HORIZON-EUROHPC-JU-2021-COE-01

Centre of Excellence in Exascale CFD

CEEC – Centre of Excellence in Exascale CFD
Grant Agreement Number: 101093393

D3.2 – Stable adjoints

WP3: Exascale algorithms

Document Information

Deliverable Number	D3.2
Deliverable Name	Stable adjoints
Due Date	31/12/2023 (PM 12)
Deliverable lead	DTU
Authors	Martin Berggren (UmU), Philipp Schlatter (FAU), Niels Aage (DTU), Roman Iakymchuk (UmU)
Responsible Author	Niels Aage (DTU) naage@dtu.dk
Keywords	adjoint solver, computation fluid dynamics, static mixer
WP	WP3
Nature	R
Dissemination Level	PU
Final Version Date	31/12/2023
Reviewed by	Ananias Tomboulides (AUTH) and Niclas Jansson (KTH)
MGT Board Approval	31/12/2023

Acknowledgment:

Funded by the European Union. This work has received funding from the European High Performance Computing Joint Undertaking (JU) and Sweden, Germany, Spain, Greece, and Denmark under grant agreement No 101093393.

Document History

Partner	Date	Comment	Version
UmU, DTU, FAU	11/10/23	Initial version: defining the scope	0.1
UmU, DTU, FAU	02/11/23	Background section	0.2
UmU, DTU, FAU	16/12/23	Initial version of strategies	0.3
UmU, DTU, FAU	29/12/23	Updated and extended strategies	0.4
DTU	01/12/23	Version ready for internal review	0.5
DTU, UmU, FAU	20/12/23	Incorporated reviewers' comments	0.6
DTU	20/12/23	Final version	1.0

Executive Summary

This document is the second deliverable of work package 3 – 'Exascale Algorithms' of the EuroHPC JU Center of Excellence in Exascale CFD (CEEC). This work package is concerned with all efforts required for improving algorithms in the CEEC codes for efficient exploitation of Exascale architectures. In particular, we focus on improving scalability and numerical properties of cornerstone algorithms in computational fluid dynamics (CFD) codes; providing mixed-precision algorithmic solutions for better energy footprint and/ or faster execution; enabling error control and assure robustness of algorithms; facilitating Exascale design optimization.

This deliverable 'D3.2 – Stable adjoints' provides a report on the known issues and challenges associated with exponentially growing solution of the temporal, linearised adjoint problem, arising as part of the sensitivity calculation for the topology optimisation of statics mixers in lighthouse case 3 (LHC 3). This report also contains an overview of possible approaches for alleviating the exponential growth in the adjoint solution, including algorithmic considerations and recommendations for next steps.

Contents

1 Introduction **6**

2 Background **6**

3 Strategies **8**

4 Summary **9**

1 Introduction

The Center of Excellence in Exascale CFD (CEEC) implements Exascale ready workflows for addressing relevant challenges for future Exascale systems, including those procured by EuroHPC. The significant improvement in energy efficiency will be facilitated through efficient exploitation of accelerated hardware architectures (GPUs) and novel adaptive mixed-precision calculations. Emphasis is furthermore given to new or improved algorithms that are needed to exploit upcoming Exascale architectures. The efforts of the center are driven by a collection of six different lighthouse cases of physical and engineering interest, ranging from aeronautical to atmospheric flows. For more details on each lighthouse case, we refer to 'Deliverable D1.1 – CEEC Exascale lighthouse cases and their need' [4].

This deliverable is directly related to the development work necessary to execute lighthouse case 3 on the topology optimization of a static mixer. The basic idea of this case is to automatically find an optimal design of a mixing device, balancing the pressure drop (i.e. the flow resistance) and the mixing ability in an optimal way. This goal is achieved using topology optimization, a modern design method that enables the determination of completely new geometries (i.e. topologies) without the need of prescribing a-priori parameterized shape function. In this way, the optimized design is very general, but may still be constraint by certain external conditions (e.g. the size of the optimization domain, the amount of solid material available etc.).

Within lighthouse case 3, the topology optimization algorithm will be developed for the code Neko [9]. Therefore, all methods will need to be efficiently implemented on both CPU and GPU architectures, which may necessitate novel algorithms developed in WP3. The aspects to be discussed in the present deliverable D3.2 deal with problem of the calculation of sensitivities for the topology, typically calculated using an adjoint method. This problem is in principle general and independent of the chosen code or numerical implementation, therefore, the below description is not tied to Neko in particular.

The roadmap for both the lighthouse case 3 (topology optimization) and the community code Neko intended for this case has already been outlined in 'Deliverable D3.1 – Analysis of the CEEC codes and underlying solvers: Requirements and strategies definition' [5], Section 3.4.

This document is organised as follows: Section 2 presents the issue with exponentially growing adjoints from a physical/mathematical perspective. Section 3 presents a number of strategies that holds the potential to remedy the issue with exploding adjoints. Emphasis is made on algorithmic complexity and suitability with respect to reaching Exascale capabilities. Section 4 summarizes and ranks the different strategies, and provides recommendations for the path forward.

2 Background

Lighthouse case 3 addresses topology optimization of static mixers, in which the geometric layout of a static mixer will be designed using density-based topology optimization. also known as the material distribution method [1]. The method is defined by its extreme level of design freedom, which is obtained by the introduction of a continuous material indicator variable, or field, which, in its standard form, is used to interpolate between no material (void/fluid) and material (solid wall/no-slip condition). Hence, it allows for topological changes (how different parts are connected) during the optimization process and is thus much more versatile than e.g. shape optimization, where only the boundary shape of a given configuration is modified. The goal is in the end to obtain a material indicator field that only specifies either fluid or solid phases, which is made feasible by use of interpolation functions tailored for the given physics being considered. However, this type of field representation of the geometry also means that the resulting optimization problem is characterized by a very large number of design variables (*e.g.* millions) but at most a handful of objective functions and constraints. For computational efficiency, it is therefore necessary to employ gradient-based optimization algorithms together with an efficient computation of the necessary gradients of the objective and constraint functions.

Since in topology optimization, all grid points describe individual degrees of freedom of the design, typically the solution of the adjoint equations are used for computing the derivatives of the objective function and constraints [3]. The reason for the popularity of the adjoint method in design optimization can be made clear from a comparison to its alternatives. In fact, it is fair to claim that the success of the topology optimization approach is due to the existence and efficiency of adjoint sensitivity analysis. As example, both direct sensitivity analysis and the simpler finite difference approximation, results in the need to perform one additional function evaluation for each design variable (linearised problem for the direct approach and full non-linear solution for the finite difference approach). When dealing with millions of variables, such approaches clearly are intractable from any practical perspective. The adjoint approach, on the other hand, only requires the solution to a single additional linearised problem for each objective and constraint function being considered. This renders the adjoint approach almost independent of the number of design variables. Also note that the adjoint-equation approach is conceptually the same as the back-propagation algorithms in machine learning. The use of adjoint equations for topology optimization is a very well-established technique in general, but the technique comes with some specific challenges when using time-dependent scale-resolving approaches like direct and large-eddy simulation (DNS and LES) to CFD.

Consider the unsteady Navier–Stokes equations for the flow of an incompressible viscous fluid, for simplicity, started from a nonzero initial condition but without any other forcing or inhomogeneous boundary conditions to the system. The velocity field $\boldsymbol{u}$ then satisfies a quite strong temporal stability condition,

$$\frac{1}{2}\frac{d}{dt}\int_\Omega |\boldsymbol{u}|^2\,dV = -\nu\int_\Omega |\nabla\boldsymbol{u}|^2\,dV, \tag{1}$$

where $\nu > 0$ is the kinematic viscosity coefficient. Thus, the time derivative of the kinetic energy will always be negative, so the velocity field will die out eventually due to the dissipative effects of viscosity.

However, when computing the effects on the velocity field of a small change in, for instance, the geometry, we are led to solving a system of equations linearized around a base flow. If the base flow is denoted $\boldsymbol{u}$ and the perturbed field $\boldsymbol{u}'$, the stability condition becomes

$$\frac{1}{2}\frac{d}{dt}\int_\Omega |\boldsymbol{u}'|^2\,dV = -\nu\int_\Omega |\nabla\boldsymbol{u}'|^2\,dV - \int_\Omega \boldsymbol{u}'\cdot(\nabla\boldsymbol{u})\boldsymbol{u}'\,dV. \tag{2}$$

The situation is now completely different from condition (1); the linearized flow field is only guaranteed to die out if the integrand in the last integral of expression (2) is positive. Unfortunately, nothing can be said in general about the sign of this integral.

The same condition holds for the corresponding adjoint variable used to compute the derivatives of a given objective function or constraint. The linearized and adjoint equations both include a term containing the gradient of the base flow, which can cause exponentially growing solutions. Computational experience indicates that exponential growth in the adjoint equations generally happens for transitional and turbulent flows when computed using a DNS or LES approach. The common interpretation of this property is that the Navier–Stokes equations constitute a chaotic system. Thus, a small change in the design variables in a DNS or LES solution typically generates a trajectory that exponentially diverges from the undisturbed flow.

However, in most applications, it is the averaged performance of a system with respect to an ensemble of realizations of the chaotic system that is of interest. Thus, the sensitivity of an individual trajectory to a change in the design is likely irrelevant.

3 Strategies

To handle the issue of potential exponential growth in the adjoint equations for DNS and LES approaches, there are a number of options.

1. Keep the flow laminar or only weakly transitional to limit the growth rate of the solution to the adjoint equation. In practice, that would correspond to the study of a mixer at low Reynolds numbers (Re), which are likely to be reached with small outer dimensions, or low speeds. However, mixers are typically operated at higher Re to enhance the mixing activity.

2. Consider only small time windows for the adjoint integration, over which the exponential growth may not yet overshadow the useful sensitivity information. However, the lack of proper convergence in time may necessitate running a large number of ensembles to extract the sensitivity, thus this method has been termed *ensemble adjoint method*. However, it turns out that the number of ensembles for a turbulent flow is very high (millions) which makes the method impractical from a computational point of view.

3. Handle the exponential growth by simply scaling the adjoint field to prevent floating-point overflow, which is easy to do due to the linearity of the adjoint equation. The corresponding gradient expression will then be very large in unscaled units, dominated by the effects of the exponentially growing modes contained in the adjoint variable. The large size of the gradient will lead to a minuscule update of the design variables since the linearization will not be accurate otherwise. In fact, the update may easily be smaller than the precision of the floating-point system, rendering it useless. In addition, the dominating exponentially growing modes will not yield the most beneficial sensitivity for an averaged quantity of interest, since these modes are only associated with a particular realization.

4. Some very involved techniques to deal with the growth issue have been developed and studied, mostly for low-dimensional chaotic systems such as the Lorenz systems. One idea, denoted Least-Squares Shadowing [10], is to differentiate the objective function only with respect to perturbations that generate trajectories that stay close to the base flow. The technique is computationally very expensive and is likely infeasible for the application at hand. Moreover, there are also recent results questioning the whole idea behind shadowing for chaotic systems [2]. So-called cumulant truncation [6] is another very computationally expensive technique that

has been proposed for differentiating low-dimensional chaotic systems.

5. Adopting a two-step approach, where the optimization is carried out on a cruder lower-dimensional model whose parameters are set by a DNS or LES simulation. This model may constitute a solution to the RANS (Reynolds-averaged Navier–Stokes) equations with a specifically tuned turbulence model. The objective function is then evaluated on the cruder model, and the DNS/LES-given parameters are frozen during the sensitivity analysis. In this way, the design updates are based solely on an adjoint equation to the crude model. After the design update, however, the DNS/LES model is rerun with the updated design to adjust the parameters in the crude model. For turbulent cases, the crude model is likely a turbulence model, *e.g.* a (steady or unsteady) RANS equation with tuned eddy viscosity. The advantage of updating the coefficients of the RANS based on the DNS/LES of the forward problem is that the potential systematic differences (modelling errors) may be minimized by these adjusted parameters. The overall predictive accuracy is thus enhanced compared to only relying on RANS alone. Specific approaches from data assimilation may be used to optimize the relevant RANS parameters. However, it must be noted that the linear systems arising from RANS-based adjoints are prone to poor conditioning, which significantly deteriorates the performance of iterative solvers.

6. Similar to the previous approach, successful applications of integrating the adjoint equation about regularized base flows have been reported [8]. Methods may include a lower resolution in both space and time of the base flow (potentially avoiding some of the exponential instabilities). Another alternative, yet unexplored, is the construction of a low-order model of the forward solution, *e.g.* based on proper orthogonal decomposition (POD). By properly selecting the modes, the small-scale sensitivity may also be avoided.

7. The addition of specific damping terms in the adjoint equations have also been proposed. In this method, the growth of the adjoint solution is monitored, and coefficients related to sink terms are dynamically adjusted to avoid energy growth [7].

It is likely that the work conducted as part of WP3 may result in alternatives to the above mentioned approaches, and hence, we do not limit our selves to these. Instead they merely act as a starting point for further algorithmic development.

4 Summary

For the case of the turbulent static mixer, it appears that only strategies 5, 6, and 7 are potential solutions. In particular, strategy 5, that is, to calculate the sensitivities based on an optimized RANS-based solution, should clearly constitute the initial step. The DNS will in this case be run using the Neko code. For the RANS-based optimization and the data assimilation for obtaining adapted RANS coefficients, either an approach within Neko or external RANS tools, e.g. Nek5000 or SU2, may be considered. On the other hand, the apparent simplicity of option 6 (regularized base flow) is striking, and thus a comparison with adjoints based on such simplified base flow is certainly instructive. Similarly, option 7 is also computationally effective as no other codes or equations need to be solved, however, it is unclear from the current literature how such a damping term is formulated. We intend to test several possible formulations, and compare the efficiency and accuracy.

Bibliography

[1] Martin P. Bendsøe and Ole Sigmund. *Topology Optimization: Theory, Methods and Applications*. Engineering online library. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[2] N. Chandramoorthy and Q. Wang. On the probability of finding nonphysical solutions through shadowing. *J. Comput. Phys.*, 440:110389, 2021.

[3] Peter W Christensen and Anders Klarbring. *An Introduction to Structural Optimization*, volume 153. 2009.

[4] CEEC Consortium. Deliverable D1.1 – CEEC exascale lighthouse cases and their needs, August 2023. Available via `https://ceec-coe.eu/`.

[5] CEEC Consortium. Deliverable D3.1 – Analysis of the CEEC codes and underlying solvers: Requirements and strategies definition, August 2023. Available via `https://ceec-coe.eu/`.

[6] J. Craske. Adjoint sensitivity analysis of chaotic systems using cumulant truncation. *Chaos Solitons Fractals*, 119:243–254, 2019.

[7] A. Garai and S. M. Murman. Stabilization of the adjoint for turbulent flows. *AIAA Journal*, 2021. https://doi.org/10.2514/1.J059998.

[8] J. Hoffman, J. Jansson, N. Jansson, and R. V. De Abreu. Towards a parameter-free method for high Reynolds number turbulent flow simulation based on adaptive finite element approximation. *Computer Methods in Applied Mechanics and Engineering*, 288:60–74, 2015.

[9] Niclas Jansson, Martin Karp, Artur Podobas, Stefano Markidis, and Philipp Schlatter. Neko: A modern, portable, and scalable framework for high-fidelity computational fluid dynamics. *ArXiv*, 2107.01243, 2021.

[10] Q. Wang, R. Hu, and P. Blonigan. Least squares shadowing sensitivity analysis of chaotic limit cycle oscillations. *J. Comput. Phys.*, 267:210–224, 2014.