VILNIUS
TECH
Vilnius Gediminas
Technical University

Numerical Validation of Probabilistic Laws to Evaluate Finite Element Error Estimates

Joël Chaskalovic[a] and Franck Assous[b]

[a] *D'Alembert, Sorbonne University*
75252 Paris, France
[b] *Department of Mathematics, Ariel University*
40700 Ariel, Israël
E-mail(*corresp.*): franckassous55@gmail.com
E-mail: jch1826@gmail.com

Abstract. We propose a numerical validation of a probabilistic approach applied to estimate the relative accuracy between two Lagrange finite elements P_k and P_m, ($k < m$). In particular, we show practical cases where finite element P_k gives more accurate results than finite element P_m. This illustrates the theoretical probabilistic framework we recently derived in order to evaluate the actual accuracy. This also highlights the importance of the extra caution required when comparing two numerical methods, since the classical results of error estimates concerns only the asymptotic convergence rate.

Keywords: numerical validation, error estimates, finite elements, Bramble-Hilbert lemma, probability.

AMS Subject Classification: 65N15; 65N30; 65N75; 65Gxx.

1 Introduction

Finite element methods and among them, error estimates play a significant role in the development of numerical methods. Very often, the success of a numerical method depends on its performance in terms of efficiency and accuracy. For this reason, it is still an active subject of research, as observed, for instance, with the considerable interest received by the discontinuous Galerkin methods in the past decades; see e.g. an introduction for elliptic problems in [3], the book [2] or the pioneering work [19]. Since the seminal papers of Strang and

Fix [26], Ciarlet and Raviart [14], Babuška [6] and Bramble and Hilbert [7], along with co-workers, a large amount of work has been published, the purpose of which was to derive and expand error estimates in different configurations. Here, we are concerned with *a priori* error estimates, that aim to find upper bounds for the error between the exact solution u and its finite element approximation u_h. More precisely, these estimates describe how the finite element error $\|u - u_h\|$, for a given norm, goes to 0 with mesh size h (i.e. the largest diameter of the elements in a given mesh). In addition, these estimates involve a constant, generally unknown, which leads to only get an upper bound for the approximation error. In addition, quantitative uncertainties do exist in finite element methods; these are based on the way the mesh grid generator creates the mesh which is used to compute the finite element approximation u_h, or since the equations are not exactly solved due to round-off errors. In previous papers [11,12], we investigated the error resulting from a partial non-control of the mesh size. For this purpose, we have considered the approximation error as a random variable, and we have evaluated the relative accuracy between two Lagrange finite elements with the help of a probabilistic approach. In the same way, one can find in [21, 22] a probabilistic approach to evaluate error bounds in numerical analysis. In this work, we *numerically* study the *a priori error* estimate due to the discretization of a linear variational problem by a finite element method, using standard polynomials. Our aim is to compare the probabilistic laws we derived with statistical results, when two different degrees of the polynomials are used, for a fixed value of the mesh size. Since the effective dependence of the accuracy on the mesh size is a central question, it could help one to understand the saturation assumption that is often used in *a posteriori* error analysis [1]. Indeed, we use here a probabilistic approach which differs from the methods involved in *a posteriori* error analysis. Nonetheless, we will show examples where P_k finite element is more likely accurate than P_m, $k < m$, which can be related to the invalidity of the saturation assumption [15]. The paper is structured as follows: Section 2 summarizes the results of [11, 12], which are necessary for one to understand the numerical experiments and their analysis. The main results are the geometrical interpretation of error estimates and the two probabilistic laws we deduced for finite element accuracy. Section 3 is devoted to the numerical results, which illustrate the new probabilistic way we propose to evaluate the accuracy between two finite elements. We basically consider two numerical problems, a stiff one and a smooth one, and we compare, for each of them, the behavior of the theoretical probabilistic models with the statistical results. Concluding remarks follow.

2 Probabilistic models and finite elements accuracy

2.1 Error estimates revisited

Consider Ω an open bounded and non-empty subset of $\mathbb{R}^n$, and let Γ denote its boundary, assumed to be C^1-piecewise. We also introduce a Hilbert space V endowed with a norm, $\|.\|_V$, and a bilinear, continuous and $V-$elliptic form $a(\cdot, \cdot)$ defined on $V \times V$. Finally, $l(\cdot)$ denotes a linear continuous form defined

on V. Let $u \in V$ be the unique solution to the second order elliptic variational formulation

$$\begin{cases} \text{Find } u \in V \text{ solution to:} \\ a(u, v) = l(v), \quad \forall v \in V. \end{cases} \tag{2.1}$$

In this paper, we will focus on the simple case where V is the usual Sobolev space $H^1(\Omega)$. More general cases can be found in [13]. Let us now introduce the finite-dimensional subspace V_h of V, and consider $u_h \in V_h$ an approximation of u, solution to the approximate variational formulation

$$\begin{cases} \text{Find } u_h \in V_h \text{ solution to:} \\ a(u_h, v_h) = l(v_h), \quad \forall v_h \in V_h. \end{cases}$$

In what follows, we are interested in evaluating error bounds for finite element methods. Hence, we first assume that domain Ω is exactly covered by a mesh $\mathcal{T}_h$ composed by N_s n-simplices $K_j, (1 \leq j \leq N_s)$, which respects classical rules of regular discretization, (see for example [8] for the bidimensional case, or [23] in $\mathbb{R}^n$). We also denote by $P_k(K_j)$ the space of polynomial functions defined on a given n-simplex K_j of degree less than or equal to k, $(k \geq 1)$. Our study relies on the results of [23]. Let $\|.\|_1$ be the classical norm in $H^1(\Omega)$ and $|.|_{k+1}$ the semi-norm in $H^{k+1}(\Omega)$, and let h be the mesh size, namely the largest diameter of the elements of the mesh $\mathcal{T}_h$. We thus have:

Lemma 1. *Suppose that there exists an integer $k \geq 1$ such that the approximation u_h of V_h is a continuous piecewise function composed by polynomials which belong to $P_k(K_j), (1 \leq j \leq N_s)$. Then, if the exact solution u belongs to $H^{k+1}(\Omega)$, we have the following error estimate:*

$$\|u_h - u\|_1 \leq \mathcal{C}_k h^k |u|_{k+1}, \tag{2.2}$$

where $\mathcal{C}_k$ is a positive constant independent of h.

Now, let us consider two families of Lagrange finite elements P_k and P_m for two values $(k, m) \in \mathbb{N}^{*2}, (k < m)$. Assuming that the solution u to (2.1) belongs to $H^{m+1}(\Omega)$, inequality (2.2) can be written as

$$\|u_h^{(k)} - u\|_1 \leq \mathcal{C}_k h^k |u|_{k+1}, \tag{2.3}$$

$$\|u_h^{(m)} - u\|_1 \leq \mathcal{C}_m h^m |u|_{m+1}, \tag{2.4}$$

where $u_h^{(k)}$ and $u_h^{(m)}$ respectively denote the P_k and P_m Lagrange finite element approximations of u. In this article, following a series of previous papers [12,13] where a theoretical analysis was performed, we are interested in numerical applications. To this end, for a given mesh size h, two independent meshes for P_k and P_m are built by a mesh generator. Usually, one considers inequalities (2.3) and (2.4) so as to conclude that, when h goes to zero, P_m is more accurate that P_k, since h^m goes faster to zero than h^k. However, in practical numerical applications, the size of the mesh is chosen according to the desired accuracy, so that h has a fixed value. Consequently, this way of comparison is no more relevant. For this reason, we mean to identify the relative accuracy between P_k and $P_m, (k < m)$, for a given value of h.

2.2 Two probabilistic laws

In [11, 12], we introduced a probabilistic approach that provides a coherent framework for modeling uncertainties in finite element approximations: such uncertainties may come from the way the meshes are created by computer algorithms, leading to a partial non-control of the mesh, even for a given maximum mesh size. In this framework, values $\|u_h^{(k)} - u\|_1$ and $\|u_h^{(m)} - u\|_1$ are viewed as two random variables, respectively denoted as $X^{(k)}(h)$ and $X^{(m)}(h)$, whose support is $\left[0, \mathcal{C}_i|u|_{i+1}h^i\right], (i = k$ or $i = m)$, according to inequalities (2.3) and (2.4). Our goal is thus to derive a probabilistic law for the event

$$\left\{X^{(m)}(h) \leq X^{(k)}(h)\right\} \equiv \left\{\|u_h^{(m)} - u\|_1 \leq \|u_h^{(k)} - u\|_1\right\},$$

which corresponds to the relative accuracy between finite elements P_k and P_m. For this purpose, we first introduce the random events A and B defined by:

$$A \equiv \left\{X^{(m)}(h) \leq X^{(k)}(h)\right\}, \quad B \equiv \left\{X^{(k)}(h) \in \left[\mathcal{C}_m|u|_{m+1}h^m, \mathcal{C}_k|u|_{k+1}h^k\right]\right\}.$$

Moreover, we proved in [12] the following result:

Lemma 1. *Let us assume that A and B are two independent events. Then, the probability law $P(A)$ of event A is given by:*

$$P(A) = \left|\begin{array}{ll} 1 & \text{if } 0 < h < h_{k,m}^*, \\ 0 & \text{if } h > h_{k,m}^*, \end{array}\right. \tag{2.5}$$

where $h_{k,m}^*$ is defined by:

$$h_{k,m}^* \equiv \left(\mathcal{C}_k|u|_{k+1}/\mathcal{C}_m|u|_{m+1}\right)^{\frac{1}{m-k}}. \tag{2.6}$$

The shape of the probabilistic distribution, called the two-steps model, is depicted in Figure 1.
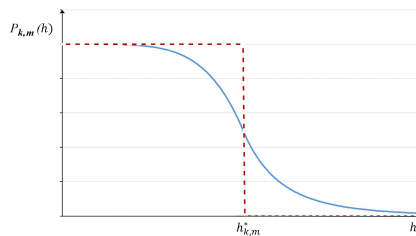

Figure 1. Case $m - k \neq 1$: shape of the sigmoid distribution (2.7) (full line) and the two steps corresponding one (2.5) (dashed line), $(P_{k,m}(h) \equiv Prob\{X^{(m)}(h) \leq X^{(k)}(h)\})$.

Basically, it expresses the fact that, for $h < h_{k,m}^*$, finite element P_m is *almost surely* more accurate than P_k, whereas for $h > h_{k,m}^*$, P_k becomes *almost surely* more accurate than P_m. To relax the independence assumption of events A and B, we also derived a second probabilistic law based on the uniform distribution of the random variable $X^{(k)}(h)$ over $\left[0, \mathcal{C}_k|u|_{k+1}h^k\right]$. In this context, we proved in [12] the following theorem:

Theorem 1. *Let us assume that $X^{(i)}(h), (i = k, m)$, are independent and uni-formly distributed on $[0, \mathcal{C}_i|u|_{i+1}h^i]$. Then, the probability $P(A)$ of event A is given by:*

$$P(A) = \left| \begin{array}{ll} 1 - 0.5 \left(h/h_{k,m}^*\right)^{m-k} & \text{if } 0 < h \leq h_{k,m}^*, \\ 0.5 \left(h_{k,m}^*/h\right)^{m-k} & \text{if } h \geq h_{k,m}^*. \end{array} \right. \tag{2.7}$$

The shape of this law, called the sigmoid model, is also plotted in Figure 1. As one can see, for $h > h_{k,m}^*$, $P(A) \leq 0.5$: in that case, finite element P_m is *probably* overqualified. The purpose of the next section is to propose numerical examples that illustrate and validate this probabilistic approach by compar-ing statistical frequencies and the corresponding probabilities determined by (2.5) or (2.7). Note that our method can be *a priori* applied for solution u smooth enough. This constraint comes from the use of classical estimates (see Lemma 1), where, for instance, the exact solution u must be at least in $H^3(\Omega)$ for comparing P_1 and P_2 finite elements. Nevertheless, from a heuristic point of view, it is usual to apply these theoretical results even for less regular solutions, where *a priori* regularity of u is not known.

3 Numerical results

In this section, we will illustrate our probabilistic approach on numerical ex-amples, by evaluating the relative accuracy of two Lagrange finite elements. We have intentionally chosen simple, standard examples, in order to help us to check numerically the relevance of the proposed probabilistic distributions. Note also that we will only consider examples where the exact solution is known. Hence, we will not deal with cases where, for instance, non-tensor product so-lutions can be constructed using a Fourier series. Indeed, such an approach would require to truncate the corresponding Fourier series, introducing in such a way an unwanted additional error, that would damage the numerical results we want to investigate.

Hence, we consider the following classical elliptic problem, with obvious notations: *Find $u \in H^1(\Omega)$ solution to*

$$\left\{ \begin{array}{rcll} -\Delta u & = & q & \text{in } \Omega, \\ u & = & h & \text{on } \partial\Omega, \end{array} \right. \tag{3.1}$$

where, for simplicity, domain Ω is the open unit square in $\mathbb{R}^2$: $\Omega =]0, 1[\times]0, 1[$. The associated variational formulation, which is analogous to (2.1), can be readily derived. According to the choice of q and h, we will consider as examples a *stiff* problem, where the solution exhibits rapid variations or, alternatively, a *very smooth* problem.

One of the main ingredients of the method is the computation of $h_{k,m}^*$, as defined by (2.6). As one will see, it will be evaluated using a *maximum likelihood estimator*; see for instance [24]. In our case, this principle is applied as follows: for a given finite element P_k, we consider a number N of different meshes with the same (maximum) mesh size h. Then, we compute:

$$\max_{N,h} \|u_h^{(k)} - u\|_1/h^k, \tag{3.2}$$

which constitutes, using estimate (2.3), the maximum likelihood estimator for $\mathcal{C}_k|u|_{k+1}$. Indeed, due to inequality (2.3), quantity $\frac{X^k(h)}{h^k}$ is also a uniform random variable whose support is $[0, \mathcal{C}_k|u|_{k+1}]$. Then, one can show [18] that for a given uniform random variable Y whose support is $[0, \theta]$, θ being an unknown real parameter, the maximum likelihood estimator $\hat{\theta}$ is given by:

$$\hat{\theta} = \max(Y_1, \dots, Y_N),$$

where $(Y_1, \dots, Y_N)$ is a sample built with independent and identically distributed random variables $(Y_i)_{i=1,N}$, with the same distribution as Y. In our case, this implies that (3.2) is the maximum likelihood estimator for $\mathcal{C}k|u|_{k+1}$, since N and h each take a finite number of values. Doing the same for another finite element P_m, we obtain that the estimator for $h_{k,m}^*$, denoted $\widehat{h_{k,m}^*}$, is defined by:

$$\widehat{h_{k,m}^*} = \left(\max_{N,h} \frac{\|u_h^{(k)} - u\|_1}{h^k} \Big/ \max_{N,h} \frac{\|u_h^{(m)} - u\|_1}{h^m} \right)^{1/m-k}. \tag{3.3}$$

Then, one can easily compute the two probability laws introduced in Subsection 2.2. Indeed, as soon as $\widehat{h_{k,m}^*}$ is computed, functions (2.5) and (2.7) are operational by replacing $h_{k,m}^*$ by $\widehat{h_{k,m}^*}$. All the numerical results below are computed in this way. In order to numerically check the validity of each model, we now compare the two probabilistic laws defined by (2.5) and (2.7) with the corresponding statistical frequencies computed on the N meshes, for each fixed value h of the mesh size. To that end, we consider for two finite elements P_k and P_m ($k < m$), the same number N of different meshes with the same (maximum) mesh size h. From there, we compute the approximate solution $u_h^{(m)}$ and $u_h^{(k)}$, and we test if $\|u_h^{(m)} - u\|_1 \leq \|u_h^{(k)} - u\|_1$. Then, we repeat the same process for different values of h, either lower or greater than $\widehat{h_{k,m}^*}$. This gives, as a function of h, the percentage of cases where the approximation error of P_m is lower than the approximation error of P_k. In all cases, we use package FreeFem++ [17] to compute the P_k and P_m finite element approximations. Basically, the mesh generator of FreeFem++ is based on the Delaunay-Voronoi algorithm, and allows us to fix the maximum mesh size h for N different meshes of the same domain Ω. Since FreeFEM++ is built using updated and modern methods of meshing, we are confident that using another mesh generator will lead to equivalent results. In the next subsection, we consider a stiff case, whereas in the following one, we deal with a very smooth example.

3.1 A first stiff case

To introduce such a stiff case, we consider the well-known Runge function $\varphi(t) = 1/(1 + \alpha t^2)$, which takes α as a parameter, the classical Runge function corresponding to $\alpha = 25$ (see [16, 25]). Since we first aim at building an exact solution $u(x, y)$ for (3.1), we consider solutions of the form $u(x, y) = f(x)g(y)$, where both $f(x)$ and $g(y)$ are Runge functions of parameter α. To compute the derivatives of $u(x, y)$, we basically need the derivatives of the Runge function

φ. After some elementary algebra, we obtain the derivatives of $\varphi(t)$ (namely $\varphi'(t) = -2\alpha t/(1+\alpha t^2)^2, \varphi''(t) = 2\alpha(3\alpha t^2 - 1)/(1+\alpha t^2)^3)$, from which the Laplacian of $u(x,y)$ can easily be derived. Indeed, by computing the second order partial derivatives u_{xx}, u_{yy}, we find that

$$- \Delta u = -(f''(x)g(y) + f(x)g''(y))$$
$$= -\frac{2\alpha(3\alpha x^2 - 1)}{(1+\alpha x^2)^3} \frac{1}{1+\alpha y^2} - \frac{1}{1+\alpha x^2} \frac{2\alpha(3\alpha y^2 - 1)}{(1+\alpha y^2)^3}. \qquad (3.4)$$

We now set the right-hand side $q(x,y)$ of (3.1) equal to expression (3.4) above, so that

$$u(x,y) = \left(\frac{1}{1+\alpha x^2}\right)\left(\frac{1}{1+\alpha y^2}\right) \qquad (3.5)$$

is the exact solution of (3.1), provided that the Dirichlet boundary condition $h(x,y)$ is taken as the trace of $u(x,y)$ on the boundary $\partial\Omega$, that is

$$\begin{cases} h(x,0) = \dfrac{1}{1+\alpha x^2}, & h(0,y) = \dfrac{1}{1+\alpha y^2}, \\ h(x,1) = \dfrac{1}{1+\alpha x^2}\dfrac{1}{1+\alpha}, & h(1,y) = \dfrac{1}{1+\alpha}\dfrac{1}{1+\alpha y^2}. \end{cases}$$

In what follows, we analyze the relative accuracy between two Lagrange finite elements, $u(x,y)$, as defined in (3.5), being the reference solution for comparison.

3.1.1 $P_2 - P_3$ comparison and α-independence

The first numerical test we present is devoted to a comparison between finite elements P_2 and P_3. We first choose $\alpha = 500$. In that case, as explained above, we computed value $\widehat{h^*_{2,3}}$ as defined in (3.3) and obtained $\widehat{h^*_{2,3}} \simeq 0.12$. For this example, we have used values of h varying from 0.05 to 0.18, and for each h we have constructed $N = 500$ different meshes with the same value of h. In Figure 2 we plot, on the same picture, the results obtained for the statistical frequencies (full line) and for the two-steps probability law (2.5) (dotted line), as a function of h.
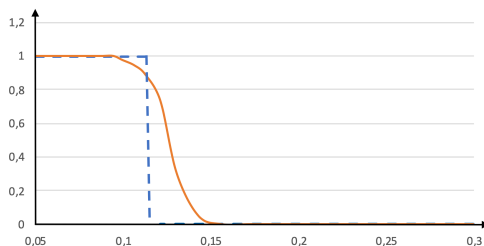

Figure 2. P_2 versus P_3 for the Runge function with $\alpha = 500$. Comparison between the statistical frequencies (full line) and the probabilistic law (2.5) (dotted line). 500 meshes are used for each value of h.

We then checked that the results do not depend on the value of α. For this purpose, we repeated the same numerical experiments for $\alpha = 25$ and $\alpha = 2000$. The results, depicted in Figure 3, show the same behavior as previously. Of course, the value of $\widehat{h^*_{k,m}}$ does depend on α and we have $\widehat{h^*_{2,3}} \simeq 0,13$ for $\alpha = 25$ and $\widehat{h^*_{2,3}} \simeq 0.07$ for $\alpha = 2000$.
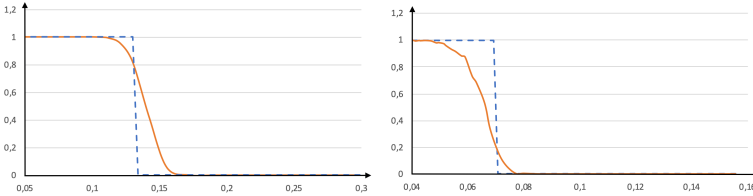

Figure 3. P_2 versus P_3 for the Runge function with $\alpha = 25$ (left) and $\alpha = 2000$ (right). Statistical frequencies (full) and probabilistic law (2.5) (dotted). 500 meshes are used for each value of h.

3.1.2 Comparison with $\mathbf{P_4}$ finite element

Next, we numerically assessed the validity of the present approach when finite element P_4 is involved. We first compared P_3 with P_4, then P_2 with P_4. To this end, we used the Runge function with $\alpha = 2000$, and we also considered $N = 500$ different meshes. The computed value of $\widehat{h^*_{k,m}}$ we obtained are $\widehat{h^*_{3,4}} \simeq 0.24$ for P_3–P_4 and $\widehat{h^*_{2,4}} \simeq 0.094$ for P_2–P_4. The results are depicted in Figure 4 and show, like previously, that the statistical frequencies behave very similarly to the two-steps probabilistic law (2.5).
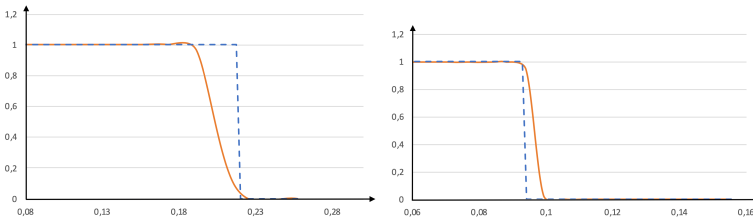

Figure 4. Comparisons P_3–P_4 (left) and P_2–P_4 (right) for the Runge function with $\alpha = 2000$. Statistical frequencies (full) and probabilistic law (2.5) (dotted). 500 meshes are used for each value of h.

The last illustration of this subsection is devoted to the comparison between the statistical frequencies and the sigmoid probabilistic law defined in (2.7). We considered comparisons between finite elements P_2 and P_4, then between P_1 and P_4. We followed the same procedure as above, again with the same parameters ($N = 500$ and $\alpha = 2000$). The results are depicted in Figure 5.

As one can see, there is a weaker fitting between the two curves than with the two-steps law (2.5), even if the trend is still correct. Remark also that the fit is better in the P_1–P_4 case than in the P_2–P_4 case. More generally, the
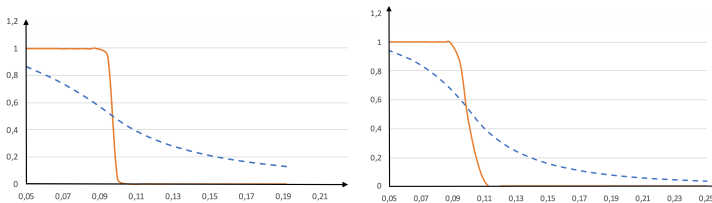
Figure 5. P_2 versus P_4 (left) and P_1 versus P_4 (right) for the Runge function with $\alpha = 2000$. Comparison between the statistical frequencies (full) and the probabilistic law (2.7) (dotted). 500 meshes are used for each value of h.

greater the $m - k$ difference, the better the match. Hence, the sigmoid model also gave a correct trend, but was less precise and satisfying than the two-steps law, in particular when $m - k = 1$, for instance when one compares P_2 with P_3, see Figure 6.
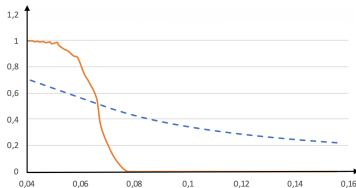

Figure 6. P_2 versus P_3 for the Runge function with $\alpha = 2000$. Comparison between the statistical frequencies (full) and the probabilistic law (2.7) (dotted). 500 meshes are used for each value of h.

Indeed, in that case, the first part of (2.7) is a linear decreasing function of h (for any given fixed $h^*_{k,m}$), and the second one decreases like $1/h$. However, if difference $m - k = 2$, for instance when comparing P_2 to P_4, the fit is better. Indeed, the first part of (2.7) is a decreasing function $\simeq -h^2$ (for any given fixed $h^*_{k,m}$), and the second one decreases like $1/h^2$. So, depending on the difference $m - k$, the sigmoid law remains to some extent relevant, where high order finite element (with m around 20–25) are sometimes used [20].

This shows that the two-steps model works well, but is a bit "rough" (essentially binary), whereas the sigmoid law is probably too "rigid" and has to be make more "flexible" to obtain a better fit with the statistical results. For this reason, we are currently working on a more general approach which corresponds to a new generation of probabilistic laws that better fit the statistical frequencies.

3.2 A smooth example

In this subsection, we illustrate the probabilistic laws for a very smooth solution to the variational problem (2.1). To build such a case, we chose $q = 2\pi^2 \sin(\pi x)\cos(\pi y)$ in (3.1), so that $u(x,y) = \sin(\pi x)\cos(\pi y)$ is the exact solution of the problem, provided that the Dirichlet boundary condition h is taken

as the trace of $u(x, y)$ on the boundary $\partial\Omega$. This can be written as:

$$\begin{cases} h(x, 0) = \sin(\pi x), & h(0, y) = 0, \\ h(x, 1) = -\sin(\pi x), & h(1, y) = 0. \end{cases}$$

As previously, we first compute $\widehat{h^*_{k,m}}$ defined by (3.3), then we compute the probabilistic models introduced above. After that, we compare these results to the statistical frequencies. For example, we consider the finite elements P_2 and P_3: we depicted in Figure 7 (left) the statistical frequencies and the probabilistic law (2.5). As in all the other numerical experiments, 500 meshes have been used for each value of h, where we found a value of $\widehat{h^*_{2,3}}$ approximately equal to 0.18. As one can see, even in this case, there is a good agreement between the statistical frequencies and the probabilistic law (2.5). However, the comparison with the sigmoid law (2.7) (right part of Figure 7) gave only a global trend and was not really accurate. Here again, as for the Runge example, it will be improved by the above-mentioned new generation of probabilistic laws.
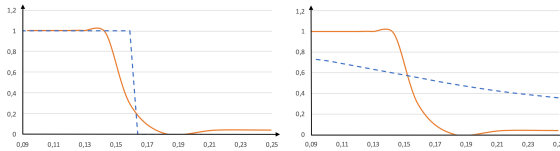

Figure 7. P_2 versus P_3 for the smooth case. Comparison between the statistical frequencies (full) and the two probabilistic laws (dotted). 500 meshes are used for each value of h. Left: comparison with the two-steps probabilistic law (2.5) – Right: comparison with the sigmoid probabilistic law (2.7).

4 Conclusions

In this paper, we proposed to apply the probabilistic approach we developed in [12] to numerical examples. It enabled us to evaluate the relative accuracy between two Lagrange finite elements P_k and P_m, $(k < m)$, for a fixed value of the mesh size h. Our approach, which is based on a geometrical interpretation of the error estimate, considers the approximation errors as random variables. Two probabilistic laws were derived, a so-called "two steps" law and a "sigmoid" one, depending on the probabilistic assumptions which were made on the corresponding random variables. For the finite elements we considered, we illustrated, using several examples, the property that, depending on the position of h with respect to the critical value $h^*_{k,m}$, we can actually estimate which of finite elements P_k and P_m is more likely accurate. This overturns the common misconception that finite elements P_m are always more precise than P_k if $m > k$, regardless of the mesh size h. In particular, this shows cases where a P_m finite element *surely* is overqualified. As a consequence, a significant reduction of implementation time and execution cost can be obtained without loss of accuracy. Such a phenomenon was already observed by using data-mining

techniques (see [4, 5, 9] and [10]). However if, in the proposed examples, the first investigated law (the two-steps law) fit the numerical results satisfactorily, the second proposed law (the sigmoid one) produces only a trend and is not accurate enough. Indeed, the results show that the statistical frequencies behave similarly to the two-steps probabilistic law, in both the smooth and stiff examples. However, there is a weaker fitting between the statistical error and the sigmoid law, particularly when difference $m - k$ is small. To address this issue, we are currently working on a new probabilistic framework which corrects the gap between the statistics and a "generalized" probability law. Finally, note that this approach is not limited to finite element methods, and can be generalized to other approximation methods: given several different numerical methods and their error estimates, it would be possible to order them by evaluating which is the most probably accurate.

References

[1] M. Ainsworth, J.T. Oden and A Posteriori. *Error Estimation in Finite Element Analysis.* John Wiley & Sons, 2000. https://doi.org/10.1002/9781118032824.

[2] J.S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications.* exts in Applied Mathematics. Springer, New York, 2008. https://doi.org/10.1007/978-0-387-72067-8.

[3] D.N. Arnold, F. Brezzi, B. Cockburn and L.D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM Journal on Numerical Analysis*, **39**(5):1749–1779, 2002. https://doi.org/10.1137/S0036142901384162.

[4] F. Assous and J. Chaskalovic. Data mining techniques for scientific computing: Application to asymptotic paraxial approximations to model ultrarelativistic particles. *Journal of Computational Physics*, **230**(12):4811–4827, 2011. https://doi.org/10.1016/j.jcp.2011.03.005.

[5] F. Assous and J. Chaskalovic. Error estimate evaluation in numerical approximations of partial differential equations: A pilot study using data mining methods. *Comptes Rendus Mécanique*, **341**(3):304–313, 2013. https://doi.org/10.1016/j.crme.2013.01.002.

[6] I. Babuška. Error-bounds for finite element method. *Numerische Mathematik*, **16**(4):322–333, 1971. https://doi.org/10.1007/BF02165003.

[7] J.H. Bramble and S.R. Hilbert. Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation. *SIAM Journal on Numerical Analysis*, **7**(1):112–124, 1970. https://doi.org/10.1137/0707006.

[8] J. Chaskalovic. *Mathematical and numerical methods for partial differential equations.* Springer Verlag, 2013. https://doi.org/10.1007/978-3-319-03563-5.

[9] J. Chaskalovic and F. Assous. Data mining and probabilistic models for error estimate analysis of finite element method. *Mathematics and Computers in Simulation*, **129**:50–68, 2016. https://doi.org/10.1016/j.matcom.2016.03.013.

[10] J. Chaskalovic and F. Assous. Probabilistic approach to characterize quantitative uncertainty in numerical approximations. *Mathematical Modelling and Analysis*, **22**(1):106–120, 2017. https://doi.org/10.1080/00036811.2019.1698727.

[11] J. Chaskalovic and F. Assous. A new mixed functional-probabilistic approach for finite element accuracy. *Computational Methods in Applied Mathematics*, **20**(4):799–813, 2020. https://doi.org/10.1515/cmam-2019-0089.

[12] J. Chaskalovic and F. Assous. A new probabilistic interpretation of the Bramble-Hilbert lemma. *Computational Methods in Applied Mathematics*, **20**(1):79–87, 2020. https://doi.org/10.1515/cmam-2018-0270.

[13] J. Chaskalovic and F. Assous. Explicit k-dependence for p_k finite elements in $w^{m,p}$ error estimates: application to probabilistic laws for accuracy analysis. *Applicable Analysis*, **100**(13):2825–2843, 2021. https://doi.org/10.1080/00036811.2019.1698727.

[14] P.G. Ciarlet and P.A. Raviart. General Lagrange and Hermite interpolation in r^n with applications to finite element methods. *Archive for Rational Mechanics and Analysis*, **46**(3):177–199, 1972. https://doi.org/10.1007/BF00252458.

[15] W. Dörfler and R.H. Nochetto. Small data oscillation implies the saturation assumption. *Numerische Mathematik*, **91**(1):1–12, 2002. https://doi.org/10.1007/s002110100321.

[16] J.F. Epperson. On the Runge example. *The American Mathematical Monthly*, **94**(4):329–341, 1987. https://doi.org/10.1080/00029890.1987.12000642.

[17] F. Hecht. New development in freefem++. *Journal of Numerical Mathematics*, **20**(3–4):251–266, 2012. https://doi.org/10.1515/jnum-2012-0013.

[18] M. Le Jeune. *Théorie de l'estimation paramétrique ponctuelle*. Springer, Paris, 2010. https://doi.org/10.1007/978-2-8178-0157-5_6.

[19] P. Lasaint and P.A. Raviart. On a finite element method for solving the neutron transport equation. In Carl de Boor(Ed.), *Mathematical Aspects of Finite Elements in Partial Differential Equations*, pp. 89–123. Academic Press, 1974. https://doi.org/10.1016/B978-0-12-208350-1.50008-X.

[20] W.F. Mitchell. How high a degree is high enough for high order finite elements? *Procedia Computer Science*, **51**:246–255, 2015. https://doi.org/10.1016/j.procs.2015.05.235.

[21] E. Novak. *Deterministic and Stochastic Error Bounds in Numerical Analysis*. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1988. https://doi.org/10.1007/BFb0079792.

[22] E. Novak. *Stochastic error bounds for some nonlinear problems in numerical analysis*, pp. 503–506. Academic Press, 01 1989.

[23] P.A. Raviart and J.M. Thomas. *Introduction á l'analyse numérique des équations aux dérivées partielles*. Lecture Notes in Mathematics. Masson, 1982. https://doi.org/10.1007/BFb0079792.

[24] R.J. Rossi. *Mathematical Statistics: An introduction to likelihood based Inference*. John Wiley & Sons, New York, 2018. https://doi.org/0.1002/9781118771075.

[25] C. Runge. Über empirische funktionen und die interpolation zwischen äquidistanten ordinaten. *Zeitschrift für Mathematik und Physik*, **46**, 1901.

[26] G. Strang and G.J. Fix. *An Analysis of the Finite Element Method*. Automatic Computation. Prentice-Hall, 1973.