

# On generalized binomial laws to evaluate finite element accuracy: toward applications for adaptive mesh refinement

Joël Chaskalovic \*      Franck Assous †

## Abstract

The aim of this paper is to provide new perspectives on the relative finite elements accuracy. Starting from a geometrical interpretation of the error estimate which can be deduced from Bramble-Hilbert lemma, we derive a probability law that evaluates the relative accuracy, considered as a random variable, between two finite elements  $P_k$  and  $P_m$ , ( $k < m$ ). We extend this probability law to get a cumulated probabilistic law for two main applications. The first one concerns a family of meshes and the second one is dedicated to a sequence of simplexes which constitute a given mesh. Both of this applications might be relevant for adaptive mesh refinement.

*keywords:* Error estimates, Probability, Finite elements, Bramble-Hilbert lemma, Mesh refinement.

## 1 Introduction

The past decades have seen the development of finite element error estimates due to their influence on improving both accuracy and reliability in scientific computing.

In the *a priori* error estimates, an unknown (in most of the cases) constant is involved which depends, among others, on the basis functions of the considered finite element and on a given semi-norm of the exact solution one wants to approximate. Moreover, error estimates are only upper bounds of the approximation error yielding that its exact value is unknown.

This was the starting point which motivated us to consider the approximation error as a random variable to therefore derive a probability law of the relative accuracy between two Lagrange finite elements  $P_k$  and  $P_m$ , ( $k < m$ ), ([3] and [4]).

Here, our aim is to extend the results obtained in [3] for a given mesh  $\mathcal{T}_h$  defined by a fixed mesh size  $h$  (corresponding to the largest diameter of all the simplexes of  $\mathcal{T}_h$ ). We will generalize our purpose either to a family of meshes or, for a given mesh, to the sequence of the local elements that constitute it.

As we will see, to proceed it, we will need to distinguish between the local and the global accuracy of two given Lagrange finite elements  $P_k$  and  $P_m$ .

Both of these points of view will be introduced as potential applications for adaptive mesh refinement.

The paper is organized as follows. We recall in Section 2 the mathematical problem we consider, the Bramble-Hilbert lemma and the resulting error estimate that allowed us to introduce

---

\*D'Alembert, Sorbonne University, Paris, France, (*email:* jch1826@gmail.com)

†Department of Mathematics, Ariel University, 40700 Ariel, Israël, (*email:* franckassous55@gmail.com).

a probability law for the relative global accuracy between  $P_k$  and  $P_m$  finite elements. Extension to a family of meshes is addressed in Section 3. In Section 4, one proposes a generalization that describes the relative local accuracy between two finite elements in a sequence of simplexes belonging to a given mesh. Then, the cumulated probabilistic law is derived either for a family of meshes or for a sequence of simplexes. Concluding remarks follow.

## 2 Comparing global finite elements accuracy by a probabilistic approach

Let  $\Omega$  be an open bounded and non empty subset of  $\mathbb{R}^n$  and  $\Gamma$  its boundary which we assumed to be  $C^1$ -piecewise. Let also  $u$  be the solution to the second order elliptic variational formulation:

$$(\mathbf{VP}) \quad \begin{cases} \text{Find } u \in V \text{ solution to:} \\ a(u, v) = l(v), \quad \forall v \in V, \end{cases} \quad (1)$$

where  $V$  is a given Hilbert space endowed with the norm  $\|\cdot\|_V$ ,  $a(\cdot, \cdot)$  is a bilinear, continuous and  $V$ -elliptic form defined on  $V \times V$ , and  $l(\cdot)$  a linear continuous form defined on  $V$ . Classically, variational problem  $(\mathbf{VP})$  has one and only solution  $u \in V$  (see for example [1]). In the sequel and for simplicity, we will restrict ourselves to the case where  $V$  is a usual Sobolev space of distributions.

Let us also consider the approximation  $u_h$  of  $u$ , solution to the approximate variational formulation:

$$(\mathbf{VP})_h \quad \begin{cases} \text{Find } u_h \in V_h \text{ solution to:} \\ a(u_h, v_h) = l(v_h), \quad \forall v_h \in V_h, \end{cases} \quad (2)$$

where  $V_h$  is a given finite-dimensional subset of  $V$ .

To state a corollary of Bramble-Hilbert's lemma (cf. [5] or [6]), we assume that  $\Omega$  is exactly covered by a mesh  $\mathcal{T}_h$  composed by  $N_K$  simplexes  $K_\mu$ , ( $1 \leq \mu \leq N_K$ ), which respect classical rules of regular discretization, (see for example [1] for the bidimensional case and [6] in  $\mathbb{R}^n$ ). Moreover, we denote by  $P_k(K_\mu)$  the space of polynomial functions defined on a given simplex  $K_\mu$  of degree less than or equal to  $k$ , ( $k \geq 1$ ).

Then, we have the following well-known result [6]:

**Lemma 2.1** *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  $P_k(K_\mu)$ , ( $1 \leq \mu \leq N_K$ ).*

*Then,  $u_h$  converges to  $u$  in  $H^1(\Omega)$ :*

$$\lim_{h \rightarrow 0} \|u_h - u\|_{1,\Omega} = 0. \quad (3)$$

*Moreover, if the exact solution  $u$  belongs to  $H^{k+1}(\Omega)$ , we have the following error estimate:*

$$\|u_h - u\|_{1,\Omega} \leq \mathcal{C}_k h^k |u|_{k+1,\Omega}, \quad (4)$$

*where  $\mathcal{C}_k$  is a positive constant independent of  $h$ ,  $\|\cdot\|_{1,\Omega}$  the classical norm in  $H^1(\Omega)$  and  $|\cdot|_{k+1,\Omega}$  denotes the semi-norm in  $H^{k+1}(\Omega)$ .*

In the sequel, we remind the probability law we derived in [3] which allowed us to evaluate the relative global accuracy, measured in  $H^1$ -norm, between two Lagrange finite elements.

We consider two families of Lagrange finite elements  $P_k$  and  $P_m$  corresponding to a set of values  $(k, m) \in \mathbb{N}^2$  such that  $0 < k < m$ .

The two corresponding inequalities given by (4), assuming that the solution  $u$  to **(VP)** belongs to  $H^{m+1}(\Omega)$ , are respectively written as:

$$\|u_h^{(k)} - u\|_{1,\Omega} \leq \mathcal{C}_k h^k |u|_{k+1,\Omega}, \quad (5)$$

$$\|u_h^{(m)} - u\|_{1,\Omega} \leq \mathcal{C}_m h^m |u|_{m+1,\Omega}, \quad (6)$$

where  $u_h^{(k)}$  and  $u_h^{(m)}$  respectively denotes the  $P_k$  and  $P_m$  Lagrange finite element approximations of  $u$ , solution to **(VP)** <sub>$h$</sub> .

In what follows, for simplicity, we set  $C_k \equiv \mathcal{C}_k |u|_{k+1,\Omega}$  and  $C_m \equiv \mathcal{C}_m |u|_{m+1,\Omega}$ . Therefore, inequalities (5) and (6) become:

$$\|u_h^{(k)} - u\|_{1,\Omega} \leq C_k h^k, \quad (7)$$

$$\|u_h^{(m)} - u\|_{1,\Omega} \leq C_m h^m, \quad (8)$$

which show that the two polynomial curves defined by  $f_k(h) \equiv C_k h^k$  and  $f_m(h) \equiv C_m h^m$  are the upper bounds of the possible values for the two norms  $\|u_h^{(k)} - u\|_{1,\Omega}$  and  $\|u_h^{(m)} - u\|_{1,\Omega}$ . More precisely, inequality (7) (resp. (8)) indicates that the norm  $\|u_h^{(k)} - u\|_{1,\Omega}$  (resp.  $\|u_h^{(m)} - u\|_{1,\Omega}$ ) is below the curve  $f_k(h)$  (resp.  $f_m(h)$ ).

As we are interested in comparing the relative positions of these two curves, we introduce their intersection point  $h^*$  defined by:

$$h^* \equiv \left( \frac{C_k}{C_m} \right)^{\frac{1}{m-k}} = \left( \frac{\mathcal{C}_k |u|_{k+1,\Omega}}{\mathcal{C}_m |u|_{m+1,\Omega}} \right)^{\frac{1}{m-k}}. \quad (9)$$

Now, as often in numerical analysis, there is no *a priori* information to surely or better specify the distance between  $\|u_h^{(k)} - u\|_{1,\Omega}$  (resp.  $\|u_h^{(m)} - u\|_{1,\Omega}$ ) and the curve  $f_k$  or its precise value in the interval  $[0, C_k h^k]$  due to (7) (resp. the curve  $f_m$  and the interval  $[0, C_m h^m]$  due to (8)).

Indeed, this situation is the consequence of two main ingredients:

1. The solution  $u$  of **(VP)** is unknown, which motivates the use of a  $P_k$  finite element method to build an approximation  $u_h^{(k)}$ ,
2. The way the mesh generator processes the mesh is *a priori* random which leads to a corresponding random approximation  $u_h^{(k)}$  too.

It is the reason why we treat the possible values of the norm  $\|u_h^{(k)} - u\|_{1,\Omega}$  as a random variable defined as follows:

Let us consider an experiment where the constitution of a random grid  $\mathcal{T}_h$  and the corresponding random approximation  $u_h^{(k)}$  are involved. Therefore, the approximation error  $\|u_h^{(k)} - u\|_{1,\Omega}$  can also be viewed as a random variable, defined by the following probabilistic framework:

- A *random trial* corresponds to the grid constitution and the associated approximation  $u_h^{(k)}$ .
- The probability space  $\mathbf{\Omega}$  contains therefore all the possible results for a given random trial, namely, for all of the possible grids that the mesh generator may processed, or equivalently, for all the corresponding approximations  $u_h^{(k)}$ .

Then, for a fixed value of  $k$ , we define the random variable  $X^{(k)}$  by:

$$\begin{aligned} X^{(k)} : \quad \Omega &\rightarrow [0, C_k h^k] \\ \omega \equiv u_h^{(k)} &\mapsto X^{(k)}(\omega) = X^{(k)}(u_h^{(k)}) = \|u_h^{(k)} - u\|_{1,\Omega}. \end{aligned} \quad (10)$$

In the sequel, for simplicity, we will set:  $X^{(k)}(u_h^{(k)}) \equiv X^{(k)}(h)$ .

Now, regarding the absence of information concerning the more likely or less likely values of the norm  $\|u_h^{(k)} - u\|_{1,\Omega}$  in the interval  $[0, C_k h^k]$ , we will assume that the random variable  $X^{(k)}(h)$  has a uniform distribution on the interval  $[0, C_k h^k]$  in the following meaning:

$$\forall (\alpha, \beta) \in \mathbb{R}_+^2, 0 \leq \alpha < \beta \leq C_k h^k : \text{Prob} \left\{ X^{(k)}(h) \in [\alpha, \beta] \right\} = \frac{\beta - \alpha}{C_k h^k}. \quad (11)$$

Equation (11) means that if one slides the interval  $[\alpha, \beta]$  anywhere in  $[0, C_k h^k]$ , the probability of the event  $\left\{ X^{(k)}(h) \in [\alpha, \beta] \right\}$  does not depend on where the interval  $[\alpha, \beta]$  is in  $[0, C_k h^k]$ ; this is the property of uniformity of the random variable  $X^{(k)}$ .

We are now able to evaluate the probability of the event

$$\left\{ \|u_h^{(m)} - u\|_{1,\Omega} \leq \|u_h^{(k)} - u\|_{1,\Omega} \right\} \equiv \left\{ X^{(m)}(h) \leq X^{(k)}(h) \right\} \quad (12)$$

to estimate the relative global accuracy between two finite elements of order  $k$  and  $m$ , ( $k < m$ ).

Let first of all define the relative global accuracy between two Lagrange finite elements  $P_k$  and  $P_m$ , ( $k < m$ ), as follows:

**Definition 2.2** *Let  $P_k$  and  $P_m$  ( $k < m$ ) be two Lagrange finite elements.*

*Then, we will say that " $P_m$  is globally more accurate than  $P_k$ " if:*

$$\|u_h^{(m)} - u\|_{1,\Omega} \leq \|u_h^{(k)} - u\|_{1,\Omega}. \quad (13)$$

We will recall now the probabilistic law established in [2] or [4] to get an estimation on the relative global accuracy between two Lagrange finite elements  $P_k$  and  $P_m$ , ( $k < m$ ), for a fixed mesh size  $h$ .

**Theorem 2.3** *Let  $u \in H^{k+1}(\Omega)$  be the solution to (1), and  $u_h^{(i)}$ , ( $i = k$  or  $i = m, k < m$ ), the two corresponding Lagrange finite element  $P_i$  approximations solution to (2).*

*We assume the two corresponding random variables  $X^{(i)}(h)$ , ( $i = k$  or  $i = m$ ), defined by (10) are independent and uniformly distributed on  $[0, C_i h^i]$ , where  $C_i$  are defined by (7)-(8).*

*Then, the probability such that " $P_m$  is globally more accurate than  $P_k$ ", as introduced in (2.2), is given by:*

$$\text{Prob} \left\{ X^{(m)}(h) \leq X^{(k)}(h) \right\} = \begin{cases} 1 - \frac{1}{2} \left( \frac{h}{h^*} \right)^{m-k} & \text{if } 0 < h \leq h^*, \\ \frac{1}{2} \left( \frac{h^*}{h} \right)^{m-k} & \text{if } h \geq h^*. \end{cases} \quad (14)$$

*Remark 1 Immediate interpretations of the non linear probability law (14) are available:*

1.  $P_m$  finite element is not only asymptotically more accurate than  $P_k$  as  $k < m$ , when  $h$  goes to 0, as usually considered and as a consequence of error estimate (4).

Indeed, for all  $h \leq h^*$ , the probability of " $P_m$  is more accurate than  $P_k$ " belongs to  $[0.5, 1]$ . It means that  $P_m$  is more likely accurate than  $P_k$  for all of these values of  $h$ , and not only for arbitrarily small values of  $h$ .

2. On the contrary when  $k < m$ , the finite element  $P_k$  is more likely accurate when  $h > h^*$ . This new point of view allows us to recommend that, for specific situations like for adaptive mesh refinement, this finite element would be more appropriated, as long as one would be able to detect that  $h > h^*$ .

The next section is devoted to a possible application of the probabilistic law (14) to a family of meshes, for example, in the process of mesh refinement.

### 3 Extension to a family of meshes

The aim of this section is to extend to a family of meshes the previous results we recalled for a given mesh  $\mathcal{T}_h$ .

For this purpose, we introduce a family of  $N$  regular meshes  $(\mathcal{T}^{(n)})_{n=1,N}$  made of simplexes, each mesh  $\mathcal{T}^{(n)}$  being characterized by its mesh size  $h_n$ . Let us also consider two Lagrange finite elements  $P_k$  and  $P_m$ , ( $k < m$ ).

Therefore, for each mesh  $\mathcal{T}^{(n)}$ , we can write the corresponding probability law (14) for the event  $\{X^{(m)}(h_n) \leq X^{(k)}(h_n)\}$ , associated to the given mesh size  $h_n$ .

Our aim is now to evaluate the probability such that exactly  $n_e$  meshes, ( $n_e = 0, 1, \dots, N$ ), satisfies " $P_m$  is more accurate than  $P_k$ ". To this end, let us introduce the sequence of  $N$  independent Bernoulli random variables  $(Y_n)_{n=1,N}$  defined by:

$$Y_n = \begin{cases} 1 & \text{if } X^{(m)}(h_n) \leq X^{(k)}(h_n), \\ 0 & \text{if not,} \end{cases} \quad (15)$$

and also the random variable  $S_N$  determined by:

$$S_N = \sum_{n=1}^N Y_n. \quad (16)$$

As each Bernoulli variable  $Y_n$  indicates if " $P_m$  is more accurate than  $P_k$ " or not on the corresponding mesh  $\mathcal{T}^{(n)}$ ,  $S_N$  describes the number of meshes on the all  $N$  meshes such that " $P_m$  is more accurate than  $P_k$ ".

*Remark 2* For any meshes  $\mathcal{T}^{(p)}$  and  $\mathcal{T}^{(q)}$  which belong to  $(\mathcal{T}^{(n)})_{n=1,N}$ , characterized by their mesh size  $h_p$  and  $h_q$ , the knowledge of the event " $P_m$  is more accurate than  $P_k$ " on  $\mathcal{T}^{(p)}$  does not enable us to conclude anything on  $\mathcal{T}^{(q)}$ .

Hence, the  $N$  random Bernoulli variables  $Y_n$ , ( $1 \leq n \leq N$ ), are considered as independent.

It comes out from the nature of the event " $P_m$  is more accurate than  $P_k$ ". Indeed, for an unknown exact solution  $u$  and the corresponding two approximations  $u_{h_p}^{(i)}$  and  $u_{h_q}^{(i)}$ , ( $i = k$  or  $i = m$ ), one cannot link the value  $Y_p$  and  $Y_q$ , associated to a mesh size  $h_p$  and  $h_q$ .

Therefore, we have the following result:

**Proposition 3.1** *The distribution of probabilities corresponding to the exact number of meshes satisfying "P<sub>m</sub> is more accurate than P<sub>k</sub>" is given by:*

$$\text{Prob}\{S_N = 0\} = (1 - \mathcal{P}(h_1)) \dots (1 - \mathcal{P}(h_N)), \quad (17)$$

$$\text{Prob}\{S_N = n_e\} = \sum_{\substack{(i_1, \dots, i_N) \in \{1, \dots, N\} \\ \text{and } i_l \neq i_q \text{ for } l \neq q}} \mathcal{P}(h_{i_1}) \dots \mathcal{P}(h_{i_n}) \dots (1 - \mathcal{P}(h_{i_{n+1}})) \dots (1 - \mathcal{P}(h_{i_N})), \quad 1 \leq n_e \leq N - 1, \quad (18)$$

$$\text{Prob}\{S_N = N\} = \mathcal{P}(h_1) \dots \mathcal{P}(h_N), \quad (19)$$

where the quantities  $\mathcal{P}(h_{i_j}) \equiv \text{Prob}\{X^{(m)}(h_{i_j}) \leq X^{(k)}(h_{i_j})\}$  are given by the probability distribution (14) of theorem 2.3.

Later in the paper, the sequence of values  $\left(\text{Prob}\{S_N = n_e\}\right)_{n_e=0, N}$  will be called the exact probabilistic distribution of the relative *global accuracy*.

*Proof* : To establish formulas (17)-(19), it is sufficient to notice that the  $N$  random Bernoulli variables  $(Y_n)_{n=1, N}$  defined by (15) are independent.

Indeed, (17)-(19) is a direct generalization of the binomial law for the variable  $S_N$  we would have to consider if all the Bernoulli variables  $(Y_n)_{n=1, N}$  had been defined by the same probability  $p$ , given by:

$$p \equiv \text{Prob}\{Y_n = 1\} = \text{Prob}\{X^{(m)}(h_n) \leq X^{(k)}(h_n)\}, \quad \forall n = 1 \text{ to } N. \quad \blacksquare$$

*Remark 3* From proposition 3.1 one can also get the cumulated distribution of the number of meshes for which "P<sub>m</sub> is more accurate than P<sub>k</sub>". Namely it corresponds to the probability such that at least  $n$  meshes,  $(n = 1, \dots, N)$ , are such that "P<sub>m</sub> is more accurate than P<sub>k</sub>".

Using the definition of the random variable  $S_N$  given by (16), it corresponds to the probabilities defined by:

$$\forall n_e = 1, \dots, N : \text{Prob}\{S_N \geq n_e\} = \sum_{j=n_e}^N \text{Prob}\{S_N = j\}. \quad (20)$$

Formula (20) is not easy to explicitly expressed due to formulas (17)-(19). However, in the next proposition, we will prove a recurrence relation which allows us to determine an algorithm to compute the exact probabilistic distribution as well as the cumulated one.

To this end, let  $p_N$  denotes the probability defined by  $p_N \equiv \text{Prob}\{Y_N = 1\}$ . Then, we have the following result:

**Proposition 3.2** *Let  $(\mathcal{T}^{(n)})_{n=1, N}$  be a family of  $N$  regular meshes composed by simplexes, each mesh  $\mathcal{T}^{(n)}$  being characterized by its mesh size  $h_n$ .*

*Then, we have:*

$$\forall n_e = 1, N : \text{Prob}\{S_N = n_e\} = p_N \text{Prob}\{S_{N-1} = n_e - 1\} + (1 - p_N) \text{Prob}\{S_{N-1} = n_e\}. \quad (21)$$

*Proof* : Formula (21) corresponds to the decomposition of the event  $\{S_N = n\}$  into two independent events which are:

$$\left(\{S_{N-1} = n_e - 1\} \cap \{Y_N = 1\}\right) \text{ and } \left(\{S_{N-1} = n_e\} \cap \{Y_N = 0\}\right).$$

Then, (21) results from classical probabilistic property. ■

*Remark 4* Proposition 3.2 enables us to process the computation of the exact probabilistic law which corresponds to  $\left(\text{Prob}\{S_N = n_e\}\right)_{n_e=0,N}$ , and therefore, the cumulated one too.

Indeed, first of all, let us notice that:

$$\forall n = 1, N : \text{Prob}\{S_n = n\} = \text{Prob}\left\{\sum_{k=1}^n Y_k = 1\right\} = \sum_{k=1}^n \text{Prob}\{Y_k = 1\} = p_1 \dots p_n, \quad (22)$$

as the bernoulli variables  $(Y_k)_{k=1,n}$ ,  $(\forall n = 1, N)$ , are independent.

Moreover, with the same arguments, we also notice that:

$$\forall n = 1, N : \text{Prob}\{S_n = 0\} = (1 - p_1) \dots (1 - p_n). \quad (23)$$

Therefore, by the help of the recurrence relation (21) and (22)-(23), one can compute the exact probabilistic law  $\left(\text{Prob}\{S_N = n_e\}\right)_{n_e=0,N}$ , step by step, as follows:

Relations (22)-(23) directly give:  $\text{Prob}\{S_n = 0\}$  and  $\text{Prob}\{S_n = n\}$ ,  $\forall n = 1, N$ .

- Step 1 :  $\text{Prob}\{S_2 = 1\}$  by (21),
- Step 2 :  $\text{Prob}\{S_3 = 1\}, \text{Prob}\{S_3 = 2\}$  by (21),
- ... : ..., ...,
- Step N :  $\text{Prob}\{S_N = 1\}, \dots, \text{Prob}\{S_N = N - 1\}$  by (21),

*Remark 5* The cumulated probabilistic distribution is then a direct consequence of proposition 3.2 due to formula (20). Finally, in the particular case corresponding to  $n = 1$ , we can derive an explicit expression of  $\text{Prob}\{S_N \geq 1\}$ . This is the purpose of the next proposition.

**Proposition 3.3** Let  $N$  be the total number of meshes which belong to a family of regular meshes  $(\mathcal{T}^{(n)})_{n=1,N}$  made of simplexes, and  $h_n$  denotes the mesh size of a given mesh  $\mathcal{T}^{(n)}$ . We assume that  $N = N_1 + N_2$ , where  $N_1$  is the number of meshes such that  $h_n \leq h^*$  and  $N_2$  those such that  $h_n > h^*$ .

Then, we have:

$$\text{Prob}\{S_N \geq 1\} = 1 - \frac{1}{2^{N_1}} \left[\frac{h_1 \dots h_{N_1}}{h^{*N_1}}\right]^{m-k} \left[1 - \frac{1}{2} \left(\frac{h^*}{h_{N_1+1}}\right)^{m-k}\right] \dots \left[1 - \frac{1}{2} \left(\frac{h^*}{h_N}\right)^{m-k}\right]. \quad (24)$$

*Proof* : By definition of the opposite event of  $\{S_N \geq 1\}$ , we can directly write:

$$\text{Prob}\{S_N \geq 1\} = 1 - \text{Prob}\{S_N = 0\} = 1 - (1 - \mathcal{P}(h_1)) \dots (1 - \mathcal{P}(h_N)), \quad (25)$$

where we used the expression of (17).

So, taking into account the probability law (14) of theorem 2.3 we can explicit the probability such that on  $N$  meshes, at least one mesh satisfies "  $P_m$  is more accurate than  $P_k$ ". Indeed, we conclude by using the definitions of  $N_1$  and  $N_2$  and the corresponding expressions in the probabilistic law (14) to get (24).  $\blacksquare$

*Remark 6* As we recall before, when  $k < m$ , it is often sought that the  $P_m$  finite element is more accurate than the  $P_k$  one. However, the probability given by (24) shows that even the event "There is a least one mesh among  $N$  meshes such that  $P_m$  is more accurate than  $P_k$ " is not a sure event. Nevertheless, one can prove that this event is asymptotically sure. It is the purpose of the next proposition based on the following lemma.

**Lemma 3.4** Let  $\beta$  be a real number such that  $0 < \beta < 1$ . Let  $p$  be a given integer and  $(x_n)_{n \geq p}$  be a sequence of real numbers satisfying:  $\forall n \geq p, 0 < x_n \leq \beta$ .

Then, the sequence product  $(\Pi_N)_{N \in \mathbb{N}}$  defines by:  $\Pi_N \equiv \prod_{n=p}^N x_n$  converges to 0 when  $N$  goes to  $+\infty$ .

*Proof :*

► Convergence of  $\Pi_N$ .

As the sequence  $(x_n)_{n \geq p}$  belongs to the interval  $]0, \beta]$  such that  $0 < \beta < 1$ , we can write that:

$$0 < \frac{\Pi_{N+1}}{\Pi_N} = x_{N+1} \leq \beta < 1. \quad (26)$$

Therefore, the sequence  $\Pi_N$  is a decreasing and bounded from below by 0, so it converges.

► Limit of  $\Pi_N$ .

To compute the limit of the sequence  $\Pi_N$  let us consider the following inequalities:

$$0 < x_p \dots x_N \leq \left( \max_{p \leq n \leq N} x_n \right)^{N-p+1} \leq \beta^{N-p+1} < 1, \quad (27)$$

since we assume that  $\beta$  belongs to  $]0, 1[$ .

As a consequence, the sequence  $(\beta^{N-p+1})_{N \in \mathbb{N}}$  goes to 0 when  $N$  goes to infinity, and due to Squeeze theorem [7], the sequence  $\Pi_N$  too.  $\blacksquare$

We are now in position to state the following proposition.

**Proposition 3.5** Let  $(\mathcal{T}^{(n)})_{n=1, N}$  be a given family of regular meshes and  $(h_n)_{n=1, N}$  the corresponding sequence of mesh sizes. We assume that there exists  $h_{max} \in R_+^*$  such that:

$$\forall n > 0 : h_n \leq h_{max} \text{ such that } 0 < \mathcal{P}(h_{max}) < 1. \quad (28)$$

Let also  $S_N$  be the random variable introduced in (16).

Then, we have:

$$\lim_{N \rightarrow +\infty} \text{Prob} \{S_N \geq 1\} = \lim_{N \rightarrow +\infty} \sum_{n_e=1}^N \text{Prob} \{S_N = n_e\} = 1. \quad (29)$$



*Proof* : We consider the expression of  $Prob \{S_N \geq 1\}$  given by (25) and we introduce the sequence  $(x_n)_{n=1,N}$  defined by:  $x_n \equiv 1 - \mathcal{P}(h_n)$ .

Then, one can write  $Prob \{S_N \geq 1\}$  as follows:

$$Prob \{S_N \geq 1\} = 1 - \prod_{n=1}^N x_n. \quad (30)$$

Now, one can check from formula (14) that

$$\forall n = 1, N : x_n = 1 - \mathcal{P}(h_n) \in ]0, 1[,$$

since  $\mathcal{P}(h_n) \neq 0$  and  $\mathcal{P}(h_n) \neq 1, \forall h_n, 0 < h_n \leq h_{\max}$ , as we assume (28).

Consequently, as the probabilistic law  $\mathcal{P}(h)$  defined by (14) is decreasing on  $\mathbb{R}_+$ , we have:

$$\forall n = 1, N : \mathcal{P}(h_n) \geq \mathcal{P}(h_{\max}), \quad (31)$$

and finally, it exists a real number  $\beta \equiv 1 - \mathcal{P}(h_{\max})$ , ( $0 < \beta < 1$ ), such that:

$$\forall n = 1, N : x_n = 1 - \mathcal{P}(h_n) \leq \beta < 1.$$

Therefore, from lemma 3.4 we obtain the claimed asymptotic behavior of  $Prob \{S_N \geq 1\}$ . ■

*Remark 7* We notice that in proposition 3.5 the assumption (28) corresponds to any concrete application where only finite mesh sizes are considered.

The idea of proposition 3.5 is that for a large number of meshes, the probability to get at least one mesh such that " $P_m$  is more accurate than  $P_k$ " goes to 1 with the number of meshes. This property could be taken into account for the adaptive mesh refinement, where a family of meshes is built to handle the large variations of the approximate solution.

In other words, this result indicates that for a large number of meshes, one has to consider that  $P_m$  finite elements will surely be more accurate, at least on one mesh. However, it is not a sufficient indication to motivate the implementation of these finite elements which are more expensive than the  $P_k$  finite elements when  $k < m$ .

Furthermore, to apply this approach to adaptive mesh refinement, one needs to transfer the results of this section which concern *global* behaviors of a family of meshes to get *local* information to refine a given mesh, mainly depending on the local gradient of the approximated solution. This is the purpose of the next section.

## 4 From a global to a local accuracy comparison of finite elements

### 4.1 A local probability law to compare accuracy of finite elements

In the previous sections, we described a probabilistic approach to estimate for a given mesh  $\mathcal{T}_h$  or for a sequence of meshes  $(\mathcal{T}^{(n)})_{n=1,N}$ , the relative global accuracy between two Lagrange finite elements  $P_k$  and  $P_m$ , ( $k < m$ ).

We obtained laws of probabilities that depend on the mesh size  $h$ , namely the size of the largest diameter in the mesh. Accordingly, the results we got are *global* and do not explicitly consider any local behavior that could play an important role, particularly when one considers adaptive

mesh refinement.

Our purpose is now to derive a *local* comparison tool between two Lagrange finite elements. This approach will be possible if one recalls that the error estimate (4), deduced from Bramble-Hilbert's lemma, is elaborated by two main ingredients. The first one is Cea's lemma and the second one is the interpolation error (*see* [6]).

So, for any simplex  $K$  belonging to a regular mesh  $\mathcal{T}_h$ , let us introduce  $\Pi_K^{(k)}$ , the local  $K$ -Lagrange interpolation operator of degree  $k$ , to define the local interpolation by the help of polynomials of degree lower or equal to  $k$  on  $K$ .

Then, one can write the global interpolation error  $\|u - \Pi_h u\|_{1,\Omega}$  as follows:

$$\|u - \Pi_h u\|_{1,\Omega} = \left( \sum_{K \in \mathcal{T}_h} \|u - \Pi_K^{(k)} u\|_{1,K}^2 \right)^{1/2}, \quad (32)$$

where  $\Pi_h$  denotes the global Lagrange interpolation operator on the mesh  $\mathcal{T}_h$ .

So, under the conditions of theorem 2.1, for each  $K \in \mathcal{T}_h$ , to obtain the next local estimate, we follow P.A. Raviart and J. M. Thomas, (*see* [6]), to get:

$$\|u - \Pi_K^{(k)} u\|_{1,K} \leq \mathcal{C}'_k h_K^k |u|_{k+1,K} \leq \mathcal{C}'_k h_K^k |u|_{k+1,\Omega}, \quad (33)$$

where  $h_K$  denotes the diameter of the simplex  $K$  and  $\mathcal{C}'_k$  a positive constant which does not depend on  $K$  and  $h_K$ , but depends on the reference element defining the Lagrange  $P_k$  finite element [6].

Then, the quantity  $\mathcal{C}'_k |u|_{k+1,\Omega}$  does not depend on  $K$  too. This independency will be further crucial when we will extend our results to the mesh refinement process.

Moreover, as a consequence of Cea's lemma, we can also deal with the following estimate:

$$\|u - u_h\|_{1,\Omega} \leq \frac{M}{\alpha} \|u - \Pi_h u\|_{1,\Omega}, \quad (34)$$

where  $M$  is the continuity constant and  $\alpha$  the ellipticity constant of the bilinear form  $a(\cdot, \cdot)$ .

Now, due to (32) and (34), we highlight that the accuracy of a given finite element  $P_k$  can be *locally* characterized by estimate (33).

As a consequence, we define the relative *local accuracy* between two finite elements  $P_k$  and  $P_m$  as the relative local interpolation accuracy on a given simplex  $K$  as follows:

**Definition 4.1** *Let  $P_k$  and  $P_m$ , ( $k < m$ ), be two Lagrange finite elements and  $K$  a given simplex which belongs to  $\mathcal{T}_h$ . We will say that " $P_m$  is locally more accurate than  $P_k$  on  $K$ " if:*

$$\|u - \Pi_K^{(m)} u\|_{1,K} \leq \|u - \Pi_K^{(k)} u\|_{1,K}. \quad (35)$$

Therefore, if we assume that the exact solution  $u$  of the variational formulation **(VP)** belongs to  $H^{m+1}(\Omega)$ , we can write inequality (33) for both of the  $\Pi_K^{(i)}$ , ( $i = k$  or  $i = m$ ), local operators and we have:

$$\|u - \Pi_K^{(k)} u\|_{1,K} \leq \mathcal{C}'_k h_K^k |u|_{k+1,\Omega}, \quad (36)$$

$$\|u - \Pi_K^{(m)} u\|_{1,K} \leq \mathcal{C}'_m h_K^m |u|_{m+1,\Omega}. \quad (37)$$

So, if we set  $C'_k \equiv \mathcal{C}'_k |u|_{k+1,\Omega}$  and  $C'_m \equiv \mathcal{C}'_m |u|_{m+1,\Omega}$ , inequalities (36) and (37) become:

$$\|u - \Pi_K^{(k)} u\|_{1,K} \leq C'_k h_K^k, \quad (38)$$

$$\|u - \Pi_K^{(m)} u\|_{1,K} \leq C'_m h_K^m, \quad (39)$$

which are the twins of inequalities (7) and (8).

The main difference between (7)-(8) and (38)-(39) is the meaning of  $h$ . Here, in (38)-(39),  $h_K$  denotes the local diameter of the simplex  $K$  whereas in (7)-(8)  $h$  is the involved maximum mesh size of  $\mathcal{T}_h$ .

As a consequence, if we introduce the random variables  $X_K^{(i)}(h_K)$ , ( $i = k$  or  $i = m$  and  $k < m$ ), defined by:

$$X_K^{(i)}(h_K) \equiv \|u - \Pi_K^{(i)} u\|_{1,K}, \quad (40)$$

we can directly get the probability of the event  $\{X_K^{(m)}(h) \leq X_K^{(k)}(h)\}$  corresponding to " $P_m$  is locally more accurate than  $P_k$  on  $K$ " defined in 4.1 by adapting formulas (14) as follows:

**Corollary 4.2** *Let  $K$  be a given simplex of diameter  $h_K$  belonging to a given regular mesh  $\mathcal{T}_h$ . Let  $u \in H^{k+1}(\Omega)$  be the solution to (1) and  $u_h^{(i)}$ , ( $i = k$  or  $i = m$ ,  $k < m$ ), the corresponding Lagrange finite element  $P_i$  approximations solution to (2).*

*We assume the two corresponding random variables  $X_K^{(i)}(h)$ , ( $i = k$  or  $i = m$ ), defined by (40) are independent and uniformly distributed on  $[0, C'_i h_K^i]$ , where  $C'_i$  are defined by (38)-(39).*

*Then, the probability such that " $P_m$  is locally more accurate than  $P_k$  on  $K$ " is given by:*

$$\text{Prob} \left\{ X_K^{(m)}(h) \leq X_K^{(k)}(h_K) \right\} = \begin{cases} 1 - \frac{1}{2} \left( \frac{h_K}{h^*} \right)^{m-k} & \text{if } 0 < h_K \leq h^*, \\ \frac{1}{2} \left( \frac{h^*}{h_K} \right)^{m-k} & \text{if } h_K \geq h^*, \end{cases} \quad (41)$$

where  $h^*$  is defined by (9), but where the constants  $C'_k$  and  $C'_m$  introduced in (38) and (39) replace  $C_k$  and  $C_m$  defined in (7) and (8).

*Remark 8* We notice that the corresponding value of  $h^*$  does not depend on the simplex  $K$  as we consider in inequality (33) the semi-norm of  $u$  in  $H^{k+1}(\Omega)$ , on the one hand, and as the constant  $\mathcal{C}'_k$ , (due to  $\mathcal{C}_k$ ), does not depend on  $K$  too, as we mention above, on the other hand.

Therefore, formula (41) gives us an evaluation of the *local* accuracy comparison between two finite elements  $P_k$  and  $P_m$  based on the *local* comparison accuracy between the corresponding  $K$ -Lagrange interpolation errors of degrees  $k$  and  $m$ .

We are now in position to extend this local result to a sequence of simplexes which belongs to a fixed mesh and which are concerned by adaptive mesh refinement process. This is the purpose of the next subsection.

## 4.2 Toward applications for adaptive mesh refinement

We now consider a given mesh  $\mathcal{T}_h$  composed by  $N$  simplexes  $K_\mu$  whose diameters are denoted by  $(h_\mu)_{\mu=1,N}$ .

For each simplex  $K_\mu$ , ( $1 \leq \mu \leq N$ ), we consider the probability law of the event " $P_m$  is locally more accurate than  $P_k$  on  $K_\mu$ " which is given by (41). Now, as in section 3, let us introduce the  $N$  Bernoulli random independent variables  $(Y_\mu)_{\mu=1,N}$  defined by (15) where we replace  $X^{(m)}(h)$  by  $X_{K_\mu}^{(m)}(h)$ , and also the corresponding random variable  $S_N$  defined by (16).

Because the total similitude between the mathematical formalism of section 3 and the present one, we directly get the analogous formulas of (17)-(19) and (24) but with a total different meaning.

Particularly, by adapting (24) of proposition 3.3 to the current situation, we get the following result:

**Proposition 4.3** *Let  $N$  be the total number of simplexes of a given mesh  $\mathcal{T}_h$  splitted so as  $N = N_1 + N_2$ , where  $N_1$  denotes the number of simplexes satisfying  $h_\mu \leq h^*$  and  $N_2$  such that  $h_\mu > h^*$ . Then, we have:*

$$Prob\{S_N \geq 1\} = 1 - \frac{1}{2^{N_1}} \left[ \frac{h_1 \dots h_{N_1}}{h^{*N_1}} \right]^{m-k} \left[ 1 - \frac{1}{2} \left( \frac{h^*}{h_{N_1+1}} \right)^{m-k} \right] \dots \left[ 1 - \frac{1}{2} \left( \frac{h^*}{h_N} \right)^{m-k} \right]. \quad (42)$$

*Remark 9* We would like to highlight that even if the formalism is totally equivalent between this section and section 3, one has to carefully distinguish the different meanings of the two situations.

Indeed, here the Bernoulli variables  $(Y_\mu)_{\mu=1,N}$  determine for a given elementary mesh  $K_\mu$  if " $P_m$  is locally more accurate than  $P_k$  on  $K_\mu$ ", while the Bernoulli variables introduced in the previous section characterize if a given mesh belonging to a sequence of meshes is such that " $P_m$  is globally more accurate than  $P_k$ " on this mesh.

Proposition 4.3 shows again that, albeit if it is usually assumed that finite elements  $P_m$  are more accurate than  $P_k$ , ( $k < m$ ), formula (42) highlights that even " $to get at least one simplex on  $N$  such that  $P_m$  is locally more accurate than  $P_k$ " is not a sure event, as its probability is different from one. Moreover, one can get a similar result we got for proposition 3.5, but first of all, we have to adapt our framework to the adaptive mesh refinement process.$

So, let us now describe what happens for a given mesh in which adaptive mesh refinement is processed, to improve the solution in the areas such that the gradient of the approximated solution becomes large.

For our purpose, we distinguish for a given mesh  $\mathcal{T}_h$  two kinds of simplexes. Those who are not going to be changed, and those which will be refined. Let us denotes by  $N'$ , ( $N' < N$ ), the number of new simplexes introduced in the mesh by the refinement process.

We also introduce  $N_1$ , ( $N_1 \leq N'$ ), the number of simplexes such that  $h_\mu \leq h^*$  and  $N_2 = N' - N_1$  the rest of simplexes. Therefore, the random variable  $S_N$  defined by (16) becomes  $S_{N'}$ , which describes now the total number of new simplexes in the refinement process such that " $P_m$  is locally more accurate than  $P_k$ ".

We remark that these considerations make sense as we neutralize the dependency of  $h^*$  toward any considered simplex  $K$  which belongs to the given mesh  $\mathcal{T}_h$ , (see remark 8).

Then, our interest is to determine the behavior of  $Prob\{S_{N'} \geq 1\}$ , (equivalently determined by (42) if one changes  $N$  by  $N'$ ), when  $N'$  goes to infinity, namely where the number of new simplexes  $N'$  concerned by the refinement process becomes large.

This situation corresponds to the framework of adaptive mesh refinement where one is usually interested by locally refine the mesh: this is performed by identifying in the mesh the areas such that the gradient of the approximated solution is high.

Therefore, we are now in position to determine for a given set of new simplexes  $N'$  which becomes large, the asymptotic probability such that "there exists at least one new simplex where  $P_m$  is locally more accurate than  $P_k$ ".

This is the purpose of the following proposition which is the twin of proposition 3.5.

**Proposition 4.4** *Let  $\mathcal{T}_h$  be a given mesh composed by  $N$  simplexes and let  $N'$ , ( $N' < N$ ), be the number of new simplexes processed by a mesh refinement. We also assume that there exists  $h_{max} \in \mathbb{R}_+^*$  which satisfies the equivalent condition of (28) for all the diameters  $h_\mu$ , ( $\mu = 1, N'$ ), defining the new simplexes.*

Then we have:

$$\lim_{N' \rightarrow +\infty} Prob \{S_{N'} \geq 1\} = 1. \quad (43)$$

This result is not surprising due to the total similarity between formulas (24) and (42), (if one replaces  $N$  by  $N'$  in (42)).

Again, it shows that one has to carefully consider the relative local accuracy between  $P_m$  and  $P_k$  Lagrange finite elements, ( $k < m$ ), except if locally, the number of simplexes becomes very large.

However, for a fixed number  $N'$  of new simplexes, this phenomena is more pronounced depending on the minimum number of simplexes  $n'_e$ , ( $n'_e = 1, N'$ ), satisfying "  $P_m$  would be more accurate than  $P_k$ ". Unfortunately, deriving the analogous formulas of (17)-(19) and (20) for a sequence of simplexes, to explicit the probability of the event  $\{S_{N'} \geq n'_e\}$  is inextricable.

So, one can compare for two different values of  $n'_e$  the behavior of the corresponding probabilities. This is the purpose of the following proposition.

**Proposition 4.5** *Let  $n'_{e,1}$  and  $n'_{e,2}$  be two integers such that  $1 \leq n'_{e,1} < n'_{e,2} \leq N'$ , and let  $S_{N'}$  be the random variable equivalently defined as (16).*

Then, we have:

$$Prob \{S_{N'} \geq n'_{e,2}\} \leq Prob \{S_{N'} \geq n'_{e,1}\}. \quad (44)$$

*Proof* : The proof of proposition 4.5 results from the following identity:

$$Prob \{S_{N'} \geq n'_{e,1}\} = \sum_{n'_e = n'_{e,1}}^{n'_{e,2}-1} Prob \{S_{N'} = n'_e\} + Prob \{S_{N'} \geq n'_{e,2}\}, \quad (45)$$

and consequently, (44) holds. ■

Therefore, proposition 4.5 clearly indicates that the more  $n'_e$  the less the probability such that at least  $n'_e$  simplexes satisfy "  $P_m$  is locally more accurate than the  $P_k$ ".

This points out that we cannot easily disqualify the  $P_k$  finite element in comparison with the  $P_m$  one, particularly for a mesh refinement process.

*Remark 10* *We finally notice that it is also possible to compute the exact distribution and the associated cumulated one by applying the same principles we processed regarding the previous laws we derived for formula (21) of proposition 3.2 and for formulas (22) and (23) too.*

## 5 Discussion and conclusion

In this paper, we proposed a new geometrical-probabilistic approach to evaluate the relative accuracy between two Lagrange finite elements  $P_k$  and  $P_m$ , ( $k < m$ ). Basically, we distinguished two cases: a global approach and a local one. Both cases are based on a probabilistic interpretation of the error estimate one can derive from Bramble Hilbert lemma.

In the global approach, we recall the probabilistic law (14) we derived in [3] to propose a first extension to a family of meshes. Regarding the local accuracy between two finite elements, we recall the two main components required to establish the *a priori* error estimate and we highlighted that it is centrally based on the local interpolation error. This leads us to transpose our global analysis to the local one to get the corresponding probability distribution (41).

Afterwards, we extended our global and local results to two principal applications, both of them could concern with adaptive mesh refinement process. The global probability law has been used to describe the case of a family of meshes while the local probabilistic one helped us to consider a fixed mesh composed by a sequence of simplexes.

These results strengthen those we got in [3] and show that even if we consider a family of meshes (respectively a sequence of simplexes), the event "to get at least one mesh (respectively, at least one simplex) such that  $P_m$  is more globally (respectively, locally) accurate than  $P_k$ " is not a sure event, (cf. propositions 3.3, 4.3 and proposition 4.5 for a more general case).

However, it is proved in Propositions (3.5) and (4.4) that, for a great number of meshes (respectively, a great number of simplexes), this event is asymptotically sure.

As we noticed, these two applications - the family of meshes or the sequence of simplexes - may be useful for adaptive mesh refinement. We also proved a recurrence relation, (see formula (3.2)), in the case of a family of meshes, which can be adapted to the case of a sequence of simplexes.

This enables to compute, for example, the probability such that "at least 50 percents of meshes (resp., at least 50 percents of simplexes) are such that  $P_m$  is more globally (resp., locally) accurate than  $P_k$ ".

Finally, we have to mention that for all concrete applications, one will have to precisely estimate the critical value  $h^*$ . Indeed, all the probabilistic laws we derived are based on formulas (14) and (41) which depend on  $h^*$ .

Since  $h^*$  strongly depends on the semi-norm  $H^{k+1}(\Omega)$  of the exact solution  $u$  to the variational problem, all the available techniques which belong to *a priori* estimation theory of solutions to partial differential equations will be involved to evaluate  $h^*$ .

**Acknowledgements:** The authors want to warmly dedicate this research to pay homage to the memory of Professors André Avez and Gérard Tronel who largely promote the passion of research and teaching in mathematics.

## References

- [1] J. Chaskalovic, Mathematical and numerical methods for partial differential equations, Springer Verlag, (2013).

- [2] J. Chaskalovic, F. Assous, A new probabilistic interpretation of Bramble-Hilbert lemma, accepted in Computational Methods in Applied Mathematics, (2018).
- [3] J. Chaskalovic, F. Assous, A new probabilistic interpretation of Bramble-Hilbert lemma, October 2018. arXiv:1803.09547 [math.NA]
- [4] J. Chaskalovic, F. Assous, Asymptotic relation between probabilistic distributions which quantify the relative accuracy between two finite elements, October 2018. arXiv:1803.09552 [math.NA]
- [5] P.G. Ciarlet, Basic error estimates for elliptic problems, in Handbook of Numerical Analysis, Vol. II, Eds. P.G. Ciarlet and J. L. Lions, North Holland, (1991).
- [6] P.A. Raviart et J.M. Thomas, Introduction à l'analyse numérique des équations aux dérivées partielles, Masson (1982).
- [7] J. Stewart, Multivariable Calculus, Cengage Learning Publisher.(2008)