Indeterminate Constants in Numerical Approximations of PDEs: A Pilot Study Using Data Mining Techniques

Franck Assous, Joël Chaskalovic

Indeterminate Constants in Numerical Approximations of PDEs: A Pilot Study Using Data Mining Techniques

Franck Assous[(1)] **and J. Chaskalovic**[(2)]

[(1)] Department of Comp. Sc. & Maths., Ariel University, Ariel, Israel

[(2)] D'Alembert, Sorbonne University, Paris, France

Abstract

Rolle's theorem, and therefore, Lagrange and Taylor's theorems are responsible for the inability to determine precisely the error estimate of numerical methods applied to partial differential equations. Basically, this comes from the existence of a non unique unknown point which appears in the remainder of Taylor's expansion. In this paper we consider the case of finite elements method. We show in details how Taylor's theorem gives rise to indeterminate constants in the *a priori* error estimates. As a consequence, we highlight that classical conclusions have to be reformulated if one considers *local* error estimate. To illustrate our purpose, we consider the implementation of P_1 and P_2 finite elements method to solve Vlasov-Maxwell equations in a paraxial configuration. If Bramble-Hilbert theorem claims that *global* error estimates for finite elements P_2 are "*better*" than the P_1 ones, we show how data mining techniques are powerful to identify and to qualify when and where *local* numerical results of P_1 and P_2 are equivalent.

1 Introduction

As it is well known, there is always a difference - *an error* - between the real system that one wants to simulate and the numerical results produced by its approximation. In [2] and [3] we proposed a way to decompose this error in four possible kinds of errors, named *the modeling error*, *the approximation error*, *the parametrization error* and *the discretization error*.

We already showed in [1] how data mining techniques could help us to evaluate *the modeling error*: As an example, we considered there the case of an asymptotic paraxial approximation that models an ultrarelativistic particle beam.

In the case we considered there, despite a convergence result proved in [10], a numerical study (see [5]) shows that the comparison of the different orders of approximations is not obvious. In other terms, the asymptotic models are often difficult to compare directly one to the other.

The main cause of this situation is due to Rolle's theorem, and therefore, Lagrange and Taylor's theorems [7] which are responsible for the inability to exactly determine constants appearing in the residual of Taylor series, which are considered in asymptotic expansions.

More precisely, this inability is mainly due to the existence of a non unique unknown point which appears in the remainder of Taylor's expansion for a concerned higher-order derivative.

The same reason is responsible to the underdetermination of the constants involved in the error estimate of numerical methods applied to partial differential equations.

Our aim in this paper is to analyze the concrete repercussion of this phenomenon by data mining techniques, which allow us to compare the effective accuracy of two numerical methods.

For this purpose, we explicit the particular case of finite elements method and we show how Taylors' theorem causes the presence of indeterminate constants which appear in the *a priori* error estimates.

As a consequence, we highlight that classical conclusions have to be reformulated if one considers *local* error estimate. To specify this point, we consider, as a pilot study, the implementation of P_1 and P_2 finite elements method to solve the Vlasov-Maxwell equations in a paraxial model.

If Bramble-Hilbert theorem (see for instance [6]) claims that *global* error estimates for the finite elements P_2 are "*better*" than the P_1 ones, we will illustrate how data mining techniques are relevant and powerful to identify when and where *local* numerical results of P_1 and P_2 are equivalent.

2 The discretization error

The type of error we introduce here - *the discretization error* - can be viewed in some sense as complementary to the *approximation error* described in [3]. Indeed, let us consider a model of equations (E) that one wants to discretize with a numerical method chosen inside a given family of approximations. The internal numerical parameters like the mesh size or the time step are fixed. For this given family of approximations - for instance the P_k finite elements, $k = 1, 2, \ldots$ - we consider two numerical methods with different order of approximation, says (MN_1) and (MN_2). To make it more concrete, we consider the example of a P_1 and a P_2 finite element method. Then, the *discretization error* is defined as the error due to the difference of order between (MN_1) and (MN_2).

At first glance, this definition appears quite elementary and the resulting approximation expectable. Indeed, the Bramble-Hilbert theorem claims that, under certain conditions of regularity of the mesh and of the solution, the results obtained by the finite elements P_2 (of order 2) will be obviously more precise than those computed by finite elements P_1 (of order 1).

However, the definition of *the discretization error* is motivated by the presence of a constant in the error estimates. Indeed, these error estimates contain multiplicative constants, unknown or difficult to estimate, before h and h^2 - where h is the average diameter of the mesh size. Using obvious notations, the error estimates in the case considered here are written for a norm *ad hoc*:

$$|u_1(x,t) - u_{exact}(x,t)| \leq \|u_1 - u_{exact}\| \leq C_1 h, \tag{1}$$

and

$$|u_2(x,t) - u_{exact}(x,t)| \leq \|u_2 - u_{exact}\| \leq C_2 h^2, \tag{2}$$

where C_1, C_2 are two different constants.

Understanding and investigating the *discretization error* with data mining techniques is motivated by the possibility to determine if there exist *local* discriminant conditions which could specify this *global* error estimates.

More precisely, we are looking for states defined by particular values of (x^*, t^*) such that:

$$|u_1(x^*,t^*) - u_{exact}(x^*,t^*)| \;\leq\; |u_2(x^*,t^*) - u_{exact}(x^*,t^*)|\,, \tag{3}$$

or at least

$$|u_1(x^*,t^*) - u_{exact}(x^*,t^*)| \;\simeq\; |u_2(x^*,t^*) - u_{exact}(x^*,t^*)|\,. \tag{4}$$

For these local cases (x^*, t^*), one cannot longer claim that the P_2 finite element method is better than the P_1 one, and as a consequence, the P_2 finite elements are overqualified.

3 Finite elements error estimate

To better understand the mechanisms which are responsible of the inability to evaluate the two constants C_1 and C_2 which appear in (1)-(2), we will now analyze how the error estimate of the finite element method is built.

To precise our purpose, we restrict ourselves to the general variational formulation **(VP)** which has the following structure:

$$\textbf{(VP)} \quad \begin{cases} \text{Find } u_{exact} \in V \text{ solution to:} \\ a(u_{exact}, v) = L(v), \forall v \in V, \end{cases} \tag{5}$$

where we assume that $(V, |.\|_V)$ is a given Hilbert space, $a(\cdot,\cdot)$ is a given bilinear, continuous and $V-$ elliptic form defined on $V \times V$, and $L(\cdot)$ a given linear continuous form defined on V.

Let us consider now the approximation solution u_h of u_{exact}, that solves the approximate variational formulation $\textbf{(VP)}_h$ defined by:

$$\textbf{(VP)}_h \quad \begin{cases} \text{Find } u_h \in V_h \text{ solution to:} \\ a(u_h, v_h) = L(v_h), \forall v_h \in V_h, \end{cases} \tag{6}$$

where V_h is a given subset of V whose dimension is finite.

Then, the first step to estimate the error between the exact solution u_{exact} and its approximation u_h is given by Céa's Lemma [6] that we recall below:

Lemma 3.1 *Let u_{exact} be solution to (5) and u_h the solution of (6). Then, the following inequality holds:*

$$\|u_{exact} - u_h\|_V \leq C \inf_{v_h \in V_h} \|u_{exact} - v_h\|_V , \tag{7}$$

where the constant C is the ratio between the continuity constant and the ellipticity constant of the bilinear form $a(\cdot,\cdot)$.

The strategy which results from Céa's lemma is that we are able to replace the estimation error $\|u_{exact} - u_h\|_V$ by another approximate problem where the approximate solution u_h does not appear anymore. In the sequel, we will see that it can be proved that the exact solution u_{exact} is well approximated by a convenient choice of functions v_h in V_h.

To make our demonstration more concrete, we restrict ourselves another time to the particular family of finite elements P_1 and we suppose that the domain Ω, where the real system (S) exists and is modeled by equations (E), (or by the variational formulation **(VP)**), is a two dimensional bounded open subset of $\mathbb{R}^2$ recovered by a mesh Π_h composed by N_T triangles $T_k, (k = 1, N_T)$, which respect classical rules of discretization, (see for example [6]).

Let us consider T, an elementary triangle in the mesh Π_h, defined by its three vertices (A_1, A_2, A_3), (see Figure 1).

Then, the relative interpolate operator π_T on the triangle T is defined by:

$$\begin{aligned} \pi_T \quad : \quad C^0(T) \ &\rightarrow \ P^1(T) \\ v \quad &\rightsquigarrow \ \pi_T(v) \equiv \sum_{k=1}^{3} v(A_k)\lambda_k, \end{aligned} \tag{8}$$

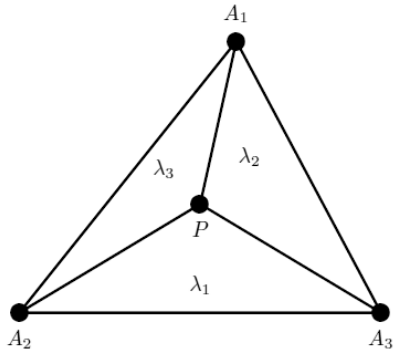

Figure 1: Example of triangle.

where $P^1(T)$ comprises the polynomial functions defined on the triangle T of degree less than or equal to one, the $\lambda_k, (k = 1, 3)$, denote the three barycentric functions defined on T [6] which satisfy the two following properties:

$$\forall P \in T \ : \ \sum_{k=1}^{3} \lambda_k(P) A_k \ = \ P \ , \ \sum_{k=1}^{3} \lambda_k(P) \ = \ 1. \tag{9}$$

We define now two geometrical parameters $diam(T)$ and $\rho(T)$ which are respectively the diameter of the triangle T, (the greatest distance between two points inside T), and the roundness of T, i.e. the diameter of the biggest circle which is inside T. We are now in position to formulate the first main result which evaluates the distance between a "smooth function" v and its interpolate $\pi_T(v)$ on T.

Theorem 3.1 *For any regular function v which belongs to $C^2(T)$, the following inequalities hold:*

$$\forall P \in T \quad : \quad |v(P) - \pi_T(v)(P)| \ \leq \ \frac{1}{2} \ diam(T)^2 \ \|D^2(v)\|_{L^\infty(T)}, \tag{10}$$

$$\forall P \in T \quad : \quad |\nabla v(P) - \nabla \pi_T(v)(P)| \ \leq \ \frac{1}{2} \ \frac{diam(T)^2}{\rho(T)} \ \|D^2(v)\|_{L^\infty(T)}, \tag{11}$$

where:

$$\|D^2(v)\|_{L^\infty(T)} \ = \ Max \left(\sup_{M \in T} \left| \frac{\partial^2 v}{\partial x^2}(M) \right|, \sup_{M \in T} \left| \frac{\partial^2 v}{\partial y^2}(M) \right|, \sup_{M \in T} \left| \frac{\partial^2 v}{\partial x \partial y}(M) \right| \right). \tag{12}$$

Proof :

- Due to our assumption that v belongs to $C^2(T)$ we can use Taylor's formula which claims that it exists an ***unknown*** point ξ_k in $]A_k, P[$ such that:

$$v(A_k) = v(P) + \nabla v(P)(A_k - P) + \frac{1}{2}(A_k - P)^t . D^2 v(\xi_k)(A_k - P). \tag{13}$$

  Then, let us compute the quantity $\pi_T(v)(P) - v(P)$:

$$\pi_T(v)(P) - v(P) \ = \ \sum_{k=1}^{3} v(A_k)\lambda_k(P) - v(P) \ = \ \frac{1}{2} \sum_{k=1}^{3} \lambda_k(P)(A_k - P)^t . D^2 v(\xi_k)(A_k - P), \tag{14}$$

where we used the two properties (9).

The control interpolation inequality (10) is therefore got because:

$$|A_k - P| \leq diam(T).$$

▶ Concerning the inequality (11), we multiply (13) by $\nabla\lambda_k$ and we sum over k:

$$\begin{aligned}\sum_{k=1}^{3} v(A_k)\nabla\lambda_k &= v(P)\sum_{k=1}^{3}\nabla\lambda_k + \sum_{k=1}^{3}\Big[\nabla v(P)(A_k-P)\Big]\nabla\lambda_k \\ &+ \frac{1}{2}\sum_{k=1}^{3}(A_k-P)^t.D^2v(\xi_k)(A_k-P)\nabla\lambda_k. \qquad (15)\end{aligned}$$

Moreover, by applying the gradient to the two properties of (9), we get:

$$\forall P\in T\ :\ \sum_{k=1}^{3}\nabla\lambda_k(P)A_k^t = \nabla P = I \text{ and } \sum_{k=1}^{3}\nabla\lambda_k(P) = 0, \qquad (16)$$

where I denotes the identity matrix.

Then, equation (15) can be written as follows:

$$\begin{aligned}\sum_{k=1}^{3} v(A_k)\nabla\lambda_k &= \left[\sum_{k=1}^{3}\nabla\lambda_k(A_k-P)^t\right]\nabla v(P) \\ &+ \frac{1}{2}\sum_{k=1}^{3}(A_k-P)^t.D^2v(\xi_k)(A_k-P)\nabla\lambda_k, \qquad (17)\end{aligned}$$

if we remark that the j^{th} component of the vector $\big[\nabla v(P)(A_k-P)\big]\nabla\lambda_k$ can be expressed by the help of the ij^{th} component of the matrix $\nabla\lambda_k(A_k-P)^t$:

$$\sum_{i=1,2}\frac{\partial v}{\partial x_i}(P)(A_k-P)_i\frac{\partial\lambda_k}{\partial x_j} = \sum_{i=1,2}\frac{\partial\lambda_k}{\partial x_j}(A_k-P)_i\frac{\partial v}{\partial x_i}(P). \qquad (18)$$

Finally, due to (16), we notice that:

$$\begin{aligned}\left[\sum_{k=1}^{3}\nabla\lambda_k(A_k-P)^t\right]\nabla v(P) &= \left[\sum_{k=1}^{3}\nabla\lambda_k A_k^t\right]\nabla v(P) - \left[\sum_{k=1}^{3}\nabla\lambda_k P^t\right]\nabla v(P), \qquad (19)\\ &= \nabla v(P). \qquad (20)\end{aligned}$$

This allows us to rewrite (17) in its final expression:

$$\sum_{k=1}^{3} v(A_k)\nabla\lambda_k = \nabla v(P) + \frac{1}{2}\sum_{k=1}^{3}(A_k-P)^t.D^2v(\xi_k)(A_k-P)\nabla\lambda_k, \qquad (21)$$

or, in another words:

$$\nabla\pi_T(v)(P) - \nabla v(P) = \frac{1}{2}\sum_{k=1}^{3}(A_k-P)^t.D^2v(\xi_k)(A_k-P)\nabla\lambda_k. \qquad (22)$$

Inequality (11) is then obtained as follows:

$$|\nabla\pi_T(v)(P) - \nabla v(P)| \leq \frac{1}{2}\, diam(T)^2 \sup_{\substack{k=1,2,3 \\ P\in T}} |\nabla\lambda_k(P)|\, \|D^2 v\|_{L^\infty(T)}, \tag{23}$$

$$\leq \frac{1}{2}\frac{diam(T)^2}{\rho(T)}\|D^2 v\|_{L^\infty(T)}, \tag{24}$$

where we used the following lemma.

■

Lemma 3.2

$$\forall P \in T,\ \forall k = 1,2,3\ :\ |\nabla\lambda_k(P)| \ \leq\ \frac{1}{\rho(T)}. \tag{25}$$

We refer to [11] for the corresponding proof.

Fundamental remark: As one can see, the two interpolation estimations (10)-(11) were derived by using Taylor's expansion where an unknown point ξ_k has been introduced. As we already mentioned, (see section 1), the fact that these points are totally unknown is due to Rolle's theorem [7] from which Taylor's theorem is established.

As a consequence, as we will see in the following result, the error estimate of finite element method is completely dependent of these ***unknown*** points.

This was what we claimed in the beginning of section, that is that Bramble-Hilbert theorem does not allowed to get a precise error estimate because the presence of ***unknown*** constants.

To precise this statement, let us introduce the global interpolate operator π_h defined by:

$$\begin{aligned} \pi_h \ :\ C^0(\Omega) &\to V_h \\ v \quad &\rightsquigarrow \quad \pi_h(v) \ \equiv\ \sum_{k=1}^{N_s} v(A_i)\varphi_i, \end{aligned} \tag{26}$$

where $A_i, (i = 1, N_s)$, are the nodes of the corresponding mesh Π_h built with triangles, φ_i the canonical functions - $\varphi_i(A_j) = \delta_{ij}$, δ_{ij} is the kronecker symbol - of the approximate functional space V_h defined by:

$$V_h \ =\ \{v_h : \Omega \to \mathbb{R},\ v_h|_{Tk} \in P_1(T_k), (k = 1, N_T)\}\,. \tag{27}$$

We are now in position to formulate the second main theorem for the error estimate of the finite element method in relation with the framework we considered.

Theorem 3.2 *Let Ω be a polygonal subset of $\mathbb{R}^2$ recovered by a "regular" mesh Π_h. Suppose that exact the solution u_{exact} of the variational formulation **(VP)** belongs to $C^2(\Omega)$ and let us consider u_h its corresponding approximation solution of $\mathbf{(VP)}_h$ based on P_1 finite elements. Then, u_h converges to u when h tends to zero and we have the following error estimations:*

$$\|u_{exact} - u_h\|_{L^2(\Omega)} \leq \beta_0 h^2\, \|D^2 u_{exact}\|_{L^\infty(\Omega)}, \tag{28}$$

$$\|u_{exact} - u_h\|_{H^1(\Omega)} \leq \beta_1 h\, \|D^2 u_{exact}\|_{L^\infty(\Omega)}, \tag{29}$$

where β_0 and β_1 are two constant which do not depend on h.

Proof :

- For simplicity we set $u_{exact} \equiv u$. Then, because of Céa's lemma, it is sufficient to find out an approximate function $v_h \in V_h$ such that:

$$\|u - v_h\|_{H^1(\Omega)} \leq Ch^2 \|D^2 u\|_{L^\infty(\Omega)}. \tag{30}$$

Obviously, we will consider for v_h the interpolate function $\pi_h(u)$ defined by (26). We first begin to evaluate the L^2-norm of the quantity $u - \pi_h(u)$:

$$\|u - \pi_h(u)\|^2_{L^2(\Omega)} = \sum_{T_k \in \Pi_h} \int_{T_k} |u - \pi_h(u)|^2 d\Omega \leq \sum_{T_k \in \Pi_h} \int_{T_k} \left| u_{|_{T_k}} - \pi_{_{T_k}}(u_{|_{T_k}}) \right|^2 d\Omega, \tag{31}$$

where we used the following property:

$$\forall\, T_k \in \Pi_h : (\pi_h\, u)_{|_{T_k}} = \pi_{_{T_k}}(u_{|_{T_k}}). \tag{32}$$

Then, because of (10) inequality (31) becomes:

$$\|u - \pi_h(u)\|^2_{L^2(\Omega)} \leq \frac{1}{4} \sum_{T_k \in \Pi_h} diam(T_k)^4 \|D^2 u\|^2_{L^\infty(T_k)}\, \mu(T_k), \tag{33}$$

$$\leq \frac{h^4}{4}\, \|D^2 u\|^2_{L^\infty(\Omega)}\, \mu(\Omega), \tag{34}$$

where $\mu(\Omega)$, (respectively $\mu(T_k)$), denotes the measure of Ω, (respectively the measure of T_k).

Then, we get (28) if set:

$$\beta_0 \equiv \frac{C}{2} \sqrt{\mu(\Omega)},$$

(C is the constant in (7) which appears in Céa's lemma).

- Concerning inequality (29), we have in the same way:

$$\|\nabla u - \nabla \pi_h(u)\|^2_{L^2(\Omega)} = \sum_{T_k \in \Pi_h} \int_{T_k} |\nabla u - \nabla \pi_h(u)|^2 d\Omega, \tag{35}$$

$$\leq \sum_{T_k \in \Pi_h} \int_{T_k} \left| \nabla u_{|_{T_k}} - \nabla \pi_{T_k}(u_{|_{T_k}}) \right|^2 d\Omega, \tag{36}$$

$$\leq \frac{1}{4} \sum_{T_k \in \Pi_h} \int_{T_k} \frac{diam(T_k)^4}{\rho(T_k)^2} \|D^2 u\|^2_{L^\infty(T_k)} \mu(T_k), \tag{37}$$

where we used inequality (11) of theorem 3.1.

Finally, to conclude, we use the regularity of the mesh Π_h, that it to say that we assume that there exists a constant σ greater or equal to 1 such that:

$$\forall\, T_k \in \Pi_h : \frac{diam(T_k)}{\rho(T_k)} \leq \sigma. \tag{38}$$

For more details, the reader would consult for example [12]. As a consequence, inequality (37) can be rewritten as follows:

$$\|\nabla u - \nabla \pi_h(u)\|^2_{L^2(\Omega)} \ \leq \ \frac{h^2}{4}\,\sigma^2 \|D^2 u\|^2_{L^\infty(\Omega)} \mu(\Omega). \tag{39}$$

Therefore, we get (29) after some calculations, if we set:

$$\beta_1 \equiv \frac{C}{2}\sqrt{(1+\sigma^2)\mu(\Omega)},$$

C is another time the constant which appears in Céa's lemma in (7).

■

Now, having proven that *global* error estimates for P_1 finite elements are based on unknown constants, we rewrite the inequality (29) as follows:

$$\|u_{exact} - u_h^{(1)}\|_{H^1(\Omega)} \ \leq \ \gamma_1 h, \tag{40}$$

where $u_h^{(1)}$ denotes the P_1 approximation and $\gamma_1 \equiv \beta_1 \|D^2 u_{exact}\|_{L^\infty(\Omega)}$.

In the same way, one can show, using the same ideas for theorem 3.2, that P_2 finite elements lead to the following result:

$$\|u_{exact} - u_h^{(2)}\|_{H^1(\Omega)} \ \leq \ \gamma_2 h^2, \tag{41}$$

where $u_h^{(2)}$ denotes the P_2 approximation and γ_2 the corresponding unknown constant.

Therefore, because the presence of the two unknown constants γ_1 and γ_2 in (40)-(41), we suspect the numerical following possibility to take place:

$$|u_{exact}(x^*,t^*) - u_h^{(1)}(x^*,t^*)| \ \leq \ |u_{exact}(x^*,t^*) - u_h^{(2)}(x^*,t^*)|\,, \tag{42}$$

or at least

$$|u_{exact}(x^*,t^*) - u_h^{(1)}(x^*,t^*)| \ \simeq \ |u_{exact}(x^*,t^*) - u_h^{(2)}(x^*,t^*)|\,, \tag{43}$$

which means that *locally*, the P_1 finite element solution might be either more accurate than the P_2 finite element solution(inequality (42)), or with the same accuracy (inequality (43)).

In the next section, our purpose will be to explore these possibilities, by implementing data mining techniques on numerical approximations collected inside a convenient database. Indeed, the aim of data mining is to explore a given database to identify, if any, and characterized homogenous subgroups which corresponds to the *local* property of data mining.

In our perspective, this *local* property in the database will be defined in relation to the time-space discretization of the physical system. As a consequence, we will construct a database in which each row will describe the numerical approximations of the physical unknowns *locally* computed for each time step and for a given space node of the considered mesh.

As we will see in the next section, data mining will appear as a "good candidate" to find out *local* features which correspond to the property we want to investigate, defined by (42) or (43).

4 Data mining methods for error source analysis

Whether of *modeling error*, *approximation error*, *parametrization error* or *discretization error*, to better demonstrate the objectives that we described above, we rather have to retain a mathematical model which will be rich enough to present a consistent number of variables, (the columns of the future databases established for each type of error), to provide an appropriate potential of exploration.

It is the reason why we retained, as a test case, the unsteady Vlasov-Maxwell equations, whose solutions can be described by asymptotic expansions, see [10].

Indeed, to model the propagation of ultrarelativistic particle beam, a systematic technique to derive approximate models, based on asymptotic analysis, was proposed on the Vlasov-Maxwell equations in the axisymmetric case [10]. The interest of such models is that they are much simpler to solve numerically than the full model.

Our aim here is to illustrate what we detailed in the previous section: the impossibility to get a precise error estimate for the finite element method, due to constants involved in the inequality of control (28)-(29) in theorem 3.2.

More precisely, using the FreeFem++ package [8], we approximate the variational formulation associated to the continuous problem, (for more details see [3]), with first a P_1, then a P_2 Lagrange finite element method.

We then collected in a suitable database all the numerical results computed by these two families of finite elements. Note also that, since the two methods do not have the same degrees of freedom, we chose to keep in the database the common data, that is the quantities computed at the vertices of the mesh. This exclude, for the P_2 finite element, the quantities computed at the middle of the edges of the mesh.

Hence, the database we consider is composed by data computed at each time step t_n and for each space node (r_j, ζ_k) of the mesh[1]. More precisely, a given row of the database parametered by the discrete time-space coordinates (r_j, ζ_k, t_n) is informed by the approximations of all of the physical unknowns, (the cylindrical components of the velocity $\mathbf{V} = (v_r, v_\theta, v_z)$, the components of the electromagnetic field $\mathbf{E} = (E_r, E_\theta, E_z)$ and $\mathbf{B} = (B_r, B_\theta, B_z)$ and the relative component of the so-called "pseudo-fields" $\mathcal{E}$, the charge density ρ and the current density $\mathbf{J} = (J_r, J_\theta, J_z)$), computed at a given time step and mesh node.

Then, the set of variables which corresponds to the columns of the database we considered is listed below:

$$v_r^{(i)}, v_\theta^{(i)}, v_\zeta^{(i)}, E_r^{(i)}, E_z^{(i)}, B_z^{(i)}, \mathcal{E}_r^{(i)}, \mathcal{E}_\theta^{(i)}, J_r^{(i)}, J_\theta^{(i)}, J_\zeta^{(i)}, \rho^{(i)}, \quad (i = 1, 2)\,, \tag{44}$$

where the exponent i specifies that the approximations are computed by the P_i Lagrange finite element.

Considering all the 100 time steps of the simulation and the 1250 space nodes of the mesh Π_h, the database we treated was composed by 125000 rows and by the 24 variables listed in (44).

Our objective was to appreciate the difference of accuracy between the P_1 and P_2 approximate problem solutions. To this end, we had to determine if there locally exist rows in the database, characterized by their (r_j, ζ_k, t_n), such that the P_1 method would bring a better approximation or, at least, an equivalent one than the P_2 finite element.

[1] we use obvious notation in axisymmetric: (r, z) for the coordinates, with $\zeta := ct - z$, c being the speed-of-light and t the time.

For doing this, and since without the exact solution, we cannot determine what "better approximation" means, we focused our attention to identify subgroups in the database such that the numerical approximations computed by the two $P_i, (i = 1, 2)$, finite elements are with "*the same order*".

For the sake of illustration, let us restrict ourselves to the radial component of the electrical field $E_r^{(1)}$ and $E_r^{(2)}$ obtained from the P_1 and P_2 finite element method respectively. Then, we are looking for local features such that:

$$|E_r^{(1)}(r_j, \zeta_k, t_n) - E_{r,\text{exact}}(r_j, \zeta_k, t_n)| \simeq |E_r^{(2)}(r_j, \zeta_k, t_n) - E_{r,\text{exact}}(r_j, \zeta_k, t_n)|. \tag{45}$$

The situation described by (45) means that, in a certain sense (will be soon defined), $E_r^{(1)}$ and $E_r^{(2)}$ have the "same numerical order". To identify such situations, let us first define the notion of "same numerical order". For this purpose, we introduce a threshold α and a new qualitative binomial variable called "$P_1 vs P_2$" as follows :

$$P_1 vs P_2 \equiv \left| \begin{array}{ll} \textit{Same Order}, & \text{if } |E_r^{(2)} - E_r^{(1)}| \leq \alpha, \\ \textit{Different Order}, & \text{if not,} \end{array} \right. \tag{46}$$

where $\alpha = 0.65$ in our example corresponding to 5% of the maximum of the absolute difference between $E_r^{(1)}$ and $E_r^{(2)}$ found in all the database. In such a way, the variable "$P_1 vs P_2$", especially its value "*Same Order*", will allow us to detect and to characterize situations where relation (45) holds, if any.

The next step is to choose a so-called *target variable* (the variable to be explained) - in our example "$P_1 vs P_2$" - and to process *ad hoc* data mining techniques to qualify the two different categories of this *target* variable.

The first result we found confirmed our suspicion: the database contains a non negligible quantity of rows such that (45) is satisfied. Indeed, as shown in Table 1, almost 14% of the rows in the dataset are in this case.

	Count	Percent
Different Order	6522	86.3
Same Order	1035	13.7

Table 1: Respective proportions of the "$P_1 vs P_2$" categories.

Now, to qualify these 14% of rows, namely the cluster composed by the "*Same Order*" of the target variable "$P_1 vs P_2$", we processed two data mining techniques. The first one we processed is called Kohonen' cards [9] and the second one, the decision trees [13].

5 Numerical Results

Let us detail now the study of *the discretization error* between $E_r^{(1)}$ and $E_r^{(2)}$ performed with data mining techniques. We are interested in qualifying the cluster composed by the "*Same Order*" of the target variable "$P_1 vs P_2$". The idea of our exploration process is to begin with the constitution of homogenous groups, called clusters, in the dataset, and this, without giving to the algorithm of Kohonen' cards the knowledge of the binomial variable "$P_1 vs P_2$" for each row of the dataset.

Therefore, when the clusters have been discovered by the Kohonen' cards, we consider again for all

of the rows which define each cluster, the value of our target variable "$P_1 vs P_2$".

The result we got is very interesting: we found, *a posteriori*, four very homogeneous clusters; two relative to the category "*Same Order*" and two relative to the category "*Different Order*".

As shown in Table 2, more than 90% of the elements of each cluster corresponds to one of the two category of the target variable "$P_1 vs P_2$".

	Group 1	Group 2	Group 3	Group 4
Different Order	93.7%	91.5%	6.5%	5.6%
Same Order	6.3%	8.5%	93.5%	94.4%

Table 2: Kohonen' clusters *versus* the target variable $P_1 vs P_2$.

We then gathered Group 1 and Group 2 as the cluster of the "*Different order*" and Group 3 and Group 4 as the cluster of the "*Same order*". So, the challenge is now to discover the rules of these two groups. This was processed by the help of a decision tree, also computed under the software *Modeler* of *IBM SPSS Inc.* Results are shown below in Figure 2 and correspond to the top part of the tree. The first segmentation which appears on the decision tree highlights the most discrim-

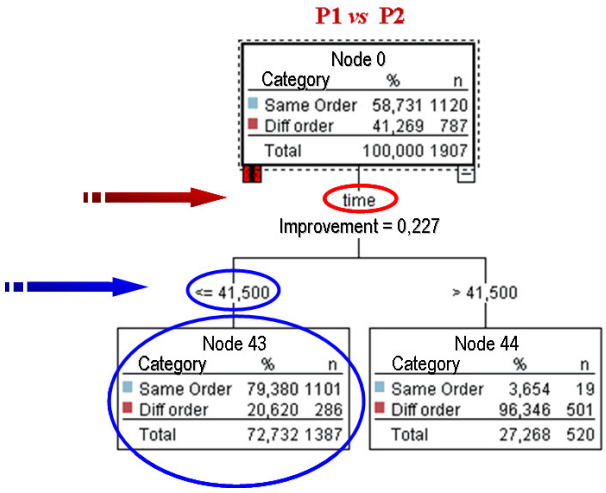

Figure 2: Decision Tree.

inated predictor variable, in the set of all the available potential predictors in the dataset. One can observe that the time is detected as this predictor, with a corresponding computed optimal threshold equal to the 42^{th} time steps over a hundred computed.

This means that when the time is smaller than the 42^{th} time step, the corresponding node 21 in the decision tree is very homogeneous regarding the value "*Same Order*". More precisely, almost 80% of the elements of this node have this value for the target variable "$P_1 vs P_2$".

We can then conclude that for time steps smaller than the 42^{th}, the implementation of P_2 finite elements was over qualified and then, the cost of the computation could not be justified anymore.

At the opposite, when one considers time steps which are greater than the 42^{th}, the corresponding cluster, which is the other node in the decision tree, (see Figure 2), is constituted by more than 98% of rows which correspond to the value "*Different Order*" of the target variable "$P_1 vs P_2$".

Unfortunately, we are not in position to give any conclusion in this case. Indeed, the criteria we choose to define the value "*Different Order*" does not allowed us to make the difference when the finite elements P_1 will lead to a better result in comparison with the P_2 ones.

This is one of the open question we have to treat, certainly by the help of above data mining techniques we already proceeded. Of course, for a more extensive study, the conclusion we got here on *the discretization error* needs to be confirmed on the other variables introduced in (44).

In other words, if data mining techniques allow us to identify critical situations where P_1 and P_2 finite elements have an equivalent accuracy, one has to observe that, until now, we are not in position to qualify if there exist situations where P_1 finite elements might be more accurate than the P_2 ones. This is one of the weaknesses of the method which is essentially due to the lack experimental results which constitute our future investigations.

More generally, we have to keep in mind that conclusions are also to be guided by the human expertise, that has to be involved in the data exploration. It also basically depends on the purpose of the performed analysis. Nevertheless, this pilot study illustrates how data mining can actually help to evaluate approximation errors for partial differential equations.

In addition, this novel even albeit heuristic approach, would help in the other sources of errors - modeling, approximation or parametrization error - to better understand and control the features of a given system in which simulations take a central part.

Beyond the particular case we treated in this paper, we suggest that data mining techniques can be applied to the analysis of any scientific computations as it is applied in a lot of other domains. Indeed, data mining would help to investigate the relevance and/or the quality of numerical simulations, particularly when a large quantity of data - as for instance with massively parallel computers - is available.

Nevertheless, several problems remain open: up to now, this approach remains heuristic and mathematical analysis to asses the method would be useful.

6 Conclusion

In this paper, we proposed a new approach based on data mining methods to evaluate error estimates of partial differential equations. Based on the fact that error estimates are derived from Taylor's expansions where unknown points appear, these estimations cannot not be precise to totally compare two numerical methods; Indeed, error estimates contain unknown constants which do not allowed to exactly evaluate the performances of the numerical considered schemes.

To specify this point, we considered as an example the finite elements method and we highlight how the unknown point introduced in Taylor's expansion causes the appearance of unknown constant in error estimates.

Then, as a pilot study, we considered what we called the *discretization error* and we restrict ourselves to a particular case: the Vlasov-Maxwell equations in a paraxial configuration. For this case, we constructed the related database made of the numerical results obtained by a P_1 and P_2 finite element method coupled with a particle method.

Therefore, based on two data mining techniques (Kohonen' cards and Decision Tree), we identified the time as the most discriminate predictor which describes the elements such that the evaluation between the two finite elements gave similar numerical results. An example of consequence is that, when the time is smaller than the identified threshold time step, both of the finite elements P_1 and

P_2 have similar results, and P_2 is overqualified

We think that, except this pilot study, this novel approach, would help to better understand and control the other sources of error and their eventual coupling, for any given system in which simulations take a central part.

References

[1] F. Assous, J. Chaskalovic, Data mining techniques for scientific computing: Application to asymptotic paraxial approximations to model ultra-relativistic particles, J. Comput. Phys., **230**, pp. 4811–4827 (2011).

[2] F. Assous, J. Chaskalovic, Error estimate evaluation in numerical approximations of partial differential equations: A pilot study using data mining methods, *C. R. Mecanique* **341** (2013) 304–313.

[3] F. Assous, J. Chaskalovic, Data mining as a tool to evaluate approximation errors for partial differential equations, (*submitted to JCP*) (2013).

[4] F. Assous, F. Tsipis, A PIC Method for Solving a Paraxial Model of Highly Relativistic Beams, *J. Comput. Appl. Math* **227-1**, pp. 136-146, 2009 (2008).

[5] F. Assous, F. Tsipis, Numerical paraxial approximation for highly relativistic beams, *Comput. Phys. Comm.* **180** 1086–1097, (2009).

[6] J. Chaskalovic, Mathematical and numerical methods for partial differential equations, Springer Verlag, 2013.

[7] R. Courant and F. John, Introduction to Calculus and Analysis, Vol. II/1 (Classics in Mathematics), Springer Verlag, 1989.

[8] F. Hecht, FreeFem++, Numerical Mathematics and Scientific Computation 3.7, Laboratoire J. L. Lions, Université Pierre et Marie Curie, http://www.freefem.org/ff++/, 2010.

[9] Kohonen, T., Self-Organizing Maps, 3rd ed., Vol. 1, Springer Series in Information Sciences, Vol. 30, Springer, Berlin, 2001.

[10] G. Laval, S. Mas-Gallic, P.-A. Raviart, Paraxial approximation of ultrarelativistic intense beams, *Numer. Math.*, **69(1)**, 33–60 (1994).

[11] B. Lucquin and O. Pironneau, Introduction to Scientific Computing , Wiley, 2001.

[12] P.A. Raviart et J.M. Thomas, Introduction à l'analyse numérique des équations aux dérivées partielles, Masson 1992.

[13] L. Rokach and O. Maimon, *Data Mining with Decision Trees: Theory and Applications*, World Scientific Publishing Company , 2001.