Variational Multiscale Stabilization of High-Order Spectral Elements for the Convection-Diffusion Equation

S. Marras\textsuperscript{a,}\textsuperscript{*}, J. F. Kelly\textsuperscript{b}, F. X. Giraldo\textsuperscript{b}, M. Vázquez\textsuperscript{a,c}

\textsuperscript{a}CASE, Barcelona Supercomputing Center (BSC-CNS) Barcelona, Spain
\textsuperscript{b}Applied Mathematics, Naval Postgraduate School. Monterey (CA), U.S.A.
\textsuperscript{c}IIIA-CSIC. Bellaterra, Spain

Abstract

One major issue in the accurate solution of convection-dominated problems by means of high-order methods is the ability of the solver to maintain monotonicity. This problem is critical for spectral elements, where Gibbs oscillations may pollute the solution. However, typical filter-based stabilization techniques used with spectral elements are not monotone. In this paper, residual-based stabilization methods originally derived for finite elements are constructed and applied to high-order spectral elements. In particular, we show that the use of the Variational Multiscale (VMS) method greatly improves the solution of the transport-diffusion equation by reducing over- and under-shoots, and can be therefore considered an alternative to filter-based schemes. We also combine these methods with discontinuity capturing schemes (DC) to suppress oscillations that may occur in proximity of boundaries or internal layers. Additional improvement in the solution is also obtained when a method that we call FOS (for First-Order Subcells) is

\textsuperscript{*}CASE, Barcelona Supercomputing Center
c. Jordi Girona 29, Barcelona 08034. Spain
simone.marras@bsc.es
Phone: +34 93 4017295
Fax: +34 93 4137721
One major issue in the accurate solution of convection-dominated problems by means of high-order methods is the ability of the solver to maintain monotonicity. This problem is critical for spectral elements, where Gibbs oscillations may pollute the solution. However, typical filter-based stabilization techniques used with spectral elements are not monotone. In this paper, residual-based stabilization methods originally derived for finite elements are constructed and applied to high-order spectral elements. In particular we show that the use of the Variational Multiscale (VMS) method greatly improves the solution of the transport-diffusion equation by reducing over- and under-shoots, and can be therefore considered an alternative to filter-based schemes. We also combine these methods with discontinuity capturing schemes (DC) to suppress oscillations that may occur in proximity of boundaries or internal layers. Additional improvement in the solution is also obtained when a method that we call FOS (for First-Order Subcells) is used in combination with VMS and DC. In the regions where discontinuities occur, FOS subdivides a spectral element of order p into p² subcells and then uses 1st-order basis functions and integration rules on every subcell of the element. The algorithms are assessed with the solution of classical steady and transient 1D, 2D, and pseudo-3D problems using spectral elements up to order 16.
used in combination with VMS and DC. In the regions where discontinuities occur, FOS subdivides a spectral element of order $p$ into $p^2$ subcells and then uses $1^{st}$-order basis functions and integration rules on every subcell of the element. The algorithms are assessed with the solution of classical steady and transient 1D, 2D, and pseudo-3D problems using spectral elements up to order 16.

*Keywords:* Spectral Elements; High-Order Methods; Variational Multiscale Stabilization (VMS); Advection-Diffusion; Numerical Weather Prediction; First Order Subcells (FOS) method; Monotonic Schemes
1. Introduction

A large number of physical applications relies on the accurate solution of the transport-diffusion equation

\[ \frac{\partial q}{\partial t} + L(q) = f, \]  

where \( q \) is the concentration of the tracer, \( L(q) = u \cdot \nabla q - \nabla \cdot (\nu \nabla q) \), \( \nu > 0 \) is a diffusion coefficient, \( u \) is a known velocity field, and \( f \) is a constant source term. The solution of (1) should respect two significant properties: (i) positivity must be preserved, and (ii) smearing at internal and boundary layers should not be excessive. These properties are extremely important in the context of transport in the atmosphere. Both limited-area and global atmospheric models for weather prediction need monotonic advection of tracers and moisture variables, otherwise the wrong amount of precipitation would be forecasted. Simple microphysics schemes, such as the Kessler parametrization [1], require three variables (water vapor, cloud water, and rain), whereas more sophisticated parameterizations include additional variables such as ice and snow [2]. Similarly, climate models require transport of hundreds of tracers, each representing a different chemical species. Regardless of the physical scales of the model, tracers must remain positive since the physical parameterizations that govern sub-grid scale processes such as auto-conversion and sedimentation, implicitly assume such a condition. These issues have been addressed for both transient and stationary problems (See, e.g., [3]) and, in the context of finite element methods, so-called stabilized methods have been an active topic of research since their introduction in the early 1980s with the Streamline-Upwind method of Hughes.
and Brooks [4]. In this paper we address the problem of solving (1) by high-order spectral element methods (SEM) without losing the ability to approach a monotone solution of the problem. Higher-order accuracy, in fact, comes at the price of aliasing phenomena in the solution [5], but the anti-aliasing filters typically used to give a stable spectral element solution do not respect conditions (i) and (ii) described above. Therefore, to achieve monotonic results with high-order spectral elements, we consider stabilization schemes originally devised for finite elements, and focus on techniques that can be derived directly from subgrid scale considerations as originally defined in [6] and [7] in the context of variational multiscale methods. These schemes assure stability by designing a diffusion-type term that is added to the Galerkin formulation of the original problem.

The first stabilized schemes based on the addition of a diffusive stabilization term to the Galerkin equation are the Artificial Viscosity methods (AV) [8] and the Streamline-Upwind method (SU) [4]. AV, as Hyper-viscosity (HV), is often used in atmospheric and ocean modeling due to the property of preserving the correct energy cascade in simulations that involve turbulence. The SU scheme uses the information in the direction of the flow to add viscosity only in the streamline direction. Both methods use a constant diffusion coefficient that does not typically change from element to element. A major improvement came by introducing the residual of the governing equation in the definition of the stabilization term. When the computed solution approaches the exact solution, the stabilization term should vanish. This strategy is known as residual weighting and generates a family of stabilization methods used mostly in FEM-based Computational Fluid Dynamics (CFD). These schemes, which are consistent in that the stabilization terms goes to zero as the numerical solution approaches the exact solution,
are considered in this paper. The most commonly used are the *Upwind-Streamline/Petrov-Galerkin* (SUPG) and the *Galerkin/Least-Squares* (GLS), devised in 1982 [9] and 1989 [10], respectively, as a consistent counterpart to SU. GLS was designed as a generalization of SUPG, but in the limit of pure advection, or for piece-wise linear elements, the GLS and SUPG methods are equivalent. Stability analysis for these two methods is detailed in [11, 12, 10]. The *Gradient Galerkin/Least-Squares* [13] for advection-diffusion with a reaction term, or the *Unusual Stabilized Finite Element Method* (USFEM) [14, 15] are a few examples. In the framework of high order methods, Petrov-Galerkin stabilization was applied by Pasquarelli and Quarteroni [16] to stabilize the convection-diffusion equation with the spectral method. Canuto used bubble functions to address the same issue [17] (See also [18, 19]).

The analyses of Hughes [6], Hughes and Stewart [20], and Hughes et al. [7] form the unifying theory of all stabilized finite element methods. According to this theory, stabilized methods are subgrid scale models where the unresolved scales are intimately related to the instabilities at the level of the resolved scales, and thus should be used in the construction of the stabilization term. These schemes are known as *Variational Multiscale (VMS)* methods. Details are given in subsection 2.2.2. VMS methods are all residual-based methods that improve the stability properties of the solution, and preserve the accuracy of the underlying numerical scheme [21]. However, Godunov’s theorem [22] implies that the latter property may be violated in the proximity of discontinuities or strong gradients. To the authors’ knowledge, the only application of VMS to spectral elements is the work of Wasberg et al. [23] in the context of large eddy simulation.
under-shoots in the proximity of sharp gradients of the solution. For this reason, *discontinuity capturing* (DC) techniques, also referred to as *Spurious Oscillations at Layers Diminishing* (SOLD) methods are used in combination with SUPG and VMS to introduce an additional term to the stabilized form of the equation. This issue was treated for the first time in [24], where details on how to build the stabilization parameter are also given, and in [25] for non-linear problems. A detailed review of most existing SOLD schemes can be found in a two-part paper by John and Knobloch [26, 27], where a modification of the *discontinuity-capturing* of Codina [28] is presented and is shown to be a promising option for FE solutions characterized by boundary layers.

All these methods strongly depend on a parameter that will be identified by $\tau$ throughout the paper. It will be also referred to as *intrinsic time*. A classical result for $\tau$ was obtained by Franca, Frey and Hughes in [29] by error analysis. Their result was reproduced by other authors using different approaches. Additional expressions for $\tau$ were found by Codina in [30, 31], by Codina, Oñate and Cervera in [32], by Harari and Hughes in [13], and by Shakib, Hughes and Johan in [33], who based the derivation on the (discrete) maximum principle. Another expression is due to Franca and Valentin [15] who based their derivation on convergence and stability analysis. Starting with the formalization of VMS methods by Hughes [6], $\tau$ has often been derived using Green’s functions, a thorough analysis of which is done by Hughes and Sangalli in [34]. Recently, Houzeaux, Eguzkitza and Vázquez [35] proposed a new way to derive the approximate subgrid scale solution, with results that are comparable to those of Hauke and García-Olivares in [36]. In [37], Codina builds $\tau$ using the Fourier analysis of the
problem; however, determining $\tau$ remains open. For this reason, we propose $\tau$ for higher-order spectral elements and use it to construct an appropriate stabilization method. To further improve the solution, we combine VMS and DC with a method that we here call FOS (for First-Order Subcells). This technique subdivides a tensor product spectral element of order $p$ and dimension $d$ into $p^d$ subcells, and then uses 1st-order basis functions and integration rules on every subcell of the element.

1.1. Main contribution of this paper

The main problem that we want to solve with the work presented here is that of stabilizing the spectral element solution of the advection-diffusion equation by sub-grid scale stabilization techniques (namely, VMS), and improve the solution by reducing the under- and overshoots that would occur if classical filters were to be used. The definition and implementation of $\tau$ in VMS stabilization may greatly affect the result. We hence adapted the method described in [35] to compute $\tau$, and applied it to spectral elements that use Legendre-Gauss-Lobatto (LGL) nodes, and show that this technique can be used in problems where spectral element filters fail. We finally apply FOS (i.e., we lower the order of interpolation by keeping the computational grid untouched) only where the solution is characterized by a propagating discontinuity. The combination of VMS, shock capturing and FOS methods will be shown to be an encouraging direction to take for constructing high-order positive-definite spectral element methods. To assess the algorithm, steady and transient advection-diffusion problems are solved on one- and two-dimensional domains using spectral elements up to order 16. We compare the performance of the method using VMS against those obtained with previous classical spectral element schemes.
1.2. Outline of this paper

The remainder of the paper is organized as follows. After this introduction, the numerical method and the corresponding stabilization are derived in Section 2. Mass conservation properties are reported in Section 3. Tests to verify the algorithm and a discussion of the results are presented in Sections 4 and 5, respectively.

2. Numerical methods

Given the space $L^2$ of real-valued functions that are square integrable in a bounded domain $\Omega \subset \mathbb{R}^2$ with boundary $\partial \Omega$, the Sobolev space $H^1$ of weakly-differentiable functions will be used. Specifically, $W \subseteq H^1$ represents the space of trial and basis functions of the Galerkin formulation to follow. In $L^2$ the inner product is given by $(\cdot, \cdot)$, and the 2-norm associated with the space is denoted by $\| \cdot \|_2$. For simplicity, we add the property that the solution $q$ vanishes on $\partial \Omega$; under this assumption, $W \subseteq H^1_0$. Given a finite element partition $\Omega^h = \bigcup_{i=1}^{n_{el}} K_i$ of the computational domain $\Omega$ into $n_{el}$ high-order conforming quadrilaterals $K_i$ of characteristic length $h$, $W^h$ is the finite dimensional projection of $W$. The discrete weak form reduces to the problem of finding the function $q^h \in (W^h; 0, t)$ such that

$$\left( \psi^h, \frac{\partial q^h}{\partial t} \right) + a(\psi^h, q^h) = (\psi^h, f) \quad \forall \psi^h \in W^h \quad (2)$$

where $a(\cdot, \cdot)$ is a bilinear form that satisfies

$$a(\psi, q) = (\psi, \mathcal{L}(q)).$$
After integrating by parts and assuming homogeneous Dirichlet boundary conditions, we have that

$$
\left( \psi^h, \partial_\psi^h \frac{\partial q^h}{\partial t} \right) = \int_{\Omega^h} \psi^h \frac{\partial q^h}{\partial t} \, d\Omega^h,
$$

$$
a(\psi^h, q^h) = \int_{\Omega^h} \psi^h u \cdot \nabla q^h \, d\Omega^h + \int_{\Omega^h} \nu \nabla \psi^h \cdot \nabla q^h \, d\Omega^h,
$$

$$
(\psi^h, f) = \int_{\Omega^h} \psi^h f \, d\Omega^h = 0.
$$

Remark 1: If advection dominates diffusion, unless $h$ is sufficiently small or the exact solution is globally smooth, the Galerkin approximation expressed by (2) is such that $q^h$ will suffer from severe/unacceptable oscillations [11, 38]. Furthermore, if the discretization relies on high-order methods, Gibbs oscillations may occur regardless of the size of the grid. Different ways to improve stability will be described in the following sections.

2.1. The spectral element method

Problem (2) is solved on a grid of quadrilateral elements of order $p$, where the element-wise solution $q^h$ is approximated by the expansion $\sum_{k=1}^{N_p} \psi_k(x) q_k^h(t)$ on $N_p = (p + 1)^2$ collocation points within the element. The expansion
functions $\psi_k$ are constructed as the tensor product from the Lagrange polynomials $h_i(\xi(x))$ and $h_j(\eta(x))$ of order $p$ as:

$$\psi_k(x) = h_i(\xi(x)) \otimes h_j(\eta(x)), \quad \forall i, j = 1, \ldots, p + 1. \quad (4)$$

$h_i(\xi(x))$ and $h_j(\eta(x))$ are the polynomials associated with the LGL points $\xi_i$ and $\eta_j$, respectively. The LGL points are the zeros of

$$(1 - \xi^2)P'_N(\xi) = 0$$

where $P'_N$ is the derivative of the $N^{th}$-order Legendre polynomial. Quadrature is performed on the reference element $\hat{\Omega}^h = [-1, 1]^2$ with LGL points that have quadrature weights $\omega$. Substitution of $\sum_{k=1}^{N_p} \psi_k(x) q^h_k(t)$ into the weak form (2) yields the semi-discrete (in space) matrix problem

$$M \frac{\partial q^h}{\partial t} + A q^h + D q^h = 0 \quad (5)$$

where $q^h$ is the array of the unknowns on the grid points, and $M$, $A$, and $D$ are the global mass, advection, and diffusion matrices, respectively. These matrices are obtained from the direct stiffness summation (DSS) of the elemental matrices $M^{el}$, $A^{el}$, and $D^{el}$ given by:

$$M_{kl}^{el} = \int_{\Omega^{el}} \psi_k \psi_l d\Omega_{el} \quad (6a)$$

$$A_{kl}^{el} = \int_{\Omega^{el}} u \cdot \nabla \psi_k \psi_l d\Omega_{el} \quad (6b)$$

$$D_{kl}^{el} = \int_{\Omega^{el}} \nu \nabla \psi_k \cdot \nabla \psi_l d\Omega_{el}. \quad (6c)$$
By construction, the mass matrix $M$ is diagonal (assuming inexact integration) and invertible. This makes these methods particularly well suited for explicit time integration schemes.

All the integrals defined above are approximated by the quadrature formula

$$
\int_{\Omega h} (\cdot) d\mathbf{x} = \int_{-1}^{1} \int_{-1}^{1} (\cdot)|J(\xi, \eta)| d\xi d\eta \approx \sum_{i=1}^{p+1} \sum_{j=1}^{p+1} (\cdot)|J(\xi, \eta)| \omega_i \omega_j,
$$

(7)

where $J$ is the Jacobian matrix associated with the map between the physical element $\Omega h(x, y)$ and the reference element $\hat{\Omega} h(\xi, \eta)$. The integration is exact up to polynomials of order $2p-1$. The $p+1$ LGL points lie along the edges and in the interior of the elements. For more on SEM see, e.g., [39, 40].

Integration in time of (5) is performed with an appropriate strong-stability preserving (SSP) time integrator. In particular, we use a five-stage explicit third-order Runge-Kutta method (RK35) [41]. SSP methods avoid the production of additional oscillations or damping.

From 3rd-order and up, the disposition of the nodes of spectral elements differs from classical finite elements in the way represented in Figure 1 for a 4th-order element.
2.2. Stabilization techniques

Aliasing and Gibbs oscillations render the Galerkin solution of (1) unstable since unwanted oscillations will pollute the numerical solution. In the framework of spectral elements, the common strategy to control these oscillations is the use of anti-aliasing filters (See, e.g., [42, 43, 44, 40, 45] and references therein). Filtering, however, suffers from non-positivity that is unacceptable in most problems that involve transport. A suitable alternative to filters may be the use of a stabilized spectral element approximation based on a residual-based diffusion-like term added to the left-hand-side (LHS) of (2). A stabilization technique should have the important property of consistency (for instance, artificial diffusion stabilizes but is not necessarily consistent). Thus, the additional viscous term should vanish as the size of the element approaches zero. The stabilized counterpart of (2) is:

\[
\left( \psi^h, \frac{\partial q^h}{\partial t} \right) + a(\psi^h, q^h) + b(\psi^h, q^h) = (\psi^h, f) \quad \forall \psi^h \in \mathcal{W}^h,
\]

where \( b(\psi^h, q^h) \) is the stabilization term. Different possible definitions of \( b(\psi^h, q^h) \) are reported in Table 1.
Table 1: Stabilized methods

<table>
<thead>
<tr>
<th>Method</th>
<th>$b(\psi^h, q^h)$</th>
<th>$\mathcal{L}(\psi^h)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AV/HV</td>
<td>$\int_{\Omega'} \mathcal{L}(\psi^h) \bar{\nu} \left[ \nabla^\alpha q^h \right] d\Omega'$</td>
<td>$\mathcal{L} = \nabla^\alpha \psi^h$</td>
</tr>
<tr>
<td>SU</td>
<td>$\int_{\Omega'} \mathcal{L}(\psi^h) \bar{\nu} \left[ u \cdot \nabla q^h \right] d\Omega'$</td>
<td>$\mathcal{L} = u \cdot \nabla \psi^h$</td>
</tr>
<tr>
<td>SUPG</td>
<td>$\int_{\Omega'} \mathcal{L}(\psi^h) \tau \left[ \partial_t q^h + u \cdot \nabla q^h - \nabla \cdot \left( \nu \nabla q^h \right) - f \right] d\Omega'$</td>
<td>$\mathcal{L} = u \cdot \nabla \psi^h$</td>
</tr>
<tr>
<td>GLS</td>
<td>$\int_{\Omega'} \mathcal{L}'(\psi^h) \tau \left[ \partial_t q^h + u \cdot \nabla q^h - \nabla \cdot \left( \nu \nabla q^h \right) - f \right] d\Omega'$</td>
<td>$\mathcal{L}' = -u \cdot \nabla \psi^h - \nu \Delta \psi^h$</td>
</tr>
<tr>
<td>VMS</td>
<td>$-\int_{\Omega'} \mathcal{L}^* (\psi^h) \tau \left[ \partial_t q^h + u \cdot \nabla q^h - \nabla \cdot \left( \nu \nabla q^h \right) - f \right] d\Omega'$</td>
<td>$\mathcal{L}^* = -u \cdot \nabla \psi^h - \nu \Delta \psi^h$</td>
</tr>
</tbody>
</table>

In Table 1 and whenever they are found below, the integrals on $\Omega'$ are defined by

$$
\int_{\Omega'} (\cdot) d\Omega' \doteq \sum_{el} \int_{\Omega_{el}} (\cdot) d\Omega_{el}, \tag{9}
$$

where $\Omega'$ is the union of the element interiors only.

In the case of AV/HV and SU, $b(\psi^h, q^h)$ is a function of a constant diffusivity coefficient, $\bar{\nu}$, and $\alpha = \beta/2$, where $\beta$ is a positive even power of the hyper-viscosity operator ($\beta = 2$ yields the usual AV). Although HV is scale-selective (i.e., it damps only higher frequencies), it is not consistent, nor is it physical. In fact, to maintain the correct physical dimensions of the hyper-viscous operator, the value of $\bar{\nu}$ must be different when different $\alpha$ are used. Its selection is hence not trivial. Furthermore, as Figure 21-d indicates, the diffusion is isotropic and spatially homogeneous. The operator does not incorporate the problem’s physics. On the other hand, in the case of
SUPG, GLS, and VMS, \( b(\psi^h, q^h) \) depends on what is referred to as intrinsic time, \( \tau \) (see Table 1). \( \tau \) is not constant or uniform; it depends on the characteristics of the grid and of the physics of the problem (e.g., velocity, diffusion). Its definition has major influence on the accuracy of the solution. A general method for finding \( \tau \) does not exist yet. The literature on the optimal selection of \( \tau \) is vast, and we refer to \([26, 27]\) for a comprehensive analysis of different definitions. Subsection 2.3 and Appendix A show how we build \( \tau \) for high order spectral elements. The method of Douglas and Wang (DW) \([46]\) was omitted from Table 1 as it can be included into the VMS method, although its derivation was specifically done in the context of Stokes problems rather than scalar advection-diffusion. SUPG and VMS will be described in the following subsections.

2.2.1. Streamline-upwind/Petrov-Galerkin (SUPG)

The SUPG method was designed by Brooks and Hughes \([9]\) and was later generalized for multidimensional problems by Hughes and Mallet \([47]\). It is a consistent alternative to the artificial diffusion approach or to the overly diffusive streamline upwind (SU) method. Its use has been ubiquitous in the solution of transport problems by the finite element method (See, e.g., \([48, 29, 49, 50, 51]\)). The application of this strategy to higher-order schemes was first tested for spectral methods by Canuto and coworkers in \([17, 18, 19, 52]\), and later by Hughes and coworkers in \([21]\) using non-uniform rational B-splines (NURBS). In this paper we show its properties when used with high-order spectral elements. SUPG is a Petrov-Galerkin method in that it does not assume that the basis and test functions live in the same space. We introduce the additional space \( \Psi^h \) of test functions \( w^h \) defined by
We have the problem of finding the function \( q^h \in (W^h; 0, t) \) such that

\[
\left( \psi^h + \tau \mathbf{u} \cdot \nabla \psi^h, \frac{\partial q^h}{\partial t} \right) + a(\psi^h + \tau \mathbf{u} \cdot \nabla \psi^h, q^h) = (\psi^h + \tau \mathbf{u} \cdot \nabla \psi^h, f) \quad \forall \psi^h \in W^h.
\] (10)

Rearrangement of (10) yields

\[
\left( \psi^h, \frac{\partial q^h}{\partial t} \right) + a(\psi^h, q^h) - (\psi^h, f) + b(\psi^h, q^h) = 0 \quad \forall \psi^h \in W^h
\] (11)

where

\[
b(\psi^h, q^h) = \int_{\Omega^h} \left[ \frac{\partial q^h}{\partial t} + \mathbf{u} \cdot \nabla q^h - \nabla \cdot (\nu \nabla q^h) - f \right] \tau \mathbf{u} \cdot \nabla \psi^h \, d\Omega^h
\] (12)

is the stabilizing term. In (12), \( \partial_t q^h + \mathbf{u} \cdot \nabla q^h - \nabla \cdot (\nu \nabla q^h) - f \) is the residual of the governing equation, and \( \tau \) is the stabilization parameter whose construction is reported below in the specific context of variational multiscale stabilization.

2.2.2. The variational subgrid scale formulation (VMS)

Let \( W^h \) be the space of resolved scales and let \( \tilde{W} \) be a space that completes \( W^h \) in \( W \) and that we will call the space of subgrid scales. VMS
relies on the decomposition $W = W^h \oplus \tilde{W}$, where $\oplus$ is the overlapping sum decomposition of the two spaces. The elements that belong to $\tilde{W}$ are $\tilde{q}$ and $\tilde{\psi}$ and are such that $q = q^h + \tilde{q}$ and $\psi = \psi^h + \tilde{\psi}$. Using such decomposition of $q$ and $\psi$ and anticipating that we will consider $\tilde{q}$ to be quasi-static [37], we re-write (2) as

$$
\left( \psi^h + \tilde{\psi}, \frac{\partial q^h}{\partial t} \right) + a(\psi^h + \tilde{\psi}, q^h + \tilde{q}) = (\psi^h + \tilde{\psi}, f) \quad \forall \psi^h \in W^h, \tilde{\psi} \in \tilde{W}.
$$

(13)

By virtue of the linear independence of $\psi^h$ and $\tilde{\psi}$ we can first take $\tilde{\psi} = 0$ and then $\psi^h = 0$ and find the split problem:

$$
\left( \psi^h, \frac{\partial q^h}{\partial t} \right) + a(\psi^h, q^h) + a(\psi^h, \tilde{q}) = (\psi^h, f) \quad \forall \psi^h \in W^h
$$

(14a)

$$
\left( \tilde{\psi}, \frac{\partial q^h}{\partial t} \right) + a(\tilde{\psi}, q^h) + a(\tilde{\psi}, \tilde{q}) = (\tilde{\psi}, f) \quad \forall \tilde{\psi} \in \tilde{W}.
$$

(14b)

We make the assumptions that $\tilde{\psi}(\partial \Omega) = 0$ and $\tilde{q}(\partial \Omega) = 0$, and that $\tilde{\psi}(\partial \Omega_{en}) = 0$.

Following [6], in (14) we integrate by parts the bilinear forms that depend on the subgrid-scales and write the following:

$$
a(\psi^h, \tilde{q}) = (\mathcal{L}^* \psi^h, \tilde{q}) \quad \forall \psi^h \in W^h \text{ and } \tilde{q} \in \tilde{W}
$$

$$
a(\tilde{\psi}, q^h) = (\tilde{\psi}, \mathcal{L} q^h) \quad \forall \tilde{\psi} \in \tilde{W} \text{ and } q^h \in W^h
$$

$$
a(\tilde{\psi}, \tilde{q}) = (\tilde{\psi}, \mathcal{L} \tilde{q}) \quad \forall \tilde{\psi} \in \tilde{W} \text{ and } \tilde{q} \in \tilde{W},
$$
where $\mathcal{L}^*$ is the dual (or adjoint) of $\mathcal{L}$. The following system is then found:

\[
\left( \psi^h, \frac{\partial q^h}{\partial t} \right) + a(\psi^h, q^h) + (\mathcal{L}^* \psi^h, \tilde{q}) = (\psi^h, f) \quad \forall \psi^h \in W^h \tag{15a}
\]

\[
\left( \tilde{\psi}, \frac{\partial q^h}{\partial t} \right) + (\tilde{\psi}, \mathcal{L} q^h) + (\tilde{\psi}, \mathcal{L} \tilde{q}) = (\tilde{\psi}, f) \quad \forall \tilde{\psi} \in \tilde{W}, \tag{15b}
\]

where

\[
(\mathcal{L}^* \psi^h, \tilde{q}) = \int_{\Omega'} \mathcal{L}^*(\psi^h) \tilde{q} \, d\Omega',
\]

\[
(\tilde{\psi}, \mathcal{L} q^h) = \int_{\Omega'} \tilde{\psi} \mathcal{L}(q^h) \, d\Omega',
\]

\[
(\tilde{\psi}, \mathcal{L} \tilde{q}) = \int_{\Omega'} \tilde{\psi} \mathcal{L}(\tilde{q}) \, d\Omega'.
\]

**Observations on time-dependent subgrid-scales:** The time-dependent approximation (13) would include a contribution from the time evolution of the subscales given by $\partial_t \tilde{q}$ if the hypothesis of *quasi-static subscales* (i.e. $\partial_t \tilde{q} \approx 0$) had not be considered. Under this hypothesis, the contribution from the subgrid scales only appears in the steady part of the Galerkin approximation. If a sufficiently small time-step is used with an explicit time integrator, we do not lose accuracy with the quasi-static hypothesis. With the use of large time-steps with semi-implicit time integrators in atmospheric simulations, tracking of the subscales is hence needed. This issue is reserved for future work by the authors.
Remark 2: It must be pointed out that, for the solution of the scalar transport equation, the expressions for SUPG and VMS are the same.

Approximation of the sub-grid scales. The unresolved quantity \( \tilde{q} \) has not been defined in any way yet. In the following, the subscales are constructed by algebraic approximation following the technique described by [35] for linear, quadratic, and cubic elements. They take \( \tilde{W} \) as the space of vanishing functions on the boundaries of each element. These are called bubble functions (see [53, 54]). By incorporating the time-dependent term of Eq. (15b) within the second inner product and by re-arranging the terms, the equation for the subgrid scales \( \tilde{q} \),

\[
\int_{\Omega'} \tilde{\psi} \mathcal{L}(\tilde{q}) \, d\Omega' = - \int_{\Omega'} \tilde{\psi} \left[ \mathcal{L}(q) - f \right] \, d\Omega' \quad \forall \tilde{\psi} \in \tilde{W}, \quad (17)
\]

is found. The equivalent strong form of (17) is

\[
\mathcal{L}(\tilde{q}) = - \left[ \mathcal{L}(q^h) - f \right] = -R(q^h), \quad (18)
\]

where \( R(q^h) \) is the residual of the equation for the resolved scale. For the solution of (18) we first define \( \tilde{q} \) as a function of the bubbles \( b(x) \); we have \( \tilde{q} = -b(x)R(q^h) \) that is plugged into (18) to find the differential problem

\[
\mathcal{L}(-b(x)R(q^h)) = -R(q^h). \quad (19)
\]

Eq. (19) is solved for \( b \) with Dirichlet boundary conditions \( b(x_1) = 0 \) and \( b(x_2) = 0 \) on every element of length \( h = |x_1 - x_2| \). By thinking that \( R(q^h) \) is
always known from the previous time-step, we can consider it as a constant and can hence be taken out of the $\mathcal{L}(\cdot)$ operator so that $L(b(x)) = 1$. For the one-dimensional steady-state advection-diffusion equation the problem is

$$\mathcal{L}(b(x)) = u b_x(x) - \nu b_{xx}(x) = 1,$$  \hspace{1cm} (20)

with exact solution

$$b(x) = \frac{x}{u} + \frac{h}{u} \frac{1 - e^{xu/\nu}}{e^{hu/\nu}}.$$  \hspace{1cm} (21)

Now that we computed the bubble functions along the element length, we construct the stabilization parameter $\tau$ as the mean value of $b(x)$ along each element. We have that

$$\tau = \frac{1}{|x_1 - x_2|} \int_{x_1}^{x_2} b(x) dx.$$  \hspace{1cm} (22)

Integration of $b(x)$ (21) in the interval $[x_1, x_2] = [0, h]$ yields the following

\footnote{The bubbles are computed only once, out of the time loop; only if they depend on non-constant coefficients, they are computed once per time-step. In either case, the problem to be solved is not time-dependent.}
definition of $\tau$:

$$
\tau = \frac{h}{2u} \left( \coth(Pe_k) - \frac{1}{Pe_k} \right), \quad (23)
$$

where

$$
Pe_k = \frac{uh}{2\nu}
$$
is the element Péclet number.

We have derived all the ingredients to define the sub-grid scales $\tilde{q}$ as the algebraic approximation [6]

$$
\tilde{q} = -\tau R(q^h). \quad (24)
$$

Eq. (15b) was used as the starting point to approximate $\tilde{q}$. Now, by plugging (24) into (15a), the VMS stabilized Galerkin method is found and expressed as follows:

Find $q^h \in W^h$ such that

$$
\left( \psi^h, \frac{\partial q^h}{\partial t} \right) + a(\psi^h, q^h) - (f, \psi^h) - \int_{\Omega'} L^*(\psi^h) \tau R(q^h) d\Omega' = 0 \quad \forall \psi^h \in W^h. \quad (25)
$$

Eq. (25) differs from Eq. (2) by the additional term that models the subgrid scales. The extra term is the viscous-like contribution that stabilizes the equation.
2.3. $\tau$ for spectral elements

For linear elements, $h = |x_1 - x_2|$ is simply the length of the element. For higher-order finite elements, $h$ becomes a fraction of the total element size if the internal nodes are equi-spaced. In the case of spectral elements, where the LGL points are unevenly distributed, the integral is computed by using $h$ as the local distance between two consecutive points. The stabilization parameter $\tau$ is built inside the element as a function of the bubbles on every segment delimited by two consecutive nodes. This means that equation $\mathcal{L}(b(x)) = 1$ is solved on every sub-element by applying homogeneous Dirichlet b.c. at the sub-element boundaries. For example, for a second-order element with one internal node we would solve $\mathcal{L}(b(x)) = 1$ on the two segments $[x_1, x_2]$ and $[x_2, x_3]$, respectively, by applying homogeneous b.c. as $b(x_1) = 0, b(x_2) = 0$ and $b(x_2) = 0, b(x_3) = 0$. With this, two $\tau$’s would be computed as

$$\tau^{i+1}_i = \frac{1}{x_{lgl}(i + 1) - x_{lgl}(i)} \int_{x_{lgl}(i)}^{x_{lgl}(i+1)} b(x) \, dx,$$  \hspace{1cm} (26)

where $x_{lgl}(i + 1)$ and $x_{lgl}(i)$ are the coordinates of two consecutive LGL points.

The most simple (but not unique) way of proceeding is that of taking the average value of all the sub-$\tau$’s as the value of the full element $\tau$. This method is used in this paper.

The uneven spacing of the element nodes is the major difference with respect to the definitions derived in previous studies. In this case the intrinsic time is non-uniform along the element, as it appears in Figure 2, where the bubbles and corresponding $\tau$’s are displayed for an element of order
7. In Appendix A we report the explicit expression to compute $\tau_k$ along a high-order spectral element.

2.4. Spurious oscillations at layers diminishing (SOLD) methods

Methods in the form of (25) may produce overshoots and undershoots in the proximity of an internal or boundary layer. These unwanted oscillations can be suppressed, without affecting the global solution, by adding an additional diffusive term of the form

$$\left(\mathbf{\nabla} \psi^h, \bar{\tau} \mathbf{\nabla} q^h\right),$$

(27)

where consistency must be respected through a proper construction of $\bar{\tau}$. We would like to have a method that does not modify the diffusion in the
streamline direction since that is already accounted for by the stabilization term, but also avoids overdamping in the crosswind direction. The comprehensive set of tests performed by John and Knobloch reveals that Codina’s [28] is among the best methods that satisfy these conditions when used with finite elements. In [28], $\tilde{\tau}$ is defined by:

$$\tilde{\tau} = \frac{1}{2} \max \left\{ 0, C - \frac{2\nu}{u_\parallel h_k} \right\} h_k \frac{R(q^h)}{||q^h||} \left( I - \frac{u \otimes u}{|u|^2} \right)$$

where $C$ is a constant, $u_\parallel$ is the velocity component in the streamline direction, and $\otimes$ indicates a tensor product. Codina suggests $C = 0.7$ for linear and bilinear elements, and $C = 0.35$ for quadratic and biquadratic elements. However, for higher order elements using LGL points, we found that the best results were obtained by setting $C = 1$, as long as $h_k$ is selected properly in the construction of both $\tilde{\tau}$ and $\tau_k$.

An alternative to (27) comes from Johnson, Schatz, and Wahlbin [55] who defined the following:

$$(\tilde{\tau}u^\perp \cdot \nabla \psi^h, u^\perp \cdot \nabla q^h) = \left( \frac{-w, u}{|u|} \right).$$

In the current work, (29) gives better results than (27), and was then used throughout. The results obtained with this technique are labeled with $DC$ for Discontinuity Capturing.

2.5. First-Order Subcells (FOS)

FOS is one additional tool that can further help the suppression of Gibbs oscillations. The concept is simple and is easily coded on structured grids. If
a solution has large gradients, the algorithm needs to identify the elements where the large gradients occur, and project the solution scheme to a $1^{st}$-order space. The gradient is sought with a proper error estimator. The simple physics of the advection-diffusion problems discussed below allows for the energy-norm of the gradient of the solution to be a sufficiently good estimator for the current study. As it is defined in this study, the error estimator depends on a parameter, $\epsilon$, that may be a function of the numerical settings (e.g., grid resolution, time step). This point must be considered in the construction of FOS and in the selection of the error estimator. We did not explore this further in this study, although it is a very important issue for the best performance of FOS.

Algorithm 1 is a simple implementation of this concept within our code. We present the pseudo-code below for the sake of clarity. The method was applied to a two-dimensional advection-diffusion problem with internal and boundary layers in a skew velocity field. Results are shown in Figures 9-12 and 14-18. A detail of Figures 18(a,b) is presented in Figure 19.

In the tests that use Algorithm 1, $\epsilon$ was set to 0.5.

3. Mass conservation

For problems in geophysical fluid dynamics and, more specifically, in atmospheric simulations, mass conservation of tracers is an important ingredient. In this section we address this issue and illustrate how the algorithm reported in this paper behaves in this respect. Under the suitable hypothesis of a divergence-free flow, Eq. (1) for the transport of the mixing ratio

$$q = \frac{\rho_{\text{tracer}}}{\rho},$$
for $iel = 1$ to $nelem$ do

// Check if the element contains a sufficiently large gradient:
if $iel$ is s.t. $||\nabla q^h||_2 > \epsilon$ then

// Treat element $iel$ as a sub-domain made of $(ngl - 1) \times (ngl - 1)$
sub-elements (isubel):
for $isubel = 1$ to $p^d$ do
  Create mass and rhs using $1^{st}$-order basis functions and
  integration rule
end for

else
  Create rhs for the high-order spectral element.
end if

end for

Algorithm 1: Compute the $1^{st}$-order rhs
where $\rho_{\text{tracer}}$ and $\rho$ are, respectively, the densities of the tracer and of the advecting fluid, is derived from the equation of conservation of mass of the tracer\(^2\)

$$\frac{\partial \rho q}{\partial t} + \nabla \cdot (u \rho q) = 0 \quad (30)$$

by elimination of $\rho$ from the conservation law (continuity equation)

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (u \rho) = 0. \quad (31)$$

In (30) and (31) the quantities that are conserved are $\rho_{\text{tracer}} = \rho q$ and $\rho$, but not $q$. In the case of $\rho(t) = \rho(t = 0) = \text{constant}$ in a non-divergent flow, Eq. (1) is equivalent to

$$\frac{\partial q}{\partial t} + \nabla \cdot (u q) = 0. \quad (32)$$

This allows the equal treatment of $\rho q$ and $q$ in the modeling of the physical system at hand [56]. Regardless of the poor conservation properties of transport expressed in advective form, Eq. (1) is of common use within atmospheric models (e.g., [1, 56, 57]). In this paper we track mass loss during the simulations to evaluate the amount of mass loss that we would run into if using SEM+VMS knowing that no method will conserve for this equation since it is not in conservation form or a conservation law.

---

\(^2\)For simplicity, the diffusion and source terms were dropped.
For this computation, we use the advection of a square wave in a periodic channel and report the results in paragraph \textit{Tr2-2D} below.

4. Numerical testing

The algorithms discussed throughout this paper are tested by using standard one- and two-dimensional problems. The problems are organized according to the nomenclature listed below:

- 1D Steady-state homogeneous advection-diffusion (\textit{St-1D})
- 1D Steady-state advection-diffusion with source (\textit{St-1D-S})
- 2D Steady-state advection-diffusion with internal and boundary layers (\textit{St-2D}).
- 2D Time-dependent advection-diffusion with “L”-shaped discontinuity (\textit{Tr1-2D}).
- 2D Time-dependent advection of a sharp tracer in a doubly periodic channel (\textit{Tr2-2D}).
- 2D Smooth solid-body rotation - Convergence study (\textit{2D Smooth solid-body rotation}).
- Pseudo-3D advection in a neutrally stratified atmospheric flow (\textit{Atmo-3D}).

\textit{St-1D}. One-dimensional steady-state advection-diffusion. The tracer $q^h$ is propagated with constant velocity $u = 1 \text{ m s}^{-1}$ and diffusivity $\nu = 1/512 \text{ m}^2 \text{ s}^{-1}$ first on two elements of order $p = 10$ (Figure 3), and then on four elements of order $p = 12$ (Figure 4). The domain is the line segment $\Omega = [-1, 1]$ with
Dirichlet boundary conditions $q^h(-1) = 0$ and $q^h(1) = 1$. We compared the filtered (top row) against the stabilized solution (bottom row) and observe a decrease of oscillations and undershoots. Also, at higher order and finer resolution, the capabilities of the filter are clearly being challenged by the presence of the boundary layer at $x = 1$. At the same time, small oscillations near the nodes of the element by the boundary layer are not completely suppressed by the stabilized method either; hence, additional localized smoothing is sought.

Figure 3: St-1D: $\nu = 1/512 \text{ m}^2 \text{ s}^{-1}$. The exact solution is dashed. The circles indicate the grid points.

Figure 4: St-1D: $\nu = 1/512 \text{ m}^2 \text{ s}^{-1}$. The exact solution is dashed. The circles indicate the grid points.
St-1D-S. Steady state advection-diffusion with source term $f = 1$. $q^h$ is propagated with constant velocity $u = 1 \text{ m s}^{-1}$ and two different diffusivities: $\nu = 5 \times 10^{-3} \text{ m}^2 \text{ s}^{-1}$ and $\nu = 5 \times 10^{-2} \text{ m}^2 \text{ s}^{-1}$. The domain is the line segment $\Omega = [0,1]$ and homogeneous Dirichlet boundary conditions are imposed. The domain is subdivided into two elements of order $p = 16$ and runs are compared using filtered SE (top row in Figures 5 and 6), and VMS (bottom row). We observe a very similar behavior of the solution among the two different cases in the smooth problem (Figure 5). The results are comparable to the ones obtained by Houzeaux et al. with their $\tau$ for quadratic and cubic elements in [35].
**Figure 5:** *St-1D-S:* $\nu = 5 \times 10^{-3} \text{ m}^2 \text{ s}^{-1}$. 2
16th-order elements. The exact solution is dashed. The circles indicate the grid points.

**Figure 6:** *St-1D-S:* $\nu = 5 \times 10^{-2} \text{ m}^2 \text{ s}^{-1}$. 2
16th-order elements. The exact solution is dashed. The circles indicate the grid points.

**St-2D.** Standard steady advection-diffusion skewed to the mesh (e.g., [28]):
a discontinuity is propagated with constant velocity $\mathbf{u} = (1, -2) \text{ m s}^{-1}$ and
diffusivity $\nu = 10^{-8} \text{ m}^2 \text{ s}^{-1}$ in the unit square $\Omega = [0, 1] \times [0, 1]$. The initial
configuration is shown in Figure 7.
Figure 7: St-2D: initial configuration of the steady-state problem

Dirichlet boundary conditions are prescribed:

\[ q^h = \begin{cases} 
1 & \text{if } y = 1, \\
1 & \text{if } x = 0 \text{ and } y \geq 0.7, \\
0 & \text{otherwise.}
\end{cases} \]

In Figures 8-12 we illustrate the run of the same case with different number of elements and order of the interpolating polynomials. For direct comparison of our solution with the ones in the existing literature of finite elements, we first run the test with linear elements \((p = 1)\), and present the results in Figure 8. The multiscale solution of this problem (see Figure 8a) shows important boundary and internal layers that are damped with the discontinuity capturing techniques of Section 2.4. The application of the discontinuity capturing (DC) scheme greatly improves the solution and yields monotonicity (see Figure 8b). In Figure 9 we maintained the same number of nodes of the previous run, but increased to 4\(^{th}\) the order of interpolation to assess the algorithm in the context of this paper (i.e. 50 elements of order
4 were used instead of 200 elements of order 1). The similar behavior of the solution with respect to the 1\textsuperscript{st}-order polynomial run suggests that the residual-based methods as implemented in this study may not be sensitive to the distribution of the interpolation nodes within the elements edges. As it appears in Figures 9\textit{c}, the behavior is completely analogous to the previous run. However, monotonicity is lost in two singular nodes: with reference to Figure 9\textit{c}, the 4\textsuperscript{th}-order solution is smooth and monotone everywhere except for the nodes represented as points A and B in Figure 7. This is not surprising: at A and B the tracer is leaving the boundary with a skew angle; an incorrect imposition of boundary conditions at these nodes may be causing the problem. The numerical singularity at this points should be addressed but it will not be done in the current work. These are fully suppressed by applying the FOS algorithm described above, as it is shown in Figure 9\textit{d}.

Decreasing the number of computational nodes by doubling the order from 4 to 8 and setting the number of elements to 10 in \(x\) and \(z\), even with a discontinuity capturing term, the solution starts to lose monotonicity. This appears in Figures 11 and 12, where extrema get larger than in the previous cases. This problem shows that the construction of the stabilizing parameter \(\tau\) should include information on the order of the interpolating polynomial.

For a better view of the problem, in Figures 10 and 12 we present a vertical slice of the solution. The boundary layers are evident. Their damping, however, is clear if \(VMS\), \(DC\), and \(FOS\) are applied.
Figure 8: St-2D: steady-state solution on $200 \times 200$ 1st-order elements. (For plotting only, the data are interpolated to a $50 \times 50$ node grid using Octave [58]).
Figure 9: St-2D: steady-state solution on 50×50 4th-order elements.
Figure 10: S1-2D: steady-state solution on 50×50 4th-order elements. Vertical slice at $z = 0.3$. 
Figure 11: St-2D: steady-state solution on $10 \times 10$ 8th-order elements.
Figure 12: St-2D: steady-state solution on $10 \times 10$ $8^{th}$-order elements. Vertical slice at $z = 0.3$.

$Tr1-2D$. Transient advection-diffusion of an L-shaped discontinuity in a flow where $\nu = 10^{-6} \, m^2 \, s^{-1}$ and the velocity $\mathbf{u}$ of magnitude $|\mathbf{u}| = 0.5\sqrt{2} \, m \, s^{-1}$ is at $45^0$ with respect to the axis $(x, z)$. The initial configuration is shown in Figure 13.
Figure 13: Tr1-2D: initial configuration of the L-shaped problem

The convex shape of the sharp discontinuity makes this problem more challenging than the previous case [59], and is chosen to analyze robustness and accuracy of the algorithm. Runs were performed at two different resolutions and two different orders of interpolating polynomials. In particular we have: approx. 100 points per side using $25 \times 25$ 4$^{th}$-order elements (Figure 14), and $12 \times 12$ 8$^{th}$-order elements (Figure 15); and approx. 200 points per side using $50 \times 50$ 4$^{th}$-order (Figure 17), and $25 \times 25$ 8$^{th}$-order (Figure 17). In the figures, Filter means that the SEM solution was filtered at every time-step. VMS and/or DC indicate that the SEM solution is stabilized by the VMS with or without a discontinuity capturing term (DC). VMS + DC + FOS indicates the contribution of FOS as well. Positivity is not preserved in the solution obtained with a filter. The sharp front, in fact, makes the filter inappropriate. However, similarly to the steady advection-diffusion test St-2D, the VMS-stabilized solution of this problem is characterized as well by the formation of internal layers that run along the edges of the tracer in the direction of the flow (See, e.g., Figure 14b), and VMS is not sufficient to preserve monotonicity unless it is supplemented by the additional DC term.
defined in (29). This effect is displayed in Figures 14, 15, 16, and 17.

The consideration made for problem St-2D on the singular peaks that form at the nodes where the tracer leaves the boundary at an angle, applies here at nodes A and B of Figure 13. This is visible in Figure 18 obtained by slicing the tracer along $z = 0$ in Figures 16 and 17, respectively. The problem is solved by the application of FOS.

As the order of interpolation is increased from 4$^{th}$ to 8$^{th}$, the smooth solution begins to lose positivity. As interpreted for St-2D, the solution is clearly being affected when the interpolation nodes are densely clustered towards the boundaries of the elements, as is the case for higher order.
Figure 14: $Tr1-2D$: $4^{th}$ order $25 \times 25$. $t = 0.25 \text{s}$.

Figure 15: $Tr1-2D$: $8^{th}$ order $12 \times 12$. $t = 0.25 \text{s}$. 
Figure 16: $Tr1-2D$: 4th order $50 \times 50$. $t = 0.25 \, s$.

Figure 17: $Tr1-2D$: 8th order $25 \times 25$. $t = 0.25 \, s$. 
Figure 18: Tr1-2D: Vertical slice at $z = 0.0$. 

(a) 4$^{th}$ - order $50 \times 50$

(b) 8$^{th}$ - order $25 \times 25$
Figure 19: Tr1-2D: detail of Figure 18. Region with undershoots. Vertical slice at $z = 0.0$. 

(a) 4th - order $50 \times 50$

(b) 8th - order $25 \times 25$
Tr2-2D. Linear advection of a 2D square wave along $x$ in the periodic domain $\Omega = [0, 1] \times [0, 1]$: the tracer is transported with velocity $u = (1/2, 0) \text{ m s}^{-1}$ for one periodic revolution along $x$. The initial concentration $q^h = 1$ is centered at $(x_c, z_c) = (0.5, 0.5)$ (Figure 20). The computational finite domain consists of $11 \times 11$ quadrilaterals of order 11.

![Figure 20: Tr2-2D: initial configuration of the pure advection problem.](image)

As in the steady case, Figures 21-23 display improvement of the solution in terms of monotonicity when the VMS method is used instead of the filter. The combination of VMS and filtering is not recommended (result not shown); although VMS alone controls the over- and under-shootings along the streamlines, the addition of the filter at the end of every time step degrades positivity in the neighborhood of large gradients.

In Figures 22 and 23 we present the streamline and crosswind sections of the solution obtained by slicing the tracer along $z = 0.5$ and $x = 0.5$, respectively. Unlike the previous problems characterized by internal and boundary layers, for pure advection the VMS preserves the maximum and minimum concentrations $q_{\text{max}}^h = 1$, and $q_{\text{min}}^h = 0$ and is free of spurious oscillations. As a point of comparison, we present the result of classical
artificial-viscosity in Figures 21-23d.

Figure 21: Tr2-2D: Surface plot of the concentration field: $\Delta t = 0.001 \text{s} \ (\text{except for HV: } \Delta t = 0.0002 \text{s})$, $11 \times 11$ elements with $11^{th}$ order polynomials. Results at $t = 2.0 \text{s}$ (after 1 periodic revolution along $x$).
Figure 22: $Tr2$-2D: Streamline cut at 0.5 m in the y-direction. $\Delta t = 0.001\,s$, $11\times11$ elements with 11th order polynomials. Results at $t = 2\,s$ (after 1 periodic revolution along x). Solid line indicates the computed solution. The dashed line is the analytic solution.
Figure 23: Tr2-2D: Crosswind cut at 0.5 m in the x-direction. $\Delta t = 0.001 \text{s}$, 11×11 elements with 11th order polynomials. Results at $t = 2 \text{s}$ (after 1 periodic revolution along x). Solid line indicates the computed solution. The dashed line is the analytic solution.

**Testing mass conservation.** Because of the periodic boundary conditions applied here, we compute mass conservation properties for this test. At every time-step, the total mass loss of $\rho q$ (for $\rho = 1$) is computed as

$$M_{\text{loss}}(t) = \frac{\int_{\Omega} (\rho q(t) - \rho q(t_0)) \, d\Omega}{\int_{\Omega} \rho q(t_0) \, d\Omega},$$  

where $\Omega$ is the domain volume and $t_0$ indicates the values at the initial
time. Figure 24 shows the evolution of the mass loss that occurs during 100 revolutions around the periodic channel of Test $T_r2-2D$. 100 revolutions happen in 100 s.

![Mass loss vs time graph](image_url)

**Figure 24:** $T_r2-2D$: evolution of the mass loss during 100 s, or 100 revolutions of the square wave around the periodic channel of Figure 20.

Although in Figure 24 there seems to be an asymptotic trend to an upper bound, this is obtained at the expenses of accuracy during long runs. Regardless of the type of equations (conservative or non-conservative), the method here proposed is certainly unable to retain all mass. Because of this, at this stage we can only think of applying this technique to short term weather forecast but not climate. This is the first application of VMS and DC to Spectral Elements to solve the advection equation; in the future we will work on a fix to this problem for better (or total) mass conservation. A first improvement of accuracy for long-time runs may be achieved using or-
orthogonal sub-grid scales (OSS) as proposed by [31]. Further analysis should be done.

2D Smooth solid-body rotation. The smooth solid-body rotation test with a smooth function is used for a grid convergence study [60, 61]. A Gaussian hill $q^h = \exp \left[ -5 \left( (x - x_c)^2 + (y - y_c)^2 \right) \right]$ is originally centered in $(x_c, y_c) = (0, 0)$ in a periodic domain $\Omega = [-\pi, \pi] \times [-\pi, \pi]$ with prescribed velocity $(u, w) = (-\pi y, \pi x)$. Convergence is computed after one full revolution ($t = 2s$). The normalized standard $L_2$ error is computed with respect to the exact solution $q_e = q^h(x, y, t = 0)$ using $N_{el} \in \{10^2, 20^2, 40^2\}$ and polynomials of degree 4 and 8. Figure 25 shows the $h$–error. The original data are reported in Table 2.

![Figure 25: 2D Smooth solid-body rotation: Log-log plots of the normalized $L_2$ error vs. $N_{el}$ using VMS or a Filter.](image)

49
Table 2: 2D Smooth solid-body rotation: normalized $L_2$ error vs. $N_{el}$. Convergence rate of every setting is reported on the last row of the table.

<table>
<thead>
<tr>
<th>$N_{el}$</th>
<th>VMS 4th</th>
<th>FILTER 4th</th>
<th>VMS 8th</th>
<th>FILTER 8th</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.7083E-02</td>
<td>0.1793E+00</td>
<td>0.3175E-05</td>
<td>0.2569E-03</td>
</tr>
<tr>
<td>400</td>
<td>0.3280E-03</td>
<td>0.2422E-01</td>
<td>0.2305E-07</td>
<td>0.4927E-05</td>
</tr>
<tr>
<td>1600</td>
<td>0.1982E-04</td>
<td>0.4127E-02</td>
<td>0.5473E-10</td>
<td>0.9388E-07</td>
</tr>
</tbody>
</table>

Convergence rate: 4.2406 2.7206 7.9120 5.7091

The experiment indicates that VMS does not affect the rate of convergence of SEM. However, the time discretization error is approximately $10^{-11}$; because of this, there is no gain in accuracy with further grid refinement from 1600 to 6400 elements unless a more accurate time discretization method is used.

Atmo-3D. The transport of a passive tracer in a neutrally stratified atmosphere in a large domain represents an idealized application to a seemingly real atmospheric problem. This final test is a proof-of-concept to verify the behavior of the methodology over larger time and spatial scales that are of relevance for real applications.

The velocity field is no longer uniform and constant, but varies non-linearly in space and time during the evolution of a rising thermal perturbation originally centered at the central lower region of the domain. The difficulty of the test is expressed by the transient character of the velocity that, in the first instant of the motion, greatly affects the stability of the solution of the advection equation. The problem is defined as follows [62]. The domain extends within $\Omega = [0, 1000] \times [0, 1000] \times [0, 1000]$ m$^3$. It is divided first into 10, and then into 20 spectral elements of order 4 along $x$ and $z$, with 1 element along $y$. The simulations final time is $t = 600$ s. A
neutral background state at uniform potential temperature $\theta_0 = 300 \, K$ is
perturbed by a cylindrical thermal bubble of radius $r_c = 250 \, m$, centered in
$(x_c, y_c, z_c) = (500, 500, 350) \, m$, and defined by

$$\theta' = A \left[ 1 + \cos \left( \frac{r \pi}{r_c} \right) \right], \quad (34)$$

where $r = \sqrt{(x-x_c)^2 + (z-z_c)^2}$ and $A = 0.5 \, K$. The top, bottom, left,
and right boundaries are modeled as non-penetrating solid walls, while pe-
riodicity is imposed on the front and back boundaries ($y$-direction).
The thermal problem is modeled by the Euler equations of inviscid com-
pressible flows and solved by the method described in [63]. The use of VMS
to solve the Euler equations falls beyond the scope of this paper, although
the authors are currently working at its implementation in the context of
spectral elements. Currently, VMS for the finite element solution of the
compressible Euler equations of dry nonhydrostatic stratified flows can be
found in [64].

At time $t = 0 \, s$, the tracer $q^h$ is centered in the same position of $\theta'$, but
is described by a cylindrical step function of maximum intensity $q^h_{\text{max}} = 0.5$
within a radius $r <= 250 \, m$. The initial state of $q^h$ and $\theta'$ is shown in Figure
26. After 600 $s$ the rising bubble has developed into the structure plotted
in Figure 27. If properly resolved, the tracer is expected to have similar
features given that the velocity field derives from the motion of the bubble.
Figure 26: *Atmo-3D*: $x - z$-slice plot at $y = 500m$ of the initial conditions of $\theta'$ (left), and $q^h$ (right).

Figure 27: *Atmo-3D*: $\theta'$ after 600 s. Grid: $20 \times 20$ elements of order 4.

The velocity field $\mathbf{u}$ is still until the warm bubble begins to move due to buoyancy. As soon as $\mathbf{u} \neq 0$, the tracer begins to move as well. The sudden change of state from rest to moving generates oscillations at the boundaries of the tracer that are more difficult to treat with respect to its analogous steady-state case. Figure 28 shows the tracer after 600 s on a grid of $10 \times 10$ elements of order 4. The filtered solution and the solution obtained with artificial diffusion (Figures 28a, 28b) have important under
and overshoots that propagate in the whole domain. The constant diffusivity coefficient used with AV is taken as the average value of $\tau$ of VMS for the same simulation. As expected from the previous tests, VMS alone is not able to fully eliminate the oscillations in the proximity of the discontinuity, however, improvement is evident (Figures 28c). The best performance is obtained with the combination VMS+DC. Using the theoretical extreme values for the tracer ($0 \leq q^{\text{exact}} \leq 0.5$), the relative error

$$\epsilon = \frac{q^h - q^{\text{exact}}}{q^{\text{exact}}}$$

is reported in Table 3. The same considerations apply for the finer-grid solution ($20 \times 20$ elements of order 4). The results are plotted in Figure 29 and the errors reported in Table 4.
Figure 28: *Atmo-3D*: Tracer after 600 sec. Grid: 10 × 10 elements of order 4.
Figure 29: Atmo-3D: Tracer after 600 sec. Grid: 20 × 20 elements of order 4.

Table 3: Atmo-3D: Relative error $\epsilon$ on the maximum (0.5) and minimum (0.0) theoretical values. Results for 10 × 10 elements of order 4.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\epsilon_{\text{min}}$</th>
<th>$\epsilon_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AV/HV ($\nu = 0.001 \text{ m}^2\text{s}^{-1}$)</td>
<td>76.89 %</td>
<td>105.37 %</td>
</tr>
<tr>
<td>AV/HV ($\nu = 0.1 \text{ m}^2\text{s}^{-1}$)</td>
<td>61.79 %</td>
<td>45.00 %</td>
</tr>
<tr>
<td>Filter</td>
<td>51.05 %</td>
<td>53.53 %</td>
</tr>
<tr>
<td>VMS</td>
<td>33.67 %</td>
<td>36.21 %</td>
</tr>
<tr>
<td>VMS + DC</td>
<td>3.53 %</td>
<td>5.20 %</td>
</tr>
</tbody>
</table>
Table 4: Atmo-3D: As for Table 3, but for $20 \times 20$ elements of order 4.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\epsilon_{min}$</th>
<th>$\epsilon_{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AV/HV ($\nu = 0.001 \text{ m}^2\text{s}^{-1}$)</td>
<td>65.26 %</td>
<td>62.59 %</td>
</tr>
<tr>
<td>AV/HV ($\nu = 0.1 \text{ m}^2\text{s}^{-1}$)</td>
<td>20.33 %</td>
<td>21.06 %</td>
</tr>
<tr>
<td>Filter</td>
<td>36.38 %</td>
<td>53.10 %</td>
</tr>
<tr>
<td>VMS</td>
<td>14.34 %</td>
<td>11.53 %</td>
</tr>
<tr>
<td>VMS + DC</td>
<td>1.24 %</td>
<td>4.63 %</td>
</tr>
</tbody>
</table>

**Remark 3: on the use of filters in the previous results** In the current work, filtering was applied in the usual way that has been used previously in SE models (see, e.g., [44, 65, 66, 67]). That is, the filtering coefficients were defined at the beginning of the simulation and applied after every time-step using the same filter matrix for all elements. It may be possible to obtain better results with filters if they are constructed in a specific way (e.g., each element uses a different filter matrix that is constructed dynamically) but a clear approach on how to do this remains an open topic since this can be viewed as a classical limiter but for Spectral Elements (see, e.g., [68]).

5. Discussion and conclusions

5.1. Conclusions

In this paper, we proposed the use of the Variational Multiscale Stabilization (VMS) method to stabilize advection-dominated problems solved with spectral elements. In the regions characterized by strong gradients, we also combine VMS, a Discontinuity Capturing technique (DC), and the First-Order Subcells method (FOS) for a better treatment of Gibbs phenomena in the proximity of boundary and internal layers. The stabilization
parameter $\tau$ that appears in the VMS scheme was computed to include the characteristics of high-order spectral elements with (non-equispaced) LGL nodes. Numerically, we demonstrated that this approach is a possible alternative to the standard filters used in the stabilization of the spectral element solvers if the suppression of unwanted under- and over-shoots is the main concern. Stabilization by these methods is obtained by introducing a diffusion-like term that is controlled and localized where the residual is important (i.e. large gradients). Where needed, the combined action of VMS and FOS yields encouraging results for high-order spectral elements.

The algorithms were evaluated on a set of standard tests of increasing difficulty. A significant improvement was observed in the performance of the spectral element solver as far as the control of extrema is concerned, both in the purely advective and in the advective-diffusive regimes. The most important features of this new approach are the following:

- Unlike hyper-viscosity, the subgrid-scale diffusion is localized and controlled.

- Under- and over-shoots are greatly suppressed relative to traditional filters.

- The VMS method does not depend on a free-parameter assigned by the user. On the other hand, in Algorithm 1 $\varepsilon$ is a free-parameter related to the simple error-estimator that was used. A more sophisticated estimator should not depend on any user-defined constant.

- Currently, the method is not fully mass-conservative. This can be an issue for long term simulations such as those for climate applications.
5.2. Application to atmospheric modeling in climate and weather prediction

In [57], a Kessler microphysics scheme was implemented within a spectral element framework that requires the advection of three moisture variables (vapor, cloud, and rain mixing ratios). This microphysics scheme will be implemented in our Nonhydrostatic Unified Model for the Atmosphere (NUMA) [63] in order to simulate both mesoscale and synoptic-scale atmospheric phenomena. As is well-known, Galerkin-based methods yield 1) higher-order accuracy and 2) excellent dispersion properties, which are both desirable for advection schemes; however, the resulting Gibbs oscillations produce strong gradients that must be remedied in some fashion. At present, a simple-minded “fixer” is applied whereby negative values of the moisture variables are set equal to zero. This fixer acts as an effective mass source, thus violating the conservation properties of the model. In addition, this fixer violates the function space that the spectral element solution inhabits. For these reasons, monotonic advection of tracer variables is essential for any atmospheric model. The proposed VMS+DC+FOS technique is a candidate since it 1) preserves monotonicity better than the standard filter approach, 2) does not significantly increase the cost of the spatial discretization scheme, and 3) is completely local in nature (i.e., no additional communications are required in a parallel environment), which is necessary for scaling on modern distributed and hierarchical memory environments. However, the method should be improved for better mass conservation, especially when the time scales at hand are large.
6. Acknowledgements

S.M. and F.X.G. gratefully acknowledge the support of the Office of Naval Research Global (ONRG) that provided support for S.M. through the Visiting Scientist Program with grant N62909-09-1-4083. The work of F.X.G. and J.F.K. was supported by the Office of Naval Research (ONR) through grant PE-0602435N. J.F.K. is supported by the National Research Council (NRC). Discussions on VMS with Guillaume Houzeaux, support by Shiva Gopalakrishnan with runs on the cluster at NPS, and comments by Luca Bonaventura on the manuscript are greatly appreciated. The authors are thankful to an anonymous reviewer for his remarks and suggestions on mass conservation.

Appendix A. Explicit expression for $\tau$ on high-order elements

In this appendix, we explicitly derive the expression for $\tau$ defined between two consecutive LGL points $[x_{lgl}(i), x_{lgl}(i + 1)]$. The bubble obtained from the integration of (20) with boundary conditions $b(x_{lgl}(i)) = 0$ and $b(x_{lgl}(i + 1)) = 0$ has expression:

$$
\begin{align*}
    b(x) &= \frac{x}{u} - \frac{x(i + 1) - x(i)}{u(e^{ux(i+1)/\nu} - e^{ux(i)/\nu})} e^{ux(i)/\nu} - \frac{x(i)e^{ux(i+1)/\nu} - x(i + 1)e^{ux(i)/\nu}}{u(e^{ux(i+1)/\nu} - e^{ux(i)/\nu})}.
\end{align*}
$$

The subscript $LGL$ is omitted to keep the long expressions simple to read.

The evaluation of the integral (26) yields the expression:

$$
\begin{align*}
    \tau_{x(i)}^{x(i+1)} = \frac{1}{x(i + 1) - x(i)} \left[ \frac{x(i) - x(i + 1)}{u} \left( \frac{\nu}{u} + \frac{x(i + 1) - x(i)}{2} \right) - \frac{e^{ux(i+1)/\nu}(x(i) - x(i + 1))^2}{e^{ux(i)/\nu} - e^{ux(i+1)/\nu}} \right].
\end{align*}
$$
When \( x(i) = 0 \) and \( x(i + 1) = h \), we have that

\[
\tau_0^h = -\frac{\nu}{u^2} - \frac{h}{2u} + \frac{he^{uh/\nu}}{e^{uh/\nu} - 1},
\]

from which, with little algebra, expression (23) is recovered:

\[
\tau = \frac{h}{2u} \left( \coth(P_{e_k}) - \frac{1}{P_{e_k}} \right).
\]


[31] R. Codina, Stabilization of incompressibility and convection through or-


66


67


