

2D inverse design of linear transport equations on unstructured grids

Mario Morales-Hernández, Enrique Zuazua

October 31, 2017

1 Adjoint estimation: low or high order?

Adjoint methods have been systematically associated to the optimal control design [5] and their applications to aerodynamics [1, 4]. During the last decades, several works were oriented to develop a robust control theory based on the concepts of observability, optimality and controllability for linear and non-linear equations and systems of equations [2, 10, 11, 12].

Since the discrete approach forces to use the discrete adjoint problem of the flow solver to numerically solve the adjoint equation, the continuous approach is adopted due to its valuable flexibility of using different solvers for the flow and adjoint equations. It is well known that problems where high frequencies play an important role such as the control of wave equation [11]

The focus is put on the 2D linear scalar transport equation, that can be expressed in a conservative form as follows:

$$\frac{\partial u(x, y, t)}{\partial t} + \nabla \cdot (\mathbf{v}u) = 0 \quad u(x, y, 0) = u_0 \quad (1.1)$$

where $\mathbf{v} = v(x, y)$ is a time-independent velocity field of propagation.

Given a target function $u^* = u^*(x, y)$, the aim in the 2D inverse design problem consists in finding u_0 such that $u(T) \sim u^*$ via the minimization of a functional J :

$$J(u) = \frac{1}{2} |u(T) - u^*|^2 \quad (1.2)$$

This problem can be easily addressed by simply solving the transport equation backwards in time because of the time reversibility of the model. But such a simple approach fails as soon as the model involves nonlinearities (leading to shock discontinuities) or diffusive terms, making the system time-irreversible.

We are thus interested in the development of gradient descent methods with the aid of the adjoint equation (solved backwards in time) that can be easily deduced

$$-\frac{\partial \sigma(x, y, t)}{\partial t} - \mathbf{v} \cdot \nabla \sigma = 0, \quad \sigma(T) = u(T) - u^* \quad (1.3)$$

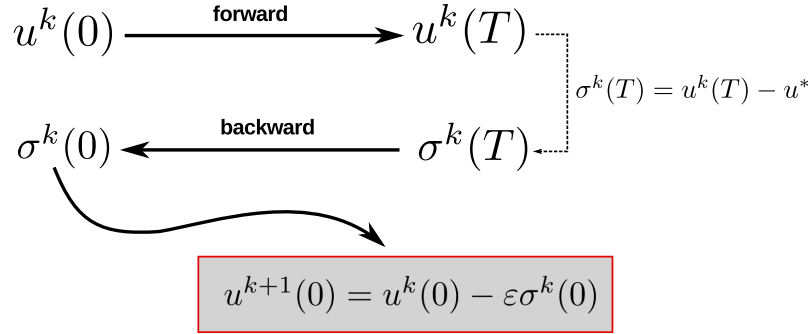


Figure 1.1: Iteration of the 2D inverse design

where $\sigma = \sigma(x, y, t)$ is the adjoint variable.

In order to achieve a good match with the continuous solutions, high order numerical schemes need to be used for the forward state equation (1.1). Here we shall use a second order scheme. Our main objective is to test the convenience of using the same order of accuracy when solving the adjoint equation or, by the contrary, to employ a low order one. When implementing the gradient descent iterations, the numerical scheme employed for solving the adjoint equation determines the direction of descent. Hence, different solvers for the adjoint system provide different results that can be compared in terms of accuracy and efficiency.

A gradient-adjoint iterative method is based on iterating a loop where the equation of state (flow equation) is solved in a forward sense while the adjoint equation, which is of hyperbolic nature as well, is solved backwards in time (see Figure 1.1).

The adequate resolution of this loop is the main question addressed in this work, paying attention not only to accuracy but also to reducing its computational complexity. One could expect that the choice of high order numerical methods to solve both the equation of state and the adjoint one should provide the best results in terms of accuracy, at the prize of a high computational cost. But this is not always true and it is possible to relax the necessity of using the same order of accuracy for the resolution of the adjoint equation, not only reducing considerably the computational time needed by the complete loop but also achieving the same order of accuracy on the approximation of the inverse design, which is our ultimate goal.

2 2D linear equation and continuous adjoint derivation

We focus on the following conservative linear transport scalar equation with an heterogeneous time-independent vector field $\mathbf{v} = \mathbf{v}(x, y)$:

$$\frac{\partial u}{\partial t} + \nabla \cdot (\mathbf{v}u) = 0. \quad (2.1)$$

Obviously, the solution $u = u(x, y, t)$ exists and it is unique and can be determined by means of the method of characteristics provided \mathbf{v} is smooth enough (say, C^1). Thus, for all initial data $u_0 \in L^2(\mathbb{R}^2)$ there exists a unique solution in the class $C([0, T]; L^2(\mathbb{R}^2))$.

Let us consider the inverse design problem: Given a target function $u^* = u^*(x, y)$ at $t = T$, to determine the initial condition u_0 such that $u(x, y, T) \equiv u^*(x, y)$.

The problem could be easily addressed solving the equation (2.1) backwards in time from the final datum u^* at $t = T$, to determine $u_0 = u(x, y, 0)$ exactly. But we are interested in addressing it from the point of view of optimal control. Let us therefore define the following functional, as a classical measure of the quadratic error with respect to the target function u^* :

$$J(u) = \frac{1}{2} \int_{\Omega} (u(T) - u^*)^2 dS \quad (2.2)$$

The derivation of the adjoint equation can be achieved by simply multiplying by $\sigma = \sigma(x, y, t)$ and integrating over a control volume $\Omega \times [0, T]$

$$\int_0^T \int_{\Omega} \sigma \left(\frac{\partial u}{\partial t} + \nabla \cdot (\mathbf{v}u) \right) dS dt = 0 \quad (2.3)$$

Integrating by parts

$$\int_{\Omega} \sigma u|_0^T dS - \int_0^T \int_{\Omega} u \frac{\partial \sigma}{\partial t} dS dt + \int_0^T \oint_{\partial \Omega} u \sigma \mathbf{v} \mathbf{n} dl dt - \int_0^T \int_{\Omega} u \mathbf{v} \nabla \sigma dS dt = 0 \quad (2.4)$$

where $\partial \Omega$ is the boundary and \mathbf{n} is the outward normal direction to the surface Ω respectively.

It is feasible to take Gateaux derivatives with respect to the variable u in this identity getting:

$$\int_{\Omega} \sigma \delta u|_0^T dS - \int_0^T \int_{\Omega} \delta u \frac{\partial \sigma}{\partial t} dS dt + \int_0^T \oint_{\partial \Omega} \delta u \sigma \mathbf{v} n dS dt - \int_0^T \int_{\Omega} \delta u \mathbf{v} \cdot \nabla \sigma dS dt = 0 \quad (2.5)$$

Gathering appropriately the terms:

$$\int_0^T \int_{\Omega} \delta u \overbrace{\left(-\frac{\partial \sigma}{\partial t} - \mathbf{v} \cdot \nabla \sigma \right)}^{(\#)} dS dt + \int_{\Omega} \sigma \delta u|_0^T dS + \int_0^T \oint_{\partial \Omega} \delta u \sigma \mathbf{v} n dS dt = 0 \quad (2.6)$$

Let us select $(\#) = 0$

$$-\frac{\partial \sigma}{\partial t} - \mathbf{v} \cdot \nabla \sigma = 0 \quad (2.7)$$

and appropriate boundary conditions for $(\#\#) = 0$. Hence (2.6) becomes

$$\int_{\Omega} [\sigma(T) \delta u(T) - \sigma(0) \delta u(0)] dS = 0 \quad (2.8)$$

Assuming $\sigma(T) = u(T) - u^*$

$$\int_{\Omega} (u(T) - u^*) \delta u(T) dS = \int_{\Omega} \sigma(0) \delta u(0) dS \quad (2.9)$$

On the other hand, it is also feasible to take the first variation of J with respect to u in (2.2)

$$\delta J(u) = \int_{\Omega} \delta u(T) (u(T) - u^*) dS \quad (2.10)$$

Finally,

$$\delta J(u) = \int_{\Omega} (u(T) - u^*) \delta u(T) dS = \int_{\Omega} \sigma(0) \delta u(0) dS \quad (2.11)$$

This derivation led to the adjoint equation $(\#)$:

$$-\frac{\partial \sigma}{\partial t} - \mathbf{v} \cdot \nabla \sigma = 0 \quad \sigma(T) = u(T) - u^*, \quad (2.12)$$

$\sigma = \sigma(x, y, t)$ being the adjoint variable.

This equation measures the sensitivity of the solution to changes in the initial condition. It is worth mentioning that the adjoint equation is solved backwards in time (from $t = T$ to $t = 0$). In fact, system (2.12) is well-posed if and only if the original system is well posed in the forward sense.

3 Numerical schemes

In order to obtain a numerical solution based on a finite volume approach, (2.1) is integrated in a control volume $\Omega_i \times [t^n, t^{n+1}]$ where Ω_i represents each computational cell of the domain and $t^{n+1} = t^n + \Delta t$

$$\int_{t^n}^{t^{n+1}} \int_{\Omega_i} \left(\frac{\partial u}{\partial t} + \nabla \cdot (\mathbf{v}u) \right) dS dt = 0 \quad (3.1)$$

Using the Gauss divergence theorem

$$\int_{\Omega_i} u(x, y, t^{n+1}) dS - \int_{\Omega_i} u(x, y, t^n) dS + \int_{t^n}^{t^{n+1}} \oint_{\partial\Omega_i} \mathbf{f} \mathbf{n} dm dt = 0 \quad (3.2)$$

where $\mathbf{f} = \mathbf{v}u$ is the flux, $\partial\Omega_i$ denotes the surface surrounding the volume Ω_i and $\mathbf{n} = (n_x, n_y)$ its unit normal vector. The last contour integral in (3.2) can be replaced by the sum of the fluxes defined at each edge k of length l_k :

$$\int_{\Omega_i} u(x, y, t^{n+1}) dS - \int_{\Omega_i} u(x, y, t^n) dS + \int_{t^n}^{t^{n+1}} \sum_{k=1}^{N_E} \mathbf{f}_k \mathbf{n}_k l_k dt = 0 \quad (3.3)$$

with N_E is the number of edges of each cell i ($N_E = 3$ is for triangular grids) with neighbouring cells j_k . Figure 3.1 clarifies the meaning of each variable.

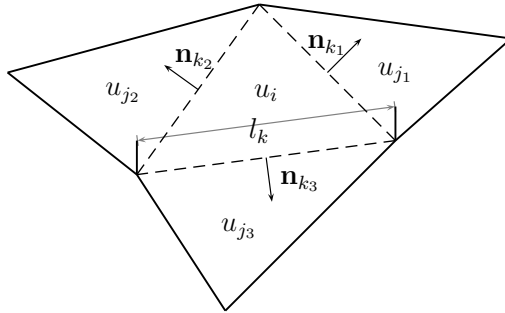


Figure 3.1: Sketch of the 2D finite volume approach

In this work two numerical schemes are proposed: a first order upwind (FOU) scheme and a second order upwind (SOU) scheme based on MUSCL-Hancock approach. The details of the implementation can be found in [7] and

[6, 9] respectively. The use of a SOU scheme for the flow equation (forward) is mandatory in order to achieve the best accuracy since the FOU scheme is unable to capture the detail for the equation of state when trying to reach the target function. The issue of using one or another approach for solving the adjoint equation and, consequently, the gradient direction approach, is worth analysing.

4 Test case: Doswell frontogenesis

This test case was firstly proposed (in a forward sense) in [3]. It symbolizes the presence of horizontal temperature gradients and fronts in the context of meteorological dynamics. The original problem consists of a computational domain $\Omega = [-5, 5]$ with open boundary conditions in which a front defined by the following initial condition:

$$u(x, y, 0) = \tanh\left(\frac{y}{\delta}\right) \quad (4.1)$$

is advected under the velocity field given by

$$\mathbf{v} = (-y g(r), x g(r)) \quad g(r) = \frac{1}{r} \bar{v} \operatorname{sech}^2(r) \tanh(r) \quad \bar{v} = 2.59807 \quad (4.2)$$

where $r = \sqrt{x^2 + y^2}$. The exact solution (the projection or restriction of the analytical solution on the computational mesh) on the whole space \mathbb{R}^2 has the following explicit form:

$$u(x, y, t) = \tanh\left(\frac{y \cos(gt) - x \sin(gt)}{\delta}\right) \quad (4.3)$$

It is important to mention that the velocity field is divergence-free. The state equation can be written similarly in divergence or non-divergence form:

$$\frac{\partial u}{\partial t} + \nabla \cdot (\mathbf{v}u) = \frac{\partial u}{\partial t} + (\nabla \cdot \mathbf{v})u = 0, \quad (4.4)$$

because

$$\nabla \cdot \mathbf{v} \equiv 0,$$

This allows us to formulate the problem of inverse design. Given the exact value of the solution at time $t = T$, i.e. giving (4.3) as a target function at $t = T$, we aim to recover by a computational version of the gradient-adjoint methodology described above, an accurate and efficient approximation of the initial datum (4.1).

In our numerical experiments we take the values $\delta = 1$ and $T = 4s$. The exact initial condition and the exact target are shown in Figure 4.1 (left and right respectively). As can be seen, the target presents a vortex-type profile, due to the heterogeneous geometry of the velocity field.

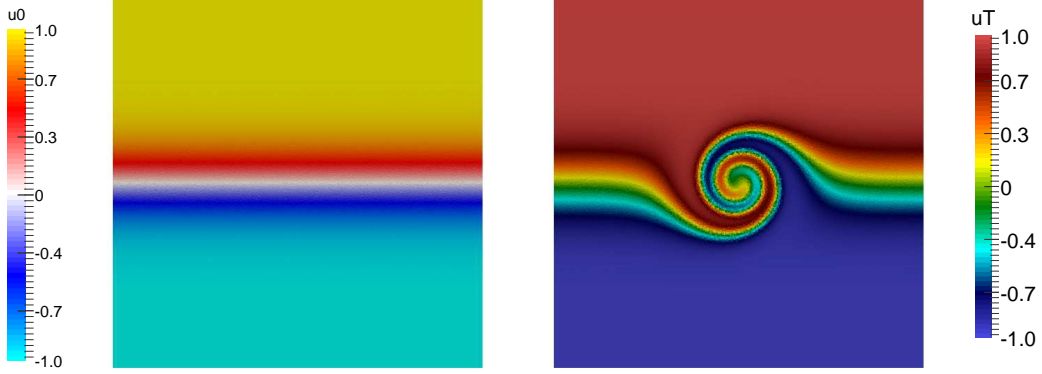


Figure 4.1: Doswell frontogenesis: Exact initial condition (left) and target function (right) for the inverse problem

Our goal is to build a numerical method capable of predicting accurately and efficiently an approximation of (4.1) out of the target (4.3) by means of a numerical version of the optimisation algorithm above.

The domain is discretized in 39998 unstructured triangles using *Triangle* [8] in a uniform Delaunay mesh. The initial condition for the first iteration is $u^0(x, y, 0) = 0$. Both the forward and backward resolution are parallelized with OPENMP in 4 cores. The maximum number of iterations for the gradient method is set to 75, a constant step $\varepsilon = 0.5$ is selected and the Courant number is set to $CFL = 0.5$.

Additionally, the step size ε is set as a constant at the beginning of each iteration. However, if the convergence is not guaranteed, i.e., $J^{k+1} \geq J^k$, the step size is halved until $J^{k+1} < J^k$. Therefore, the stopping criteria for the iterative algorithm are:

1. To reach the maximum number of iterations
2. To achieve the maximum allowable error of the functional, in this case, 10^{-9}
3. To descend below $\varepsilon_{min} = 10^{-6}$ when halving the step size ε trying to achieve convergence.

The numerical results are shown in Figures 4.2, 4.3 and 4.4 for the target $u(x, y, T)$, initial condition $u(x, y, 0)$ and adjoint $\sigma(x, y, 0)$ respectively at the last iteration of the gradient method. On the left side the ∇J^{FOU} approach is displayed while the ∇J^{SOU} resolution is illustrated on the right side.

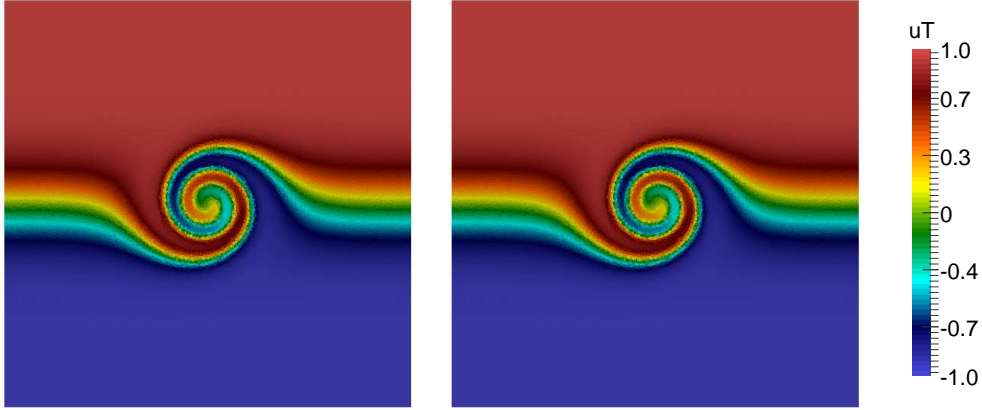


Figure 4.2: Doswell frontogenesis: Numerical target $u(x, y, T)$ achieved by ∇J^{FOU} (left) and ∇J^{SOU} (right)

In order to complete the qualitative results, the error computed as the difference between the exact and the numerical approximation is displayed in Figures 4.5 and 4.6 for the target $u(x, y, T) - u^*(x, y)$ and the initial datum $u(x, y, 0) - u^e(x, y)$.

Quantitative results can be obtained by means of the computation of L_2 and L_∞ norms, not only with respect to the target function but also to the known exact initial condition u^e in order to evaluate the results achieved by ∇J^{FOU} and ∇J^{SOU} :

$$\begin{aligned}
 L_2(u^T) &= \sqrt{\sum_i^N (u(x_i, T) - u^*(x_i))^2} & L_\infty(u^T) &= \max_i |u(x_i, T) - u^*(x_i)| \\
 L_2(u^0) &= \sqrt{\sum_i^N (u(x_i, 0) - u^e(x_i))^2} & L_\infty(u^0) &= \max_i |u(x_i, 0) - u^e(x_i)|
 \end{aligned}
 \tag{4.5}$$

Table 4.1 condenses this information as well as the number of iterations and the CPU time for each proposed scheme. For the sake of clarity and for each comparison, the best results are highlighted in bold font.

Note that the number of iterations for the ∇J^{SOU} approach does not reach the maximum number of iterations set at the beginning of the inverse design process because at iteration 45, the functional is not able to diminish

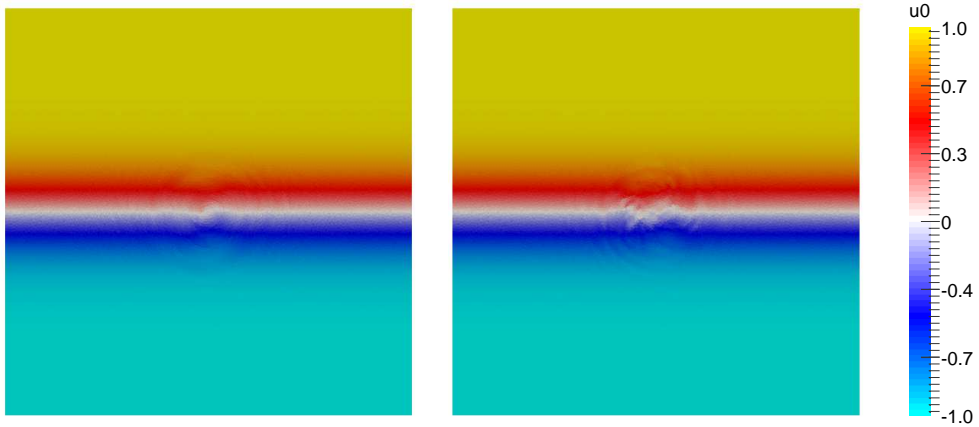


Figure 4.3: Doswell frontogenesis: Numerical initial condition $u(x, y, 0)$ achieved by ∇J^{FOU} (left) and ∇J^{SOU} (right)

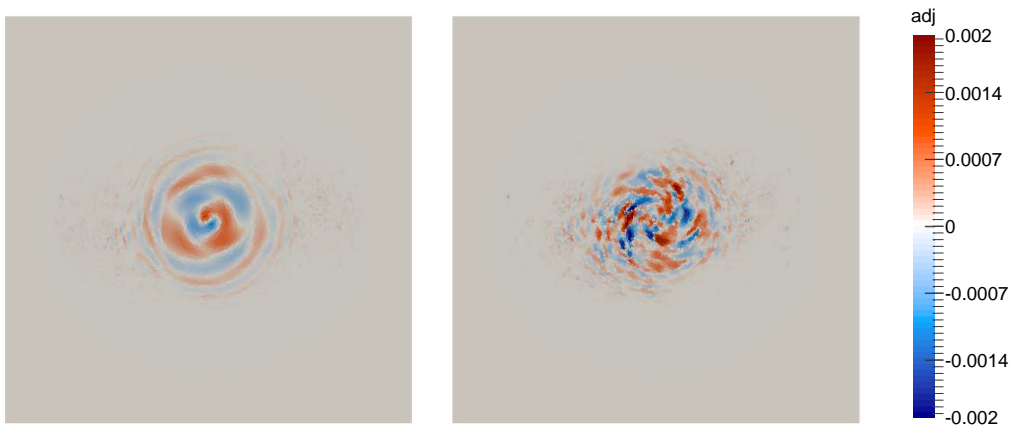


Figure 4.4: Doswell frontogenesis: Numerical adjoint variable $\sigma(x, y, 0)$ achieved by ∇J^{FOU} (left) and ∇J^{SOU} (right)

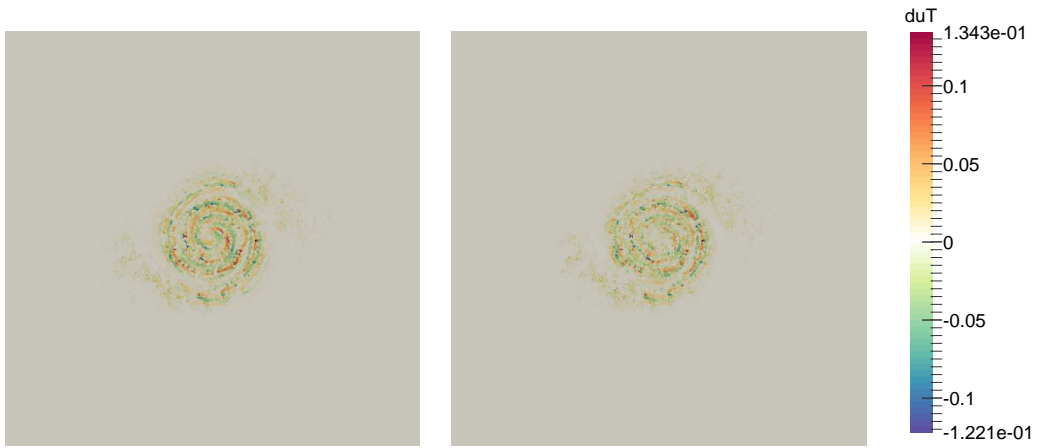


Figure 4.5: Doswell frontogenesis: Target error $u(x, y, T) - u^*(x, y)$ achieved by ∇J^{FOU} (left) and ∇J^{SOU} (right)

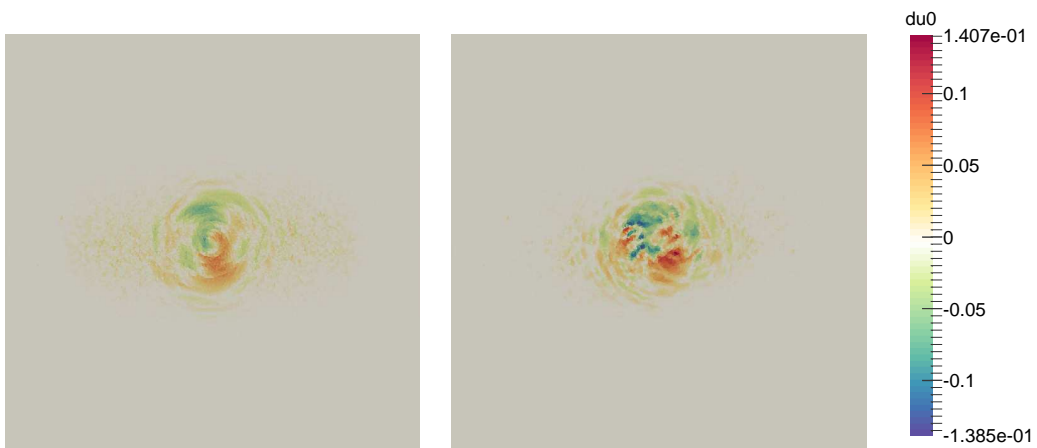


Figure 4.6: Doswell frontogenesis: Initial datum error $u(x, y, 0) - u^e(x, y)$ achieved by ∇J^{FOU} (left) and ∇J^{SOU} (right)

Scheme	$L_2 (u^T)$	$L_\infty (u^T)$	$L_2 (u^0)$	$L_\infty (u^0)$	nIters	CPU Time (s)
FOU	1.37127e+00	1.37078e-01	1.36043e+00	8.64870e-02	75	366.26
SOU	1.23889e+00	1.34271e-01	1.88749e+00	1.40718e-01	45	488.61

Table 4.1: Doswell Frontogenesis: L_2 and L_∞ norms, number of iterations and CPU time for the ∇J^{FOU} and ∇J^{SOU}

its value to the next iteration due to the deterioration of the initial condition by the high frequencies.

Bibliography

- [1] C. CASTRO, C. LOZANO, F. PALACIOS, E. ZUAZUA, *Systematic Continuous Adjoint Approach to Viscous Aerodynamic Design on Unstructured Grids*, AIAA Journal, 45-9 (2007) 2125-2139
- [2] C. CASTRO, F. PALACIOS, E. ZUAZUA, *An alternating descent method for the optimal control of the inviscid Burgers equation in the presence of shocks*, Mathematical Models and Methods in Applied Sciences, 18 (2008), 369-416
- [3] C.A. DOSWELL, *A kinematic analysis of frontogenesis associated with a nondivergent vortex*, J. Atmos. Sci. 41 (1984), 1242-1248
- [4] M.B. GILES AND N.A. PIERCE, *An Introduction to the Adjoint Approach to Design*, Flow, Turbulence and Combustion, 65 (2000), 393-415
- [5] M. HERTY, A. KURGANOV, D. KUROCHKIN, *Numerical method for optimal control problems governed by nonlinear hyperbolic systems of PDE's*, Commun. Math. Sci, 13 No.1 (2015) 15-48
- [6] M.E. HUBBARD, *Multidimensional Slope Limiters for MUSCL-Type Finite Volume Schemes on Unstructured Grids*, Journal of Computational Physics, 155 (1999), 54-74
- [7] MORALES-HERNÁNDEZ M. GARCÍA-NAVARRO P., BURGUETE J. AND BRUFAU P., *A conservative strategy to couple 1D and 2D models for shallow water flow simulation*, Computers & Fluids 81, (2013) 26-44.
- [8] J.R. SHEWCHUK, *Delaunay Refinement Algorithms for Triangular Mesh Generation*, Computational Geometry: Theory and Applications 22, (2002) 21-74.
- [9] P.K. SWEBY, *High resolution schemes using flux-limiters for hyperbolic conservation laws*, SIAM J. Num. Anal. 21 (5) (1984) 995-1011.
- [10] E. ZUAZUA, *Controllability of partial differential equations and its semi-discrete approximations*, Discrete and continuous dynamical systems, 8 (2), (2002) 469-513

- [11] E. ZUAZUA, *Propagation, observation, and control of waves approximated by finite difference methods*, SIAM Review, 47 (2), (2005) 197-243
- [12] E. ZUAZUA, *Controllability and Observability of Partial Differential Equations: Some results and open problems*, Handbook of differential equations: evolutionary equations, 3, (2007) 527-621