A posteriori error estimation via differences of numerical solutions

Alekseev A.K.^[0000-0001-8317-8688], Bondarev A.E.^[0000-0003-3681-5212] and Kuvshinnikov A.E.^[0000-0003-1667-6307]

Keldysh Institute of Applied Mathematics RAS, Moscow, Russia aleksey.k.alekseev@gmail.com, bond@keldysh.ru, kuvsh90@yandex.ru

Abstract. In this work we address the problem of the estimation of the approximation error that arise at a discretization of the partial differential equations. For this we take advantage of the ensemble of numerical solutions obtained by independent numerical algorithms. To obtain the approximation error, the differences between numerical solutions are treated in the frame of the Inverse Problem that is posed in the variational statement with the zero order regularization. In this work we analyse the ensemble of numerical results that is obtained by five OpenFOAM solvers for the inviscid compressible flow around a cone at zero angle of attack. We present the comparison of approximation errors that are obtained by the Inverse Problem, and the exact error that is computed as the difference of numerical solutions and a high precision solution.

Keywords: approximation error, ensemble of numerical solutions, differences of solutions, Inverse Problem, Euler equations, flow around a cone, Open-FOAM.

1 Introduction

An estimation of the approximation (discretization) error is a problem of the high current interest due to the need for the verification of software and numerical solutions [1,2,3]. Let us consider the system of the partial differential equations written in the operator form $A\tilde{u} = f$ and a discrete algorithm $A_h u_h = f_h$ that approximates the system on some grid. There are two main approaches to the estimation of the approximation error $\Delta u = u_h - \tilde{u}$ [3]. A priori error estimation has the appearance $||\Delta u|| \leq Ch^n$ (*h* is the step of discretization over space (or time), *n* is the order of approximation, *C* is an unknown constant). It is commonly used at a design and the theoretical analysis of the finite-difference or finite element algorithms (mainly from the standpoint of the convergence order determination). Unfortunately, the approach cannot be used in applications due to an unknown constant. A posteriori error estimation has the form $||\Delta u|| \leq E(u_h)$ and contains some computable error indicator $E(u_h)$ that depends on the previously computed numerical solution u_h . The approach has no unknown constants and can be easily applied to practical computations. Rather often, it has a non-intrusive form of certain postprocessor. The highly efficient technique is

developed for a posteriori error estimation in the domain of the finite-element analysis [3]. Unfortunately, for the problems of Computational Fluid Dynamics, the progress in a posteriori error estimation is limited so far due to irregularity of solutions (shock waves, contact surfaces etc). As the main tool for the verification, the standards [1,2] recommend the Richardson extrapolation. This method provides the pointwise approximation of the error field, however, at the cost of the extremely high computational burden [6]. There exist some computationally efficient approaches for the approximation error norm estimation, for example [4,5]. However, these approaches do not provide the pointwise information on the error.

Thus, the need for a computationally inexpensive a posteriori estimation of approximation error still exists. By this reason we consider herein the computationally cheap approach to a posteriori error estimation that is based on the ensemble of numerical solutions obtained by independent methods. By independent methods we mean the numerical algorithms with different computational properties (an inner structure or the order of the approximation). The approximation error is estimated in the pointwise approach using the difference of solutions at every grid node separately. For this purpose, the Inverse Ill-posed Problem is posed in the variational statement that includes the Tikhonov zero order regularization [7,8]. We provide the results of the numerical tests for compressible Euler equations that demonstrate both the estimated error and the exact error (obtained by the comparison of numerical solution with the precise solution [9]). These results demonstrate the applicability of the considered approach.

2 The estimation of approximation error using the differences of numerical solutions

The approximation error estimation may be performed by different methods [3-6] including the Richardson extrapolation [6] that uses the set of numerical solutions obtained *for consequently refined grids*. In contrast, we consider an ensemble of numerical solutions $u^{(i)} = \tilde{u} + \Delta u^{(i)}$ (i = 1...n), obtained by *n* independent algorithms *on the same grid*. Let_us note the exact (unknown) solution as \tilde{u} and the approximation error (unknown) for *i*-th solution as $\Delta u^{(i)}$. The differences of numerical solutions $d_{ij} = u^{(i)} - u^{(j)} = \tilde{u} + \Delta u^{(i)} - \tilde{u} - \Delta u^{(j)} = \Delta u^{(i)} - \Delta u^{(j)}$ are computable. These differences are equal to the differences of errors and, hence, contain some information regarding the unknown approximation errors $\Delta u^{(i)}$. We get $N = \frac{n(n-1)}{2}$ equations that relate unknown errors and computable differences of numerical solutions

$$D_{ii}\Delta u^{(j)} = f_i , \qquad (1)$$

where D_{ij} is a rectangular $N \times n$ matrix, the summation over the repeating index is implied starting from this point of the presentation.

The approximation errors may be estimated as

$$\Delta u^{(j)} = (D_{ij})^{-1} f_i \,. \tag{2}$$

Formally, this system of equations may be resolved for the dimensionality n that is equal or greater three. For n = 3 we use the option that follows:

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} d_{12} \\ d_{13} \\ d_{23} \end{pmatrix} = \begin{pmatrix} \Delta u^{(1)} - \Delta u^{(2)} \\ \Delta u^{(1)} - \Delta u^{(3)} \\ \Delta u^{(2)} - \Delta u^{(3)} \end{pmatrix} = \begin{pmatrix} u^{(1)} - u^{(2)} \\ u^{(1)} - u^{(3)} \\ u^{(2)} - u^{(3)} \end{pmatrix}$$
(3)

The equation (1) for this case has the form

$$\begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \Delta u^{(1)} \\ \Delta u^{(2)} \\ \Delta u^{(3)} \end{pmatrix} = \begin{pmatrix} u^{(1)} - u^{(2)} \\ u^{(1)} - u^{(3)} \\ u^{(2)} - u^{(3)} \end{pmatrix}$$
(4)

We numerically solve Equation (1) for three variables in the form (4) and for five variables in the similar form, omitted for brevity, by the method considered in following section.

3 Inverse Problem for the estimation of approximation error

One may easy see the solution of system (1) to be invariant relatively a shift transformation $\Delta u^{(j)} = \Delta \tilde{u}^{(j)} + b \ (\Delta \tilde{u}^{(j)})$ is the exact error) for any $b \in (-\infty, \infty)$. It is caused by the usage of the difference of solutions as the input data. Thus, the problem $\Delta u^{(j)} = (D_{ij})^{-1} f_i$ is underdetermined and, hence, ill-posed. The steady and bounded solution of the ill-posed problems requires a regularization ([7,8]) that we consider as the zero order Tikhonov one by the following reasons. It is natural to search for solutions of the minimum shift error |b| (in an ideal, $|b| \rightarrow 0$). We consider the search for the minimal L₂ norm of $\Delta u^{(j)}$ that provides restrictions on the absolute value of b:

$$\min(\delta) = \min \sum_{j=1}^{n} (\Delta u^{(j)})^2 / 2 = \min \sum_{j=1}^{n} (\Delta \widetilde{u}^{(j)} + b)^2 / 2$$
 (5)

We consider this expression as the regularizing term in further analysis.

By accounting

$$\Delta\delta(b) = \sum_{j}^{n} (\Delta \widetilde{u}^{(j)} + b) \Delta b, \qquad (6)$$

one may see that the minimum over b occurs at b that equals the mean error with the opposite sign:

$$b = -\frac{1}{n} \sum_{j}^{n} \Delta \tilde{u}^{(j)} = -\Delta \bar{u} .$$
⁽⁷⁾

So, the relation (5) addresses the deviation of the exact error from the mean $\Delta u^{(j)} = \Delta \tilde{u}^{(j)} - \Delta \bar{u}$ and corresponds the minimum of the error dispersion. Fortunately, the assumption of δ minimality (5) ensures the boundedness of *b*. Unfortunately, the maximum attainable accuracy of $\Delta u^{(j)}$ is restricted by the mean error value.

In accordance with [8] we pose the Inverse Problem for $\Delta u^{(j)}$ estimation in the variational statement that assumes both the equation (1) to be valid and $\|\Delta u\|_{L_2}$ to be minimal (Eq. 5). This statement implies the minimization of the functional:

$$\varepsilon(\Delta u) = 1/2(D_{ij}\Delta u^{(j)} - f_i) \cdot (D_{ik}\Delta u^{(k)} - f_i) + \alpha/2(\Delta u^{(j)}E_{jk}\Delta u^{(k)}).$$
(8)

Herein, the first term of (8) defines the discrepancy of the predictions and observations, the second term of (8) poses the zero order Tikhonov regularization, α is the regularization parameter, E_{jk} is the unite matrix. Eq. (8) demonstrates the standard statement of the variational Inverse Problem [8] that unify the observations and the a priori information (in present case, the information regarding the boundedness of solution).

To obtain the solution that minimize the functional we use the gradient based approach:

$$\Delta u^{(j),m+1} = \Delta u^{(j),m} - \tau \nabla \varepsilon \,. \tag{9}$$

In this expression *m* is the number of the iteration. In the numerical tests that follow the gradient is obtained by the direct numerical differentiation. The iterations stop at small enough value of the functional $\varepsilon \leq \varepsilon_*$ (in these tests $\varepsilon_* = 10^{-8}$). This optimization method is very primitive one (although quite operational) and easily may be replaced by any more advanced algorithm.

The solution of Eq. (8) is dependent on the regularization parameter α . The function $|\Delta u^{(j)}(\alpha)|$ is not bounded at $\alpha = 0$, so, it is not acceptable. At $\alpha \to \infty$ the asymptotics $|\Delta u^{(j)}(\alpha)| \to 0$ occurs that is not acceptable also. However, there exists a range of α where the weak dependence of the solution on α exhibits. In this range of the regularization parameter, the solution $\Delta u^{(j)}(\alpha)$ is close to the exact one $\Delta \tilde{u}^{(j)}$ and may be accepted [8].

4 The test problem

We consider the approximation error estimation for the test problem governed by two dimensional compressible Euler equations. The flowfield around cone at zero angle of attack $\alpha = 0^{\circ}$ (Fig. 1) in the uniform supersonic flow of ideal gas is analyzed. The test problem statement is very close to one described by [5] from the gasdynamics view-point, and the same solvers are used.

The precise solution by [9] is used for estimation of the exact error.



Fig. 1. Flow scheme

The cone of the half angle $\beta = 20^{\circ}$ is considered with the Mach number 2.

5 **OpenFOAM solvers**

The following solvers from the OpenFOAM software package [10] were used similarly to [5]:

• *rhoCentralFoam* (rCF), based on a central-upwind scheme, which is a combination of central-difference and upwind schemes [11,12]. The essence of the central-upwind schemes consists in a special choice of a control volume containing two types of domains: around the boundary points, the first type; around the center point, the second type. The boundaries of the control volumes of the first type are determined by means of local propagation velocities. The advantage of these schemes is the possibility to achieve a good resolution for discontinuous solutions in gas dynamics, using the appropriate technique for the numerical viscosity reducing.

- sonicFoam (sF), based on the PISO algorithm (Pressure Implicit with Splitting of Operator) [13]. The basic idea of the method is the application of two difference equations to calculate the pressure for the correction of the pressure field obtained from discrete analogs of the equations of moments and continuity. This approach takes to account that the velocities changed by the first correction may not satisfy the continuity equation, therefore, a second corrector is introduced, which enables us to calculate the velocities and pressures satisfying the linearized equations of momentum and continuity.
- *rhoPimpleFoam* (rPF), based on the PIMPLE algorithm, which is a combination of the PISO and SIMPLE (Semi-Implicit Method for Pressure-Linked Equations) algorithms. In this method, an external loop is added to the PISO algorithm, due to which the method becomes iterative one and allows to count with the Courant number greater than 1.
- *pisoCentralFoam* (pCF), which is a combination of a Kurganov-Tadmor scheme [11] with the PISO algorithm [14].
- *QGDFoam* (QGDF), which is based on the implementation of quasi-gas dynamic equations [15].

All these solvers have the same approximation order, however, their inner structure is quite different. This circumstance engenders the hope that the exact approximation errors $\Delta \tilde{u}^{(j)}$ are independent and the mean error $\Delta \bar{u} = \frac{1}{n} \sum_{j=1}^{n} \Delta \tilde{u}^{(j)}$ has the acceptable magnitude and decays as the ensemble of solutions expands.

6 Numerical results

6.1 Initial and boundary conditions

The computational domain and boundaries are provided in Fig. 2 similarly to [5].



Fig. 2. Computational domain and boundaries

ICCS Camera Ready Version 2020

To cite this paper please use the final published version: DOI: 10.1007/978-3-030-50436-6_37

6

On the upper boundary ("top") the zero gradient condition for the gas dynamic functions is specified. The same conditions are set on the right border ("outlet"). On the left border ("inlet"), the inflow parameters are set for Mach number M = 2: pressure P = 101325 Pa, temperature T = 300 K, speed U =694.5 m/s. On the cone surface, the condition of zero normal gradient is posed for the pressure and the temperature, and the condition "slip" is posed for the speed, corresponding to the non-penetration condition.

The special "wedge" condition is used for the front ("front") and back ("back") borders to model the axisymmetric geometry in the OpenFoam package. The Open-Foam package also employs the special "empty" boundary condition. This condition is specified in cases when calculations in the given direction are not carried out. In our case, this condition is used for the "axis" border.

6.2 Solvers settings

In the OpenFOAM package, there are two options for approximating differential operators: directly in the solver's code or using the fvSchemes and fvSolution configuration files. In order the comparison to be correct, we used the same parameters, where possible. The parameters are the same as in [5]. In the fvSchemes file: ddtSchemes – *Euler*, gradSchemes – *Gauss linear*, divSchemes – *Gauss linear*, laplacianSchemes – *Gauss linear corrected*, interpolationSchemes– *vanLeer*. In the fvSolution file: solver – *smoothSolver*, smoother *symGaussSeidel*, tolerance – 1e-09, nCorrectors – 2, nNonOrthogonalCorrectors – 1.

6.3 The results of the approximation error estimation in comparison with the exact error

For the estimation of approximation error, we minimize the functional (8) using iterations described by the Expression (9) for all flow parameters { ρ , u, v, p} separately at every grid point. Below, we present the results mainly for pressure, as the most important and expressive gas-dynamic function for inviscid flows with shock waves.

Fig. 3 presents the pressure error estimation $\Delta p^{(IP)}$ obtained by the Inverse Problem for the numerical solution computed by rCF [11,12].

Fig. 4 presents the "exact error" for the pressure calculated as the difference of numerical solution, computed by rCF, and the precise solution by [9] $\Delta p^{(rCF)} = p^{(rCF)} - p^{(precise)}$.

The comparison of Figs. 3 and 4 demonstrate the satisfactory proximity of error estimates to the exact error.



Fig. 3. The error of the pressure estimated by the Inverse Problem (for the flowfield, computed by rCF)



Fig. 4. The exact error of the pressure (for the flowfield, computed by rCF)

The Fig. 5 presents the piece of vectorized grid function of total pressure error (computed by rCF) obtained by the Inverse Problem (using five solutions) in comparison with the exact error. The index along abscissa axis $i = N_x(k_x - 1) + m_y$ is de-

fined by indexes along $X(k_x)$ and $Y(m_y)$. The periodical jump of solution variables corresponds the transition through the shock wave. One may see that the error at the shock wave remains underestimated. This behavior is expected since the error at a shock tends to be singular at the mesh refinement. The scales in Figures 3 and 4 are chosen to be smaller if compare with Figure 5 in order to increase the visibility.



Fig. 5. The comparison of the pressure error (rCF), estimated by the Inverse Problem, with the exact error



Fig. 6. The comparison of the density error (rCF), estimated by the Inverse Problem, with the exact error. IP-3 corresponds 3 solutions, IP-5 - five solutions

The Fig. 6 presents the piece of vectorized grid function of density error (for the solution computed by rCF) obtained by Inverse Problem in comparison with the exact error for the ensembles of solutions containing three and five samples. Herein, IP-3 corresponds three solutions (pCF, QGDF, rCF), IP-5 corresponds all five solutions.

In accordance with [3], the global quality of a posteriori error estimation may be described by the *effectivity index* that equals the relation of the estimated error norm to the exact error norm:

$$I_{eff,k} = \left\| \Delta \vec{\rho}^{(k)} \right\| / \left\| \Delta \vec{\tilde{\rho}}^{(k)} \right\|.$$
(10)

In the contrast to above used pointwise (scalar) variables, the vectors $\Delta \vec{\rho}^{(k)}, \Delta \vec{\rho}^{(k)} \in \mathbb{R}^M$ (*M* is the number of grid nodes) in this relation denote the grid functions. Thus, the norms, herein, average the pointwise errors over the total flow-field.

To provide the reliability of the error estimation, this index should be greater the unit. On the other hand, the estimation should be not too pessimistic, so the value of the effectivity index should be not too great. According [3], the range $1 \le I_{eff} \le 3$ is acceptable for the finite elements in the domain of elliptic equations. The solutions for the supersonic Euler equations (of hyperbolic type), considered herein, are less smooth (contain shock waves), so the acceptable range of the effectivity index may be greater and should be determined in numerical experiments. In the applications, the error of some valuable functionals may be of the primal importance. So, additional

efforts should be made to estimate the bounds of the effectivity index from the viewpoint of the error of valuable functionals. However, this subject is far away from the scope of present paper. Some information on the relation of the approximation error and the uncertainty of valuable functionals may be found in [16].

The corresponding values of the effectivity index are provided in the Table 1 for the sets of 3 and 5 solutions. L_2 norm of the global error (averaged over the total grid) is computed for the effectivity index calculation. To ensure the reliable error estimation ($I_{eff} \ge 1$) certain safety coefficient (about three) may be used.

One may expect the improvement of the error estimation quality as the number of solutions increases. This assumption is partly supported by the comparison of the effectivity index values for three and five solutions presented in the Table 1.

	I_{eff}^{pCF}	I_{eff}^{QGDF}	I_{eff}^{rCF}	I_{eff}^{rPF}	I_{eff}^{sF}
five solutions	0.35	0.37	0.38	0.39	0.26
three solutions	0.24	0.29	0.26	-	-

Table 1. Effectivity indexes of error estimation for different solutions

7 Discussion

Unlike the computationally cheap estimation of the error norm [4,5], also based on the set of numerical solutions, the above discussed approach provides the pointwise error estimation.

Certainly, this approach is less accurate if compared with the Richardson extrapolation due to the presence of the irremovable error, proportional to the mean error over the ensemble of solutions. Nevertheless, in contrast to the Richardson extrapolation, the considered postprocessor is much more computationally inexpensive since it implies only single grid computations (without a mesh refinement). Additionally, it has some natural parallelization properties since the solutions are obtained by the independent codes.

It should be noted that the calculations of the set of solutions for the flow around the cone were carried out using the approach of the generalized computational experiment [17,18]. This approach is based on the parallel solution (in the multitask mode) of the same problem for the ensemble of input data corresponding a variation of the determining parameters. This allows one to analyze a solution as the element of some ensemble engendered by the class of problems, which is set by the choice of determining parameters. The problem at hand can be considered as the generalized computational experiment, where the choice of a solver plays the role of a determining parameter.

The OpenFOAM algorithms are considered herein as independent ones from the standpoint of their logics and inner structure, however, the independence from the viewpoint of the mean error decay at the ensemble enhancement (important from the viewpoint of irremovable error diminishing) needs for a further analysis.

8 Conclusion

The Inverse Problem is stated for the estimation of the point-wise approximation error occurring at a discretization of the Partial Differential Equations. The differences between solutions, obtained by independent numerical algorithms, are used as the input data. The variational statement with the zero order regularizing term is considered for the problem.

The numerical tests demonstrate the feasibility for the estimation of the point-wise approximation error via the ensemble of numerical solutions obtained using the OpenFOAM software package.

The minimal number of numerical solutions that is necessary for the error estimation is equal to three. Five solutions provide a bit better results from the viewpoint of effectivity index.

The considered approach is less accurate in comparison with the Richardson extrapolation. However, the proposed approach is much more computationally cheap and ensures the satisfactory accuracy.

Acknowledgments

This work was supported by grant of Russian Science Foundation № 18-11-00215.

References

- Guide for the Verification and Validation of Computational Fluid Dynamics Simulations, American Institute of Aeronautics and Astronautics, AIAA-G-077-1998, Reston, VA, 1998.
- Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer, ASME V&V 20-2009, 2009
- Repin, S.I.: A posteriori estimates for partial differential equations. Vol. 4. Walter de Gruyter (2008). https://doi.org/10.1515/9783110203042
- Alekseev, A.K., Bondarev, A.E., Navon, I.M.: On Triangle Inequality Based Approximation Error Estimation. arXiv:1708.04604 [physics.comp-ph], August 16, 2017.
- Alekseev A.K., Bondarev A. E., Kuvshinnikov A. E.: Verification on the Ensemble of Independent Numerical Solutions, In: Rodrigues J. et al. (eds) Computational Science – ICCS 2019. ICCS 2019. Lecture Notes in Computer Science, Springer, Cham, 11540, 315–324 (2019).
- Alexeev, A.K., Bondarev, A.E.: On Some Features of Richardson Extrapolation for Compressible Inviscid Flows. Mathematica Montisnigri XL, 42–54 (2017).
- Tikhonov, A.N., Arsenin, V.Y.: Solutions of Ill-Posed Problems. Winston and Sons, Washington DC (1977).
- Alifanov, O.M., Artyukhin, E.A., Rumyantsev S.V.: Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems. Begell House (1995).
- Babenko, K.I., Voskresenskii, G.P., Lyubimov, A.N., Rusanov, V.V.: Three-dimensional ideal gas flow past smooth bodies. Nauka, Moscow (1964). (In Russian)

12

- 10. OpenFOAM, http://www.openfoam.org, last accessed 2020/01/30.
- Kurganov, A., Tadmor, E.: New high-resolution central schemes for nonlinear conservation laws and convection-diffusion equations. J. Comput. Phys. 160(1), 241–282 (2000). https://doi.org/10.1006/jcph.2000.6459
- Greenshields, C., Wellerr, H., Gasparini, L., Reese, J.: Implementation of semi-discrete, non-staggered central schemes in a colocated, polyhedral, finite volume framework, for high-speed viscous flows. Int. J. Numer. Meth. Fluids 63(1), 1–21 (2010). https://doi.org/10.1002/fld.2069
- Issa, R.: Solution of the implicit discretized fluid flow equations by operator splitting. J. Comput. Phys. 62(1), 40–65 (1986). https://doi.org/10.1016/0021-9991(86)90099-9
- Kraposhin, M., Bovtrikova, A., Strijhak, S.: Adaptation of Kurganov-Tadmor numerical scheme for applying in combination with the PISO method in numerical simulation of flows in a wide range of Mach numbers. Procedia Computer Science 66, 43–52 (2015). https://doi.org/10.1016/j.procs.2015.11.007
- Kraposhin, M.V., Smirnova, E.V., Elizarova, T.G., Istomina, M.A.: Development of a new OpenFOAM solver using regularized gas dynamic equations. Computers & Fluids, 166, 163–175 (2018). https://doi.org/10.1016/j.compfluid.2018.02.010
- Alekseev A.K., Bondarev A. E., Kuvshinnikov A. E., On Uncertainty Quantification via the Ensemble of Independent Numerical Solutions, Journal of Computational Science (2020), https://doi.org/10.1016/j.jocs.2020.101114
- 17. Bondarev, A.E.: On the Construction of the Generalized Numerical Experiment in Fluid Dynamics. Mathematica Montisnigri, XLII, 52–64 (2018).
- Bondarev, A.E., Galaktionov, V.A.: Generalized Computational Experiment and Visual Analysis of Multidimensional Data. Scientific Visualization 11(4), 102–114 (2019). https://doi.org/10.26583/sv.11.4.09