of the difference in the coefficient vector between two subsequent computed solutions has been reduced by a factor of 10^{-2} , relative to the corresponding quantity computed between the initial guess and the first iterate computed using (36); furthermore we set $\alpha = 1/2$. Here we consider the impact of static condensation, as per

§4.2.3, as well as when both MUMPS and LAPACK are applied to the full matrix problem arising in (36). In Figures 1(a) & 2(a) we plot the $L^2(\Omega)$ -norm of the gradient of the error $u - \mathcal{U}_{n(p)}$ against the third root of the number of degrees of freedom in the FE space $\mathbb{W}_{N(p)} \equiv \mathbb{S}_{\mathcal{D}}^p(\Omega; \mathcal{T}_p)$ on a semi-log plot; here we observe that for both examples, as p is increased (and as the mesh is concurrently geometrically refined towards the corners of Ω), the convergence lines become straight, thereby indicating exponential convergence, cf. Theorem 3. Of course, all four numerical approaches produce identical results, as we would expect. Figures 1(b) & 2(b) present analogous results for both examples where we now plot $\|\nabla(u - \mathcal{U}_{n(p)})\|_{L^2(\Omega)}$ against the seventh root of the work, measured in CPU seconds, required to compute the hp -ILG solution $\mathcal{U}_{n(p)}$ on each FE space. As expected, when static condensation is employed both MUMPS and LAPACK lead to exponential rates of convergence, which is in agreement with Theorem 3. Furthermore, when LAPACK is exploited *without* static condensation, then we clearly observe a degeneration in the performance of the proposed approach. In contrast, we do not observe a similar degradation in the performance of the MUMPS solver applied to the full matrix (i.e., the assembled global stiffness matrix, without static condensation); we postulate that this is due to the fact that MUMPS can automatically exploit the structure of the matrix within the initial analyse phase, which is performed prior to computation of the factorisation. This ensures that exponential rates of convergence, *in terms of the work required to compute the hp -ILG solution*, are retained in this setting. Finally, we observe that the MUMPS solver, which stores the matrix in sparse format, is typically slightly more efficient than LAPACK which utilizes dense matrix storage.

6. CONCLUSIONS

For a model semilinear elliptic boundary value problem in a polygonal domain $\Omega \subset \mathbb{R}^2$, with a monotone, polynomial nonlinearity and analytic in $\bar{\Omega}$ source term f in (1), we proved that the underlying solution $u \in H_{\mathcal{D}}^1(\Omega)$ can be approximated numerically to accuracy $\varepsilon > 0$ with no more than $\mathcal{O}((1 + |\log(\varepsilon)|)^7)$ many operations. This is, up to the constant hidden in $\mathcal{O}(\cdot)$, the same asymptotic complexity than resulting for the corresponding linear elliptic problem, if a direct solver is used for the associated (linear) Galerkin system. Key in these results is analytic regularity of the weak solution (16) in scales of corner-weighted Sobolev spaces. The generalization to diffusion matrices with analytic in $\bar{\Omega}$ coefficients is possible at the expense of a quadrature error analysis in the (linear) principal part of (1) by means of an auxiliary result of Strang type. By similar arguments the presence of nonlinear reactions with variable analytic coefficients may be realized as well (not considered here). It could also be of interest to combine the present results with hp -adaptive refinement, similar to (7), where fixed order, Langrangian FE discretizations were studied.

REFERENCES

1. P. R. Amestoy, I. S. Duff, J. Koster, and J.-Y. L'Excellent, *A fully asynchronous multifrontal solver using distributed dynamic scheduling*, SIAM Journal on Matrix Analysis and Applications **23** (2001), no. 1, 15–41.
2. P. R. Amestoy, I. S. Duff, and J.-Y. L'Excellent, *Multifrontal parallel distributed symmetric and unsymmetric solvers*, Comput. Methods Appl. Mech. Eng. **184** (2000), 501–520.
3. P. R. Amestoy, A. Guermouche, J.-Y. L'Excellent, and S. Pralet, *Hybrid scheduling for the parallel solution of linear systems*, Parallel Computing **32** (2006), no. 2, 136–156.
4. M. Amrein and T. P. Wihler, *Fully adaptive Newton-Galerkin methods for semilinear elliptic partial differential equations*, SIAM J. Sci. Comput. **37** (2015), no. 4, A1637–A1657.
5. E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK users' guide*, third ed., Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999.
6. I. Babuška and B. Q. Guo, *Regularity of the solution of elliptic problems with piecewise analytic data. I. Boundary value problems for linear elliptic equation of second order*, SIAM J. Math. Anal. **19** (1988), no. 1, 172–203. MR 924554

7. 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), no. 4, 2193–2225. MR 4609880
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), no. 1, 329–347. MR 3620224
9. C. Bernardi, J. Dakroub, G. Mansour, and T. Sayah, *A posteriori analysis of iterative algorithms for a nonlinear problem*, J. Sci. Comput. **65** (2015), no. 2, 672–697. MR 3411283
10. S. Beuchler and J. Schöberl, *New shape functions for triangular p -FEM using integrated Jacobi polynomials*, Numer. Math. **103** (2006), no. 3, 339–366. MR 2221053
11. S. Congreve and T.P. Wihler, *Iterative Galerkin discretizations for strongly monotone problems*, J. Comput. Appl. Math. **311** (2017), 457–472. MR 3552717
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. Engrg. **200** (2011), no. 37–40, 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), no. 4, A1761–A1791.
14. M. Feischl and Christoph Schwab, *Exponential convergence in H^1 of hp -FEM for Gevrey regularity with isotropic singularities*, Numer. Math. **144** (2020), no. 2, 323–346.
15. F. Fuentes, B. Keith, L. Demkowicz, and S. Nagaraj, *Orientation embedded high order shape functions for the exact sequence elements of all shapes*, Computers & Mathematics with Applications **70** (2015), no. 4, 353–458.
16. Y. He and Ch. Schwab, *Analytic regularity and solution approximation for a semilinear elliptic partial differential equation in a polygon*, Calcolo **61** (2024), no. 1, Paper No. 11, 23. MR 4688254
17. 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), no. 2, 407–422. MR 4235817
18. P. Heid and T.P. Wihler, *Adaptive iterative linearization Galerkin methods for nonlinear problems*, Math. Comp. **89** (2020), no. 326, 2707–2734. MR 4136544
19. P. Houston and T. P. Wihler, *An hp -adaptive Newton-discontinuous-galerkin finite element approach for semilinear elliptic boundary value problems*, Math. Comp. **89** (2020), 2707–2734.
20. J. Nečas, *Introduction to the theory of nonlinear elliptic equations*, Teubner-Texte zur Mathematik [Teubner Texts in Mathematics], vol. 52, BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1983, With German, French and Russian summaries. MR 731261
21. Ch. Schwab, *p - and hp -finite element methods*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 1998, Theory and applications in solid and fluid mechanics. MR 1695813
22. P. Solin, K. Segeth, and I. Dolezel, *Higher-order finite element methods*, Studies in advanced mathematics, Chapman & Hall/CRC, Boca Raton, London, 2004.
23. B. Szabó and I. Babuška, *Finite element analysis*, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1991. MR 1164869
24. E. Zeidler, *Nonlinear functional analysis and its applications. II/B*, Springer-Verlag, New York, 1990.