Time-Consistent Filtering in Spectral and Spectral Element Methods

Alex Kanevsky [a], Mark H. Carpenter [b], and Jan S. Hesthaven [a]

[a] *Division of Applied Mathematics, Brown University, Box F, Providence, Rhode Island 02912, USA*

[b] *Aeronautics and Aeroacoustic Methods Branch, NASA Langley Research Center, Hampton, VA 23681-0001, USA*

E-mail: kanevsky@dam.brown.edu;Mark.H.Carpenter@nasa.gov;Jan.Hesthaven@Brown.edu

Abstract

The comparison of numerical results for semi-implicit [17] and fully explicit Runge-Kutta [16] time integration methods for a nozzle flow problem shows that filtering can significantly degrade the accuracy of the numerical solution for long-time integration problems. We demonstrate analytically and numerically that filtering-in-time errors become additive for $\|u_N(x,t+k\Delta t) - u_N(x,t)\| \ll \|u_N(x,t)\|$ when nonconsistent filters are used, and suggest the development and implementation of time-consistent filters.

Key Words: modal filters; exponential; sharp-cutoff; spectral element method; discontinuous Galerkin method; high-order accuracy; implicit-explicit; Runge-Kutta; time-consistent; Euler equations; staircase effect; net filter.

1 Introduction

In the last decade, significant attention has been paid to modal filtering in spectral methods [2, 3, 8, 9]. Filtering is popular in spectral and spectral element-type methods for several reasons. Firstly and most importantly, it stabilizes the numerical approximation and results in a more robust method. Furthermore, for discontinuous functions, filtering can recover high-order accuracy at the points of discontinuity [11, 10] and in the smooth regions away from the discontinuity [21]. In the early 1990s, Gottlieb, Shu, Solomonoff and Vandeven [11] showed that the Gibbs phenomenon, which is associated with the reconstruction of discontinuous functions, could be overcome by accelerating the rate of convergence of the reconstruction using Gegenbauer polynomials. Since then, a lot of work has followed along similar lines. A recent review of filtering in spectral methods can be found in [8, 12].

However, there are still many unresolved issues related to filtering, e.g., it is not clear how to choose a filter for the problem at hand. What filter order should one use? Should the filter be applied once or more per time step or perhaps once every several time steps? What is the effect of applying a filter repeatedly on the accuracy of our approximation?

We aim to address some of these issues in this paper, and give some guidelines concerning how filters should be designed and applied in practice for spectral and spectral element methods. We will restrict our theory and numerical examples to two types of low-pass filters [4]: the sharp-cutoff or step-function filter used in classical dealising methods [3] and the exponential filter. However, the ideas and analysis presented may be applied to a more general class of filters.

One of the outcomes of comparing numerical test results using explicit Runge-Kutta (ERK) time-integration methods [16] to those using implicit-explicit (IMEX-RK) methods [17] is the realization that filtering may severely degrade the accuracy of the approximation when applied for a large number of time-steps. We can see in Fig. 1.1 that the IMEX-RK tests, which were run at time steps an order of magnitude larger (on average) than the ERK

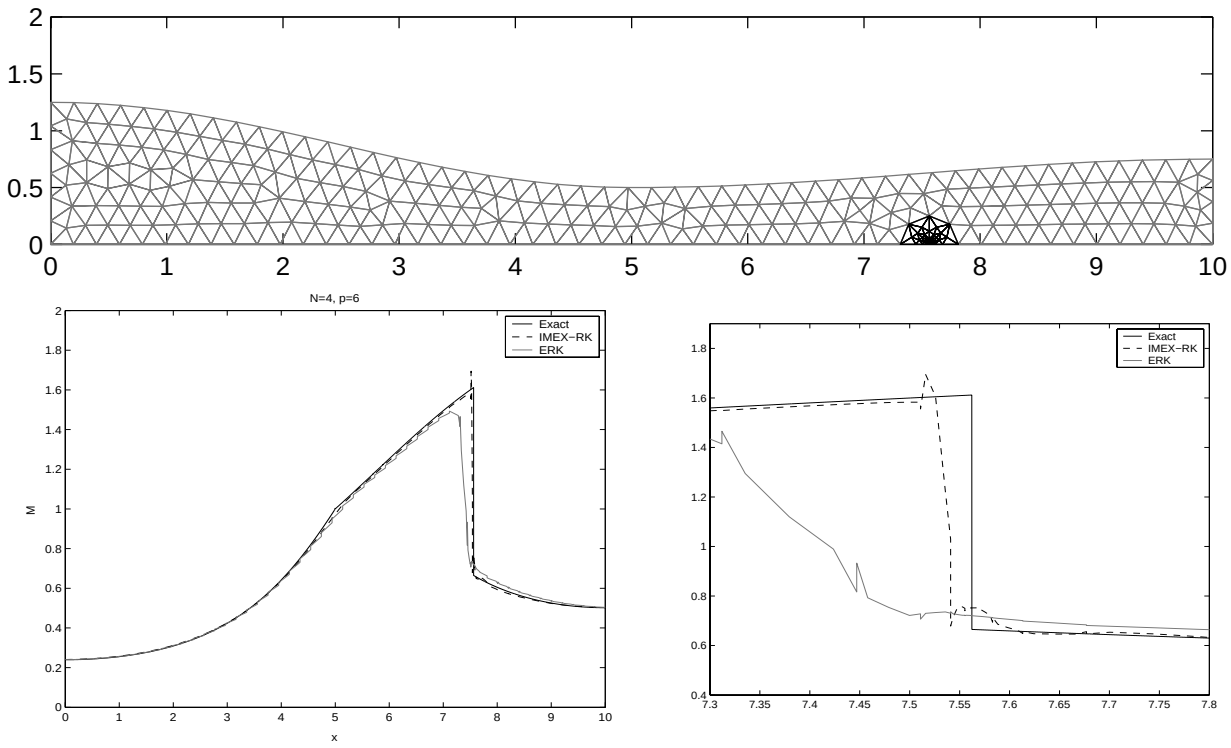

Figure 1.1: The top image is the mesh used for the two-dimensional steady-state converging-diverging nozzle flow tests. The black region is solved implicitly when using the IMEX-RK scheme, and explicitly when using the ERK scheme. The exact steady-state solution to the Euler equations has a shock at $x \cong 7.56$. The two figures below compare the Mach number profiles at the centerline of the nozzle, $y = 0$, for the ERK, IMEX-RK, and analytic solutions. The bottom-right figure is a close-up of the shock region in the bottom-left figure. The polynomial degree $N = 4$ on each element and the order of the exponential filter $p = 6$.

tests ($\Delta t_{ERK} \approx 8.56 \times 10^{-4}, \Delta t_{IMEX-RK} \approx 6.69 \times 10^{-3}$), show superior accuracy, especially in the case of relatively low-order exponential filters [15]. The details of the physical and numerical setup for the nozzle flow test case appear in the Appendix.

In Section 2, we review the underlying theory behind filtering in spectral methods. In Section 3, we develop a set of tools for analyzing the net effect of filtering on the accuracy of the numerical approximation, and conjecture that filtering a numerical approximation $\|u_N(x, t + k\Delta t) - u_N(x, t)\| \ll \|u_N(x, t)\|$ results in an additive filtering process. This observation leads us to suggest that time-dependent, time-consistent filters be developed to control filtering-in-time errors.

We propose nearly consistent exponential filters in Section 4, and construct filters whose

filter order, $p(t)$, is a function of time. We then show that if $p(t)$ can be properly controlled, we can overcome the potentially additive net effect of filtering in time. The control strategy for $p(t)$ is based on the worst-case scenario, purely additive filtering.

Finally, we discuss the theoretical and numerical results in Section 5.

2 Filtering

Our goal is to approximate the exact solution $u(x,t)$ of a conservation law in the form

$$\frac{\partial u(x,t)}{\partial t} + \frac{\partial f(u(x,t))}{\partial x} = 0. \tag{2.1}$$

We express $u(x,t)$ as an infinite series of basis functions $\phi_n(x)$

$$u(x,t) = \sum_{n=0}^{\infty} \hat{u}_n(t)\phi_n(x). \tag{2.2}$$

In spectral methods, we project $u(x,t)$ to the finite-dimensional space $P_N \in \{\phi_n(x)\}_{n=0}^{N}$ and get the truncated approximate solution

$$u_N(x,t) = \mathcal{P}_{\mathcal{N}} u(x,t) = \sum_{n=0}^{N} \hat{u}_n(t)\phi_n(x). \tag{2.3}$$

$\mathcal{P}_{\mathcal{N}} u(x,t) \in P_N$, where P_N is spanned by the smooth basis functions $\phi_n(x)$, which form an L^2-complete basis.

For nonlinear problems, the nonlinear interaction of modes often results in nonlinear instability and leads to the unbounded growth of the high frequency energy in time. We can

add a term to our original PDE (2.1) that dissipates the high-frequency energy components and therefore controls the instability

$$\frac{\partial u(x,t)}{\partial t} + \frac{\partial f(u(x,t))}{\partial x} = \epsilon(-1)^{p+1}\frac{\partial^{2p} u}{\partial x^{2p}}. \tag{2.4}$$

Although adding artificial dissipation to (2.1) will stabilize the method, it is costly and may introduce restrictions on the stable time step.

Instead, we follow the approach originally introduced in [18, 19] and add dissipation by applying a modal filter to the numerical approximation at regular intervals. The filtered approximation is

$$\mathcal{F}_N u_N(x,t) = \sum_{n=0}^{N} \sigma\left(\frac{n}{N}\right) \hat{u}_n(t)\phi_n(x), \tag{2.5}$$

where $\sigma(\eta)$ is the filter kernel.

Let us introduce two commonly used filter functions, which we will refer to in subsequent sections:

1. *Exponential Filter*

$$\sigma\left(\frac{n}{N}\right) = \begin{cases} 1, & 0 \le n \le N_c \\ \exp\left[-\alpha\left(\frac{n-N_c}{N-N_c}\right)^p\right], & N_c < n \le N \end{cases}, \tag{2.6}$$

   where p is the order of the filter, $\alpha = -\log\epsilon$ (ϵ is the machine zero), and N_c is the cutoff mode.

2. *Sharp-Cutoff* or *Step-Function Filter*

$$\sigma\left(\frac{n}{N}\right) = \begin{cases} 1, & 0 \le n \le N_c \\ 0, & N_c < n \le N \end{cases}, \tag{2.7}$$

where N_c is the cutoff mode.

3 From Multi-Modes To Uni-Mode

Let us assume that $\|u_N(x,t+\Delta t) - u_N(x,t)\| \ll \|u_N(x,t)\| \Longrightarrow u_N(x,t+\Delta t) \approx u_N(x,t)$. For now, we assume that the time step, Δt, is constant, and that the filter kernel, $\sigma(\eta)$, does not change with time.

After applying the filter once at the end of the time step, the filtered approximation becomes:

$$\begin{aligned} \tilde{u}_N(x,t+\Delta t) &= \mathcal{F}_{\mathcal{N}} u_N(x,t+\Delta t) \cong \mathcal{F}_{\mathcal{N}} u_N(x,t) && (3.8) \\ &= \sum_{n=0}^{N} \sigma\left(\frac{n}{N}\right) \hat{u}_n(t)\phi_n(x). && (3.9) \end{aligned}$$

If we filter once again at the end of the next time step, we have

$$\begin{aligned} \tilde{u}_N(x,t+2\Delta t) &= \mathcal{F}_{\mathcal{N}} u_N(x,t+2\Delta t) \cong \mathcal{F}_{\mathcal{N}} \tilde{u}_N(x,t+\Delta t) && (3.10) \\ &\cong \mathcal{F}_{\mathcal{N}}(\mathcal{F}_{\mathcal{N}} u_N(x,t)) = \sum_{n=0}^{N} \sigma\left(\frac{n}{N}\right)\left(\sigma\left(\frac{n}{N}\right)\hat{u}_n(t)\phi_n(x)\right) && (3.11) \\ &= \sum_{n=0}^{N} \sigma^2\left(\frac{n}{N}\right)\hat{u}_n(t)\phi_n(x) && (3.12) \\ &= {\mathcal{F}_{\mathcal{N}}}^2 u_N(x,t). && (3.13) \end{aligned}$$

Repeating this process k times, and assuming that $\|u_N(x,t+k\Delta t)-u_N(x,t)\| \ll \|u_N(x,t)\|$ we have

$$\begin{aligned}\tilde{u}_N(x,t+k\Delta t) \cong {\mathcal{F}_{\mathcal{N}}}^k u_N(x,t) &= \sum_{n=0}^{N} \sigma^k\left(\frac{n}{N}\right)\hat{u}_n(t)\phi_n(x) &(3.14)\\ &= \sum_{n=0}^{N} \tilde{\sigma}_{k\Delta t}\left(\frac{n}{N}\right)\hat{u}_n(t)\phi_n(x). &(3.15)\end{aligned}$$

The net effect of filtering k times is represented by the *net filter* $\tilde{\sigma}_{k\Delta t}(\eta)$, which we define as

$$\tilde{\sigma}_{k\Delta t}(\eta) = \sigma^k(\eta). \tag{3.16}$$

We now assume that our filter kernel is an exponential filter with $N_c = 0$

$$\sigma(\eta) = \exp\left(-\alpha\eta^p\right). \tag{3.17}$$

Therefore

$$\begin{aligned}\tilde{u}_N(x,t+k\Delta t) \cong {\mathcal{F}_{\mathcal{N}}}^k u_N(x,t) &= \sum_{n=0}^{N} \sigma^k\left(\frac{n}{N}\right)\hat{u}_n(t)\phi_n(x) &(3.18)\\ &= \sum_{n=0}^{N} \exp\left(-\alpha k\left(\frac{n}{N}\right)^p\right)\hat{u}_n(t)\phi_n(x) &(3.19)\end{aligned}$$

i.e. the net filter becomes

$$\tilde{\sigma}_{k\Delta t}(\eta) = \sigma^k(\eta) = \exp\left(-\alpha k\eta^p\right). \tag{3.20}$$

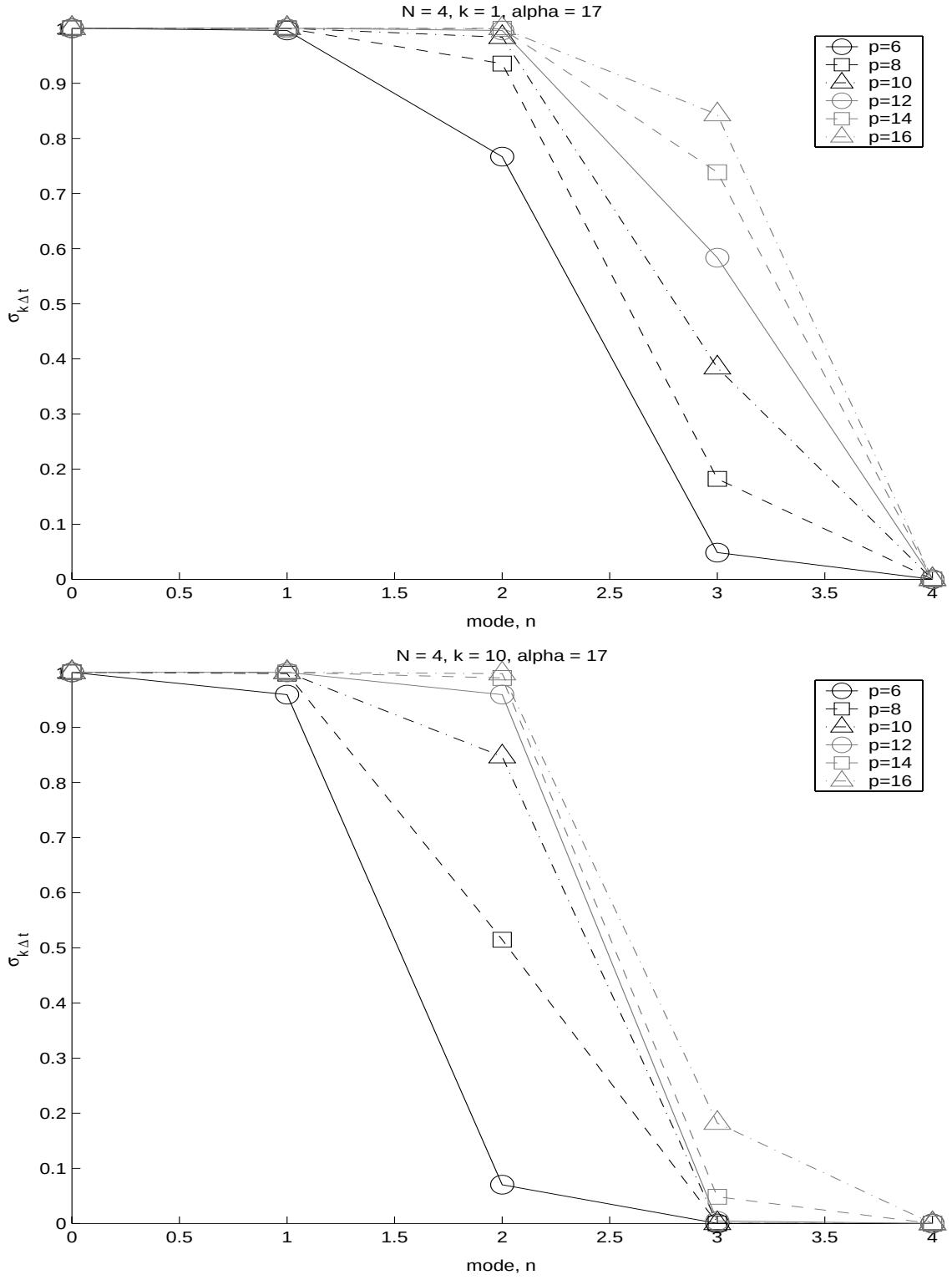

Figure 3.2: Net filter strength versus mode number for nonconsistent exponential filter (3.20) of order $p = 6, 8, 10, 12, 14, 16$. Number of time steps $k = 1$ (top), 10 (bottom). Polynomial degree $N = 4$, $\alpha = 17$.

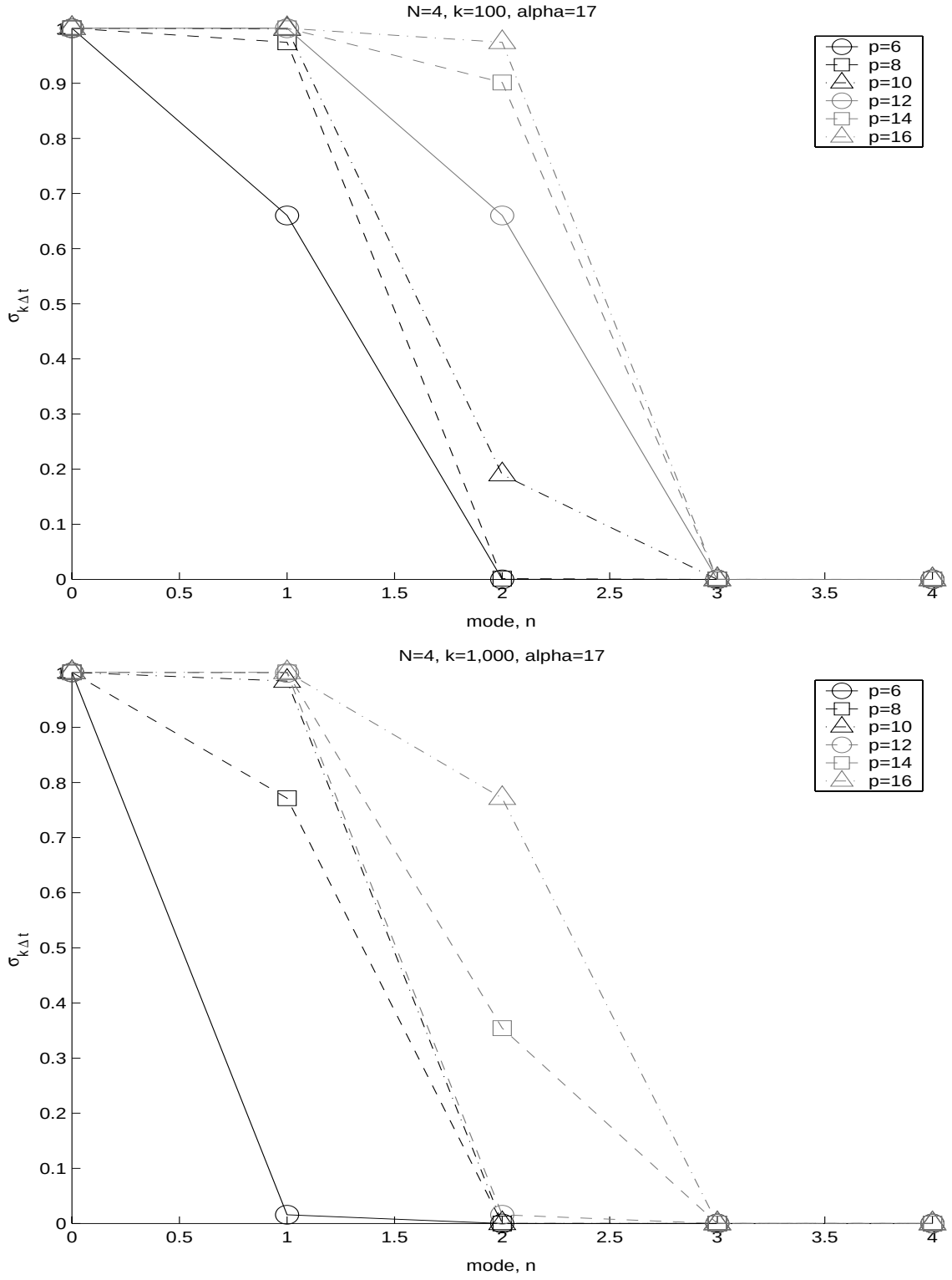

Figure 3.3: Net filter strength versus mode number for nonconsistent exponential filter (3.20) of order $p = 6, 8, 10, 12, 14, 16$. Number of time steps $k = 100$ (top), $1,000$ (bottom). Polynomial degree $N = 4$, $\alpha = 17$.

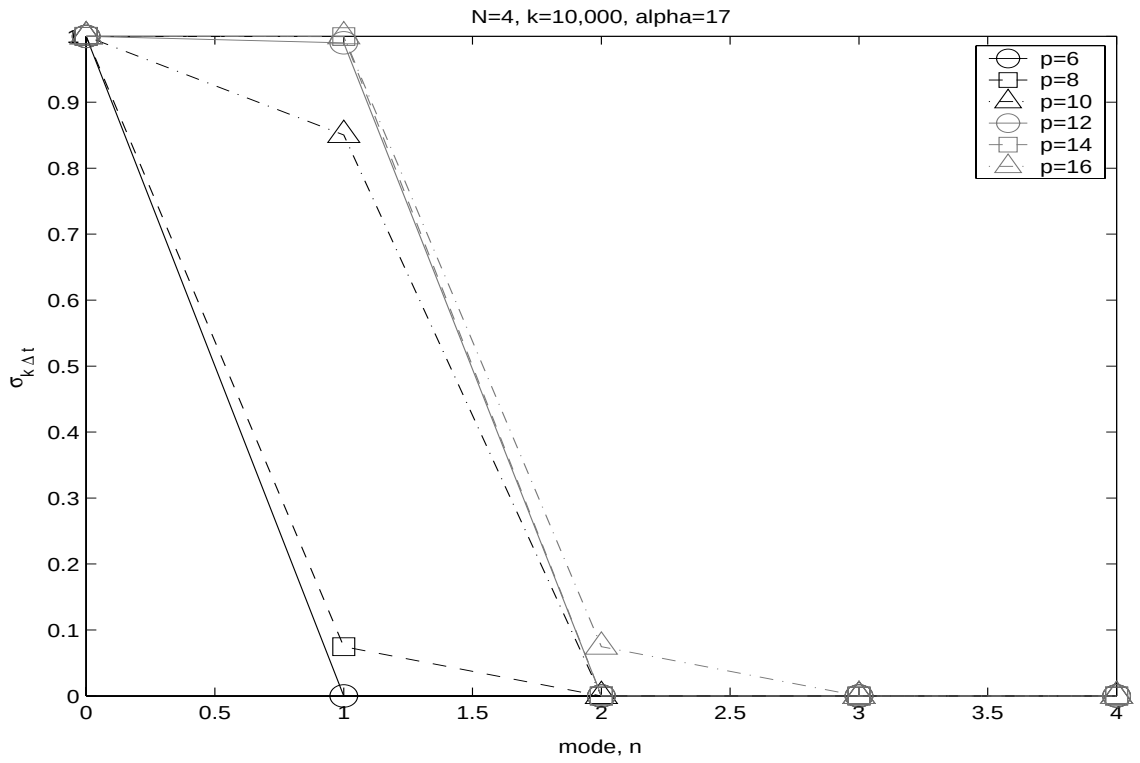

Figure 3.4: Net filter strength versus mode number for nonconsistent exponential filter (3.20) of order $p = 6, 8, 10, 12, 14, 16$. Number of time steps $k = 10,000$. Polynomial degree $N = 4$, $\alpha = 17$.

Filtering repeatedly results in an additive process, under the assumptions stated at the beginning of this section. In fact, a purely additive filtering process, which is represented by the net filter, is the upper-bound on filtering time-dependent problems, resulting in a net filter kernel that grows larger with each time step, and is equal to σ^k after k steps. We can clearly see from this analysis that as the number of time steps k grows large, which we expect in long-time simulations, the net filter kernel can have an extremely strong, crippling effect on the accuracy of the numerical solution.

For spectral element methods, it is common to use polynomial approximations of degree $N = 4$ to $N = 16$ on each subdomain. Filtering regularly can effectively zero out many of the modes, thereby potentially reducing polynomial approximations of degree $N = 4$ to $N = 16$ to much lower-degree polynomials. The same holds for classical spectral methods with polynomial degree $N = 128, 256, 512...$, although it is more difficult to "see" the loss in

accuracy.

In order to validate our analysis, we perform the following numerical experiment. We repeat the nozzle flow test case using small time steps, $\Delta t = 1 \times 10^{-6}$, and polynomials of degree $N = 4$ on each element. We use an initial condition, which is very far from the exact steady-state solution (the IC is a linear profile connecting the inflow and outflow BCs), and are therefore solving a time-dependent problem. Since the time steps are very small, the solution will change very slowly with respect to the time step number, k, and should adhere to the above theory. We apply a nonconsistent exponential filter (3.17) and plot the results after $k = 10,\ 100,\ 1,000$ and $10,000$ time steps.

Figs. 3.5-3.7 show that the nonconsistent exponential filter degrades the accuracy of the numerical solution, just as we predicted above. In fact, we see a staircase phenomenon develop. After $k = 10,000$ time steps, the approximations, which originally start out as polynomials of degree $N = 4$, are filtered into 0th-degree polynomials on each subdomain (piecewise-constants) by the nonconsistent filter of order $p = 6$ and $p = 8$, resulting in a solution that looks flat on each element (Figs. 3.5-3.6). The staircase-looking results from this numerical experiment support Figs. 3.2-3.4, which show plots of the net filter versus mode number for the nonconsistent exponential filter based on (3.20) for varying number of time steps. Fig. 3.4, corresponding to $k = 10,000$ time steps, shows that the filters of order $p = 6$ and $p = 8$ zero out virtually all but the first mode for polynomials of degree four. Even the $p = 16$ nonconsistent exponential filter zeros out all but the first 2 modes after 10,000 steps.

Additional evidence of the staircase phenomenon can be seen in the ERK Mach number profile in Fig. 1.1. The onset of staircasing is evident in the ERK approximation (gray) which was integrated for $46,741$ time steps. The IMEX-RK approximation (dashed) was integrated for $5,975$ time steps and does not exhibit staircasing at this scale, although it does exhibit staircasing at higher magnifications.

After contemplating the potentially severe consequences of filtering, it is natural to ask

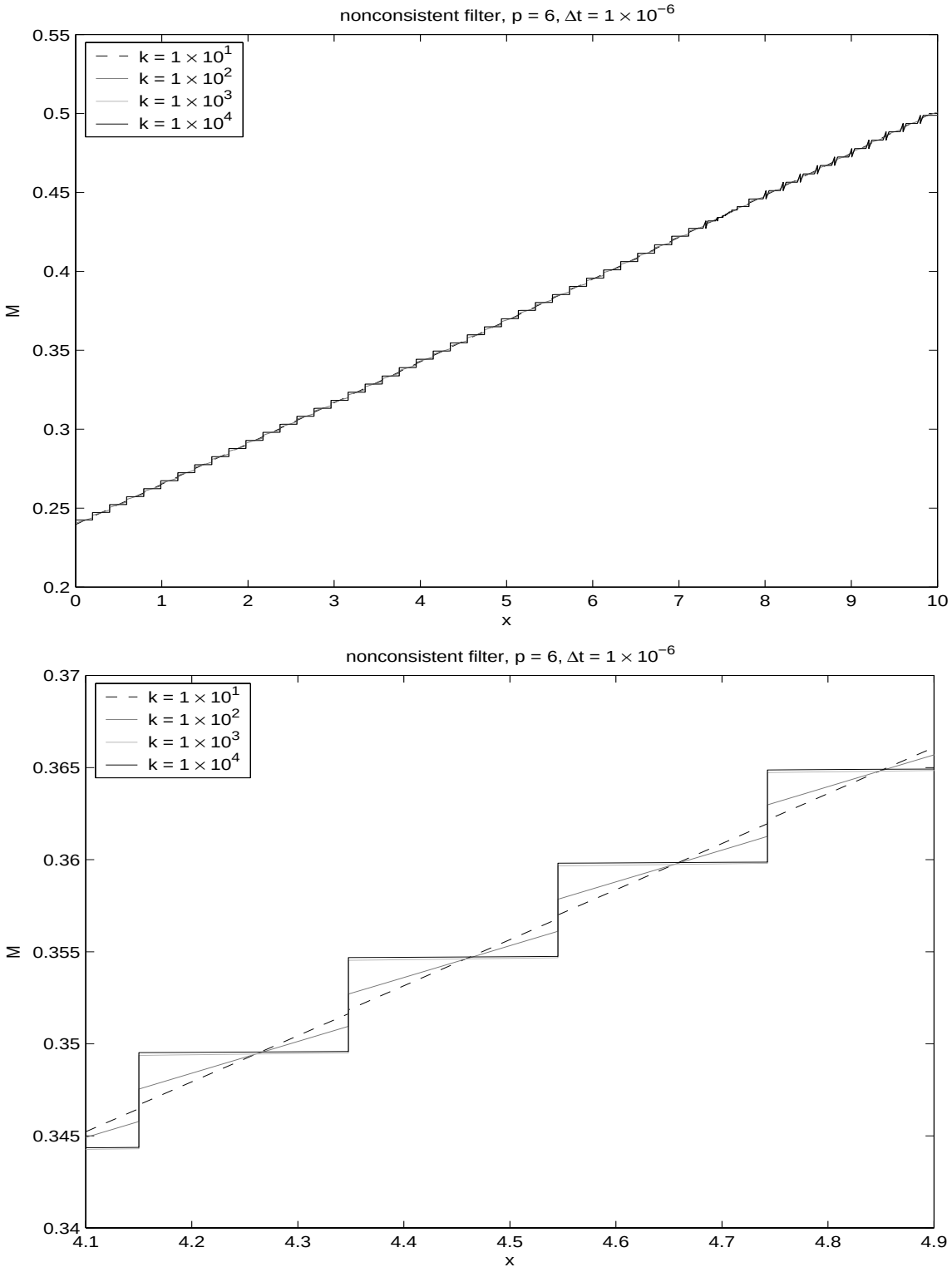

Figure 3.5: Staircase effect for nonconsistent exponential filter with $\Delta t = 1 \times 10^{-6}$ and filter order $p = 6$. Bottom plot is a blow-up of the top plot, and shows the staircase effect in more detail. All figures show the Mach number profile at the centerline of the nozzle, $y = 0$.

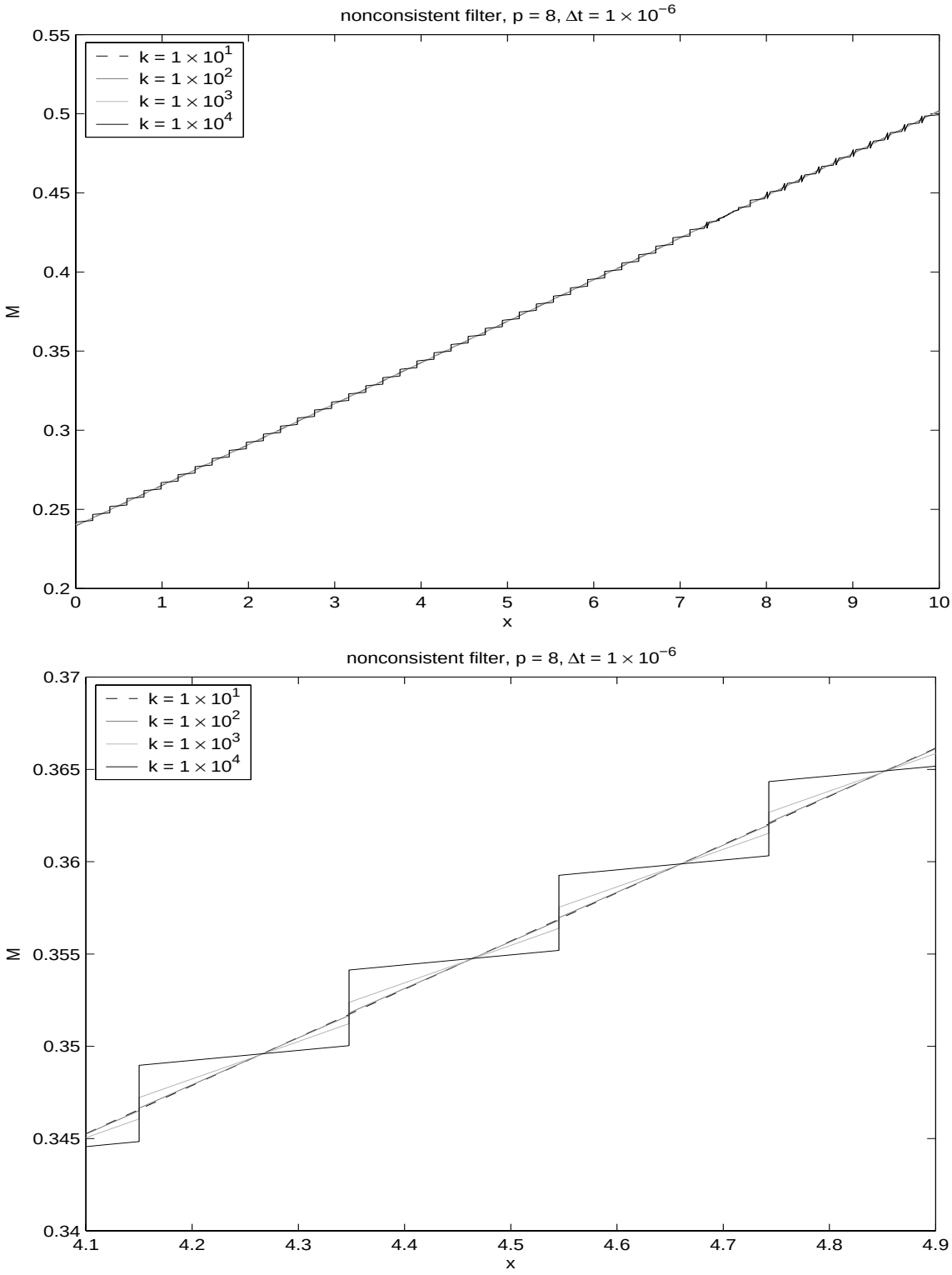

Figure 3.6: Staircase effect for nonconsistent exponential filter with $\Delta t = 1 \times 10^{-6}$ and filter order $p = 8$. Bottom plot is a blow-up of the top plot, and shows the staircase effect in more detail. All figures show the Mach number profile at the centerline of the nozzle, $y = 0$.

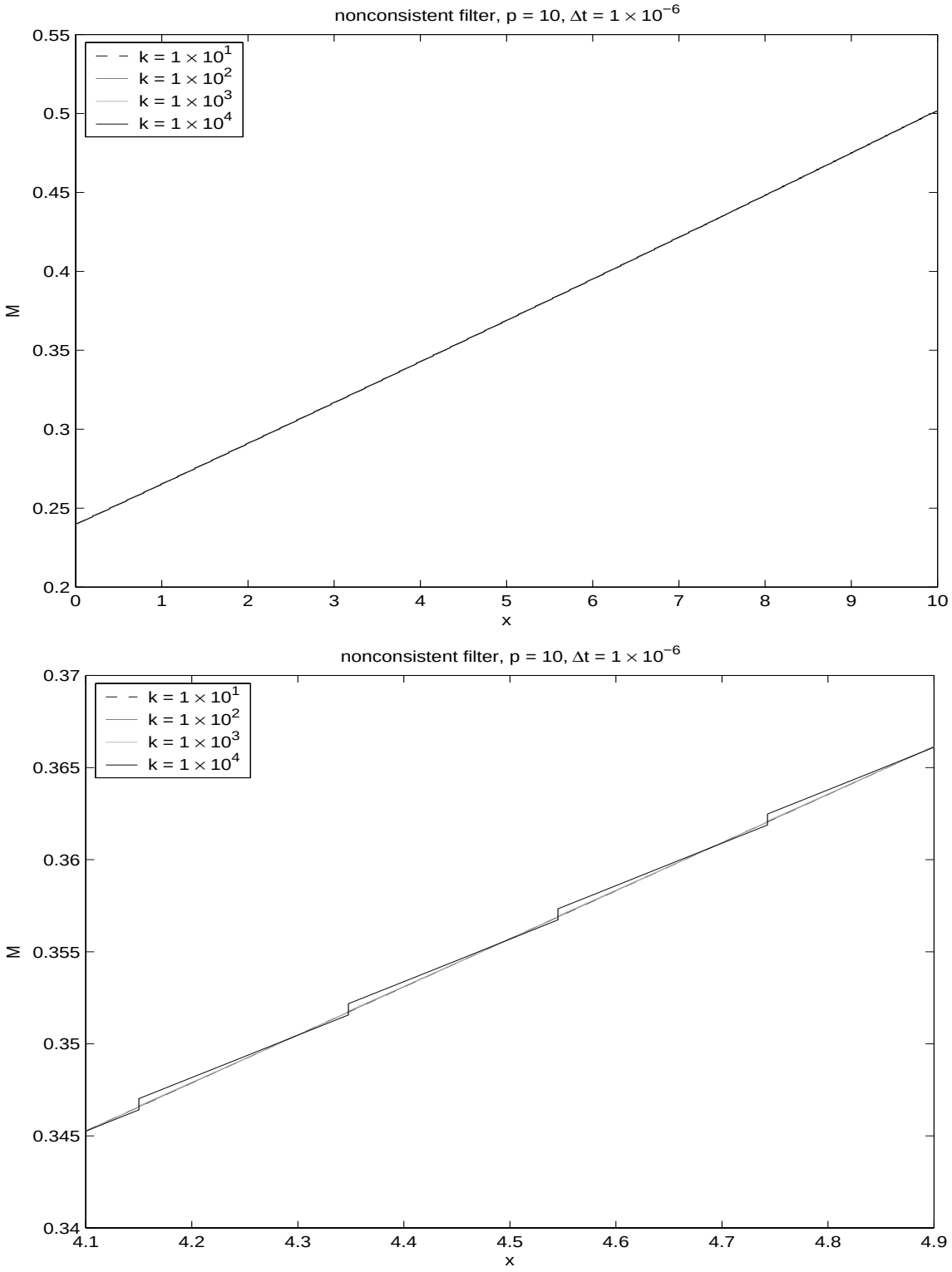

Figure 3.7: Staircase effect for nonconsistent exponential filter with $\Delta t = 1 \times 10^{-6}$ and filter order $p = 10$. Bottom plot is a blow-up of the top plot, and shows the staircase effect in more detail. All figures show the Mach number profile at the centerline of the nozzle, $y = 0$.

the following question: Is there any way we can counterbalance the net effect of filtering in long-time simulations?

4 Time-Consistent Filters

Definition: Time-Consistent Filter

A time-consistent filtering process is one for which the net filter after k_1 steps is equal to the net filter after k_2 steps is equal to the net filter after one step

$$\tilde{\sigma}_{k_1 \Delta t}(\eta) = \tilde{\sigma}_{k_2 \Delta t}(\eta) = \tilde{\sigma}_{\Delta t}(\eta). \tag{4.21}$$

Equivalently, a time-consistent filtering process will result in the same net filter at final time $t = k_1 \Delta t_1 = k_2 \Delta t_2 = T$

$$\tilde{\sigma}_{k_1 \Delta t_1}(\eta) = \tilde{\sigma}_{k_2 \Delta t_2}(\eta). \tag{4.22}$$

If the filter is truly time-consistent, it should have the same net effect after one million time steps as it does after one time step.

In the case of resolved, smooth solutions, we are able to achieve a time-consistent filtering process by choosing a *sharp-cutoff* or *step-function* filter

$$\sigma\left(\frac{n}{N}\right) = \begin{cases} 1, & 0 \le n \le N_c \\ 0, & N_c < n \le N \end{cases}, \tag{4.23}$$

Proof: The net filter is

$$\tilde{\sigma}_{k\Delta t}\left(\frac{n}{N}\right) = \sigma^k\left(\frac{n}{N}\right) \tag{4.24}$$

$$= \begin{cases} 1^k, & 0 \le n \le N_c \\ 0^k, & N_c < n \le N \end{cases} \tag{4.25}$$

$$= \begin{cases} 1, & 0 \le n \le N_c \\ 0, & N_c < n \le N \end{cases} \tag{4.26}$$

$$= \sigma\left(\frac{n}{N}\right) \tag{4.27}$$

$$= \tilde{\sigma}_{\Delta t}\left(\frac{n}{N}\right). \tag{4.28}$$

The maximum possible filtering-in-time error after k steps is equal to the maximum possible filtering-in-time error after one step.

We perform the same exact numerical experiment we carried out for the nonconsistent exponential filter, to test the consistent sharp-cutoff filter defined above. We use $N_c = N-1$, polynomial degree $N = 4$ and $\Delta t = 1 \times 10^{-6}$. We can see from the results in Fig. 4.8 that the sharp-cutoff filter is indeed time-consistent and does not result in staircasing, since it only cuts off the highest mode $n = N$. However, if the solution is smooth and well-resolved, filtering should not be used.

In the case of discontinuous solutions, we would like to recover high-order accuracy in smooth regions away from the discontinuity. Also, in the case of under-resolved smooth solutions, we want to retain the high rate of convergence. Vandeven [21] showed that the following sufficient conditions need to be met in order to guarantee this. Let $\sigma(\eta) \in C^\infty : R^+ \to [0,1]$ such that

$$\sigma(\eta) : \begin{cases} \sigma(0) = 1 & \\ \sigma^{(k)}(0) = 0 & k \le p \\ \sigma(1) = 0 & \\ \sigma^{(k)}(1) = 0 & k \le p \end{cases} \tag{4.29}$$

We now consider the exponential filter function. The exponential filter does not meet all of the conditions in (4.29). However, it appears to recover high-order accuracy in practice [12].

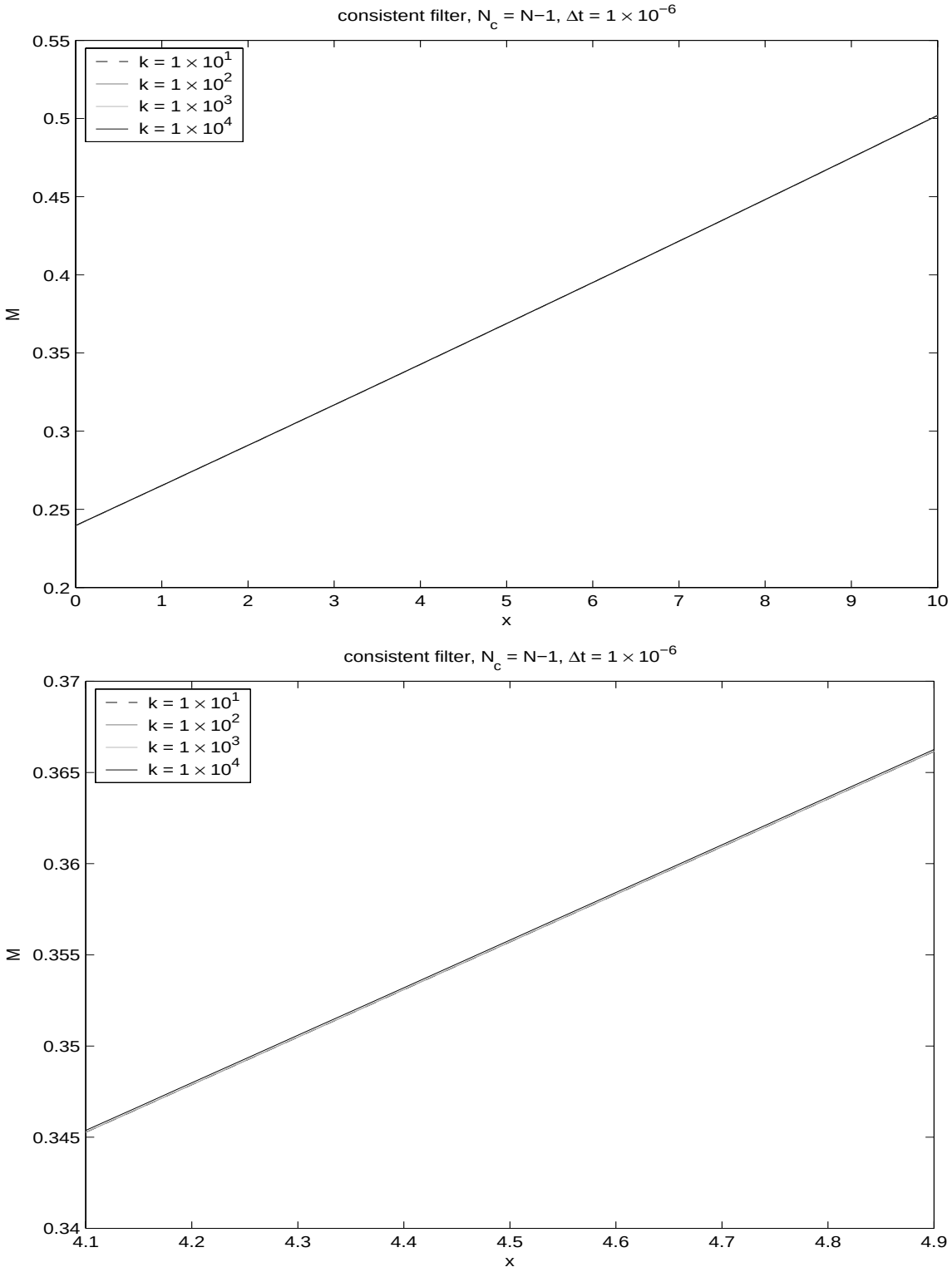

Figure 4.8: Numerical experiment for the consistent sharp-cutoff filter. $N_c = N - 1$, polynomial degree $N = 4$ and $\Delta t = 1 \times 10^{-6}$. There is no staircasing. Both figures show the Mach number profile at the centerline of the nozzle, $y = 0$.

We propose the following: Let us make our filter kernel time-dependent by allowing the order of the filter, p, to be a function of time

$$p = p(t). \tag{4.30}$$

The time-dependent exponential filter now becomes

$$\sigma\left(\frac{n}{N}\right) = \begin{cases} 1, & 0 \le n \le N_c \\ \exp\left[-\alpha\left(\frac{n-N_c}{N-N_c}\right)^{p(t)}\right], & N_c < n \le N \end{cases}. \tag{4.31}$$

For generality let $p_i : p_0, p_1, ..., p_{k-1}$ be a sequence of filter orders such that p_0 is the order of the filter applied after integrating one time step, while p_{k-1} is the order of the filter after integrating k time steps. We assume that the parameters α, N are constant for all times. The sequence of filter kernels is

$$\sigma_i\left(\frac{n}{N}\right) = \begin{cases} 1, & 0 \le n \le N_c \\ \exp\left[-\alpha\left(\frac{n-N_c}{N-N_c}\right)^{p_i}\right], & N_c < n \le N \end{cases}, \quad i = 0, ..., k-1. \tag{4.32}$$

After k time steps, our filtered approximation becomes

$$\begin{aligned} \tilde{u}_N(x, t + k\Delta t) &= \sum_{n=0}^{N} (\sigma_0\sigma_1\sigma_2...\sigma_{k-1})\hat{u}_n(t)\phi_n(x) && (4.33) \\ &= \sum_{n=0}^{N} (\prod_{i=0}^{k-1} \sigma_i)\hat{u}_n(t)\phi_n(x) && (4.34) \\ &= \begin{cases} \sum_{n=0}^{N} \hat{u}_n(t)\phi_n(x), & n \le N_c \\ \sum_{n=0}^{N} \exp\left[-\alpha \sum_{i=0}^{k-1} \left(\frac{n-N_c}{N-N_c}\right)^{p_i}\right] \hat{u}_n(t)\phi_n(x), & n \le N \end{cases} && (4.35) \\ &= \sum_{n=0}^{N} \tilde{\sigma}_{k\Delta t}\hat{u}_n(t)\phi_n(x). && (4.36) \end{aligned}$$

Definition: Net Filter

In general, the net filter is defined as

$$\tilde{\sigma}_{k\Delta t}(\eta) = \prod_{i=0}^{k-1} \sigma_i(\eta). \tag{4.37}$$

The net filter is therefore

$$\tilde{\sigma}_{k\Delta t}\left(\frac{n}{N}\right) = \begin{cases} 1, & n \leq N_c \\ \exp\left[-\alpha \sum_{i=0}^{k-1} \left(\frac{n-N_c}{N-N_c}\right)^{p_i}\right], & n \leq N \end{cases}. \tag{4.38}$$

We now ask ourselves the following question: How should we choose the filter order p_i after each time step? Initially, $p = p_0$ and the inital filter kernel is

$$\sigma_0\left(\frac{n}{N}\right) = \begin{cases} 1, & 0 \leq n \leq N_c \\ \exp\left[-\alpha \left(\frac{n-N_c}{N-N_c}\right)^{p_0}\right], & N_c < n \leq N \end{cases}. \tag{4.39}$$

In order for our filtering process to remain time-consistent, we must conserve the initial net filter

$$\tilde{\sigma}_{k\Delta t}(\eta) = \tilde{\sigma}_{\Delta t}(\eta). \tag{4.40}$$

Therefore, we must choose the sequence p_i such that

$$\exp\left[-\alpha \sum_{i=0}^{k-1} \left(\frac{n-N_c}{N-N_c}\right)^{p_i}\right] = \exp\left[-\alpha \left(\frac{n-N_c}{N-N_c}\right)^{p_0}\right], \quad n = N_c+1, ..., N, \tag{4.41}$$

or

$$\sum_{i=0}^{k-1}\left(\frac{n-N_c}{N-N_c}\right)^{p_i}=\left(\frac{n-N_c}{N-N_c}\right)^{p_0},\quad n=N_c+1,...,N. \tag{4.42}$$

However, we can see that the left-hand side is greater than the right-hand side, and that the filtering process cannot be purely time-consistent according to the definition

$$\left(\frac{n-N_c}{N-N_c}\right)^{p_0}+\sum_{i=1}^{k-1}\left(\frac{n-N_c}{N-N_c}\right)^{p_i}\geq\left(\frac{n-N_c}{N-N_c}\right)^{p_0}, \tag{4.43}$$

since $\frac{n-N_c}{N-N_c}\geq 0$. Nevertheless, we add coefficients a_k to the right hand side, which can allow us to make the process nearly time-consistent

$$\sum_{i=0}^{k-1}\left(\frac{n-N_c}{N-N_c}\right)^{p_i}=a_k\left(\frac{n-N_c}{N-N_c}\right)^{p_0}. \tag{4.44}$$

We can rewrite this as

$$\left(\frac{n-N_c}{N-N_c}\right)^{p_{k-1}}+\sum_{i=0}^{k-2}\left(\frac{n-N_c}{N-N_c}\right)^{p_i}=a_k\left(\frac{n-N_c}{N-N_c}\right)^{p_0}. \tag{4.45}$$

Now solving for p_{k-1}

$$p_{k-1}=\frac{\log\left(a_k\left(\frac{n-N_c}{N-N_c}\right)^{p_0}-\sum_{i=0}^{k-2}\left(\frac{n-N_c}{N-N_c}\right)^{p_i}\right)}{\log\left(\frac{n-N_c}{N-N_c}\right)},\quad k\geq 2,\quad N_c+1\leq n\leq N. \tag{4.46}$$

We have an expression for p_{k-1} which depends on $n=N_c+1,...,N$, N_c, p_0 and a_k. In order to retain the structure of the original filter function, which in this case happens to be

the exponential filter, we cannot make p_{k-1} a function of n. In other words, we should not compute a different value of p_{k-1} for each of the $N+1$ values of n. Instead, we need to choose one value of n which will be used to compute p_{k-1}. Let us refer to this value of n as n_*.

The time-dependent exponential filter becomes

$$\sigma_{k-1}\left(\frac{n}{N}\right) = \begin{cases} 1, & 0 \le n \le N_c \\ \exp\left[-\alpha\left(\frac{n-N_c}{N-N_c}\right)^{p_{k-1}}\right], & N_c < n \le N \end{cases}. \tag{4.47}$$

$$p_{k-1} = \frac{\log\left(a_k\left(\frac{n_*-N_c}{N-N_c}\right)^{p_0} - \sum_{i=0}^{k-2}\left(\frac{n_*-N_c}{N-N_c}\right)^{p_i}\right)}{\log\left(\frac{n_*-N_c}{N-N_c}\right)}. \tag{4.48}$$

Proposition: Time-Consistent Exponential Filter

The exponential filter described by (4.47-4.48) is time-consistent for coefficients

$$a_k = \lim_{m\to\infty} k^{1/m}$$

Proof: Let $a_k = k^{1/m}$

$$\sum_{i=0}^{k-1}\left(\frac{n-N_c}{N-N_c}\right)^{p_i} = a_k\left(\frac{n-N_c}{N-N_c}\right)^{p_0} = k^{1/m}\left(\frac{n-N_c}{N-N_c}\right)^{p_0}. \tag{4.49}$$

Now taking the limit $\lim_{m\to\infty}$ of (4.49)

$$\sum_{i=0}^{k-1}\left(\frac{n-N_c}{N-N_c}\right)^{p_i} = \left(\frac{n-N_c}{N-N_c}\right)^{p_0}. \tag{4.50}$$

Therefore, the net filter is conserved since

$$\lim_{m\to\infty} \tilde{\sigma}_{k\Delta t}(\eta) = \tilde{\sigma}_{\Delta t}(\eta).$$

Choosing $a_k = k^{1/m}, m \gg 1$ may not be practical due to stability reasons. A better choice may be $m = \mathcal{O}(1)$, which will make the numerical scheme more stable and nearly consistent.

The optimal choice of a_k, p_0, N_c and n_* for both accuracy and stability is still an open problem.

5 Conclusions and Future Work

We have identified the mechanism by which filtering can become an additive process in this paper. Filtering a numerical approximation $\|u_N(x, t + k\Delta t) - u_N(x, t)\| \ll \|u_N(x, t)\|$ will result in a growing filtering-in-time error which will erase modes from the Fourier or polynomial approximation when nonconsistent filters are used. Purely additive filtering (4.37) is the worst-case scenario for time-dependent problems, and will occur if the time step is very small (e.g. due to severe stable time-step restrictions, etc...), the solution is at or near steady-state, or a combination thereof. The theory developed in Section 3 holds for time-dependent problems as long as $k\Delta t \ll t_c$, where t_c is the characteristic time-scale.

In general, filtering will erode the accuracy of the numerical approximation, but at a slower rate than that defined by the net filter (purely additive) as we can see in Fig 1.1. We conjecture that the level of additivity is a function of the time step Δt. The slight staircasing in Fig. 1.1 for $k = 46,741$ time steps looks very much like that in Fig. 3.3 for $k = 100$ time steps for the filter of order $p = 6$, corresponding to the net filter $\tilde{\sigma}_{100\Delta t}$. This can be seen in Fig. 5.5, where the results from Fig. 1.1 and Fig. 3.3 at the section of the nozzle where the Mach number profiles have a similar slope are plotted. The ratio of number of time steps is of the order of the ratio of the average time step sizes $\frac{46,741}{100} = \mathcal{O}\left(\frac{8.56\times10^{-4}}{1\times10^{-6}}\right)$. Therefore, the filtering error for the $\Delta t \approx 8.56 \times 10^{-4}$ test is roughly $\frac{8.56\times10^{-4}}{1\times10^{-6}} = 856$ times less additive than that for the $\Delta t = 1 \times 10^{-6}$ purely additive test case. This means that a filtering process that is even 3 orders of magnitude less additive than a purely additive process still leads to significant loss of accuracy for long-time integration problems. A careful study needs

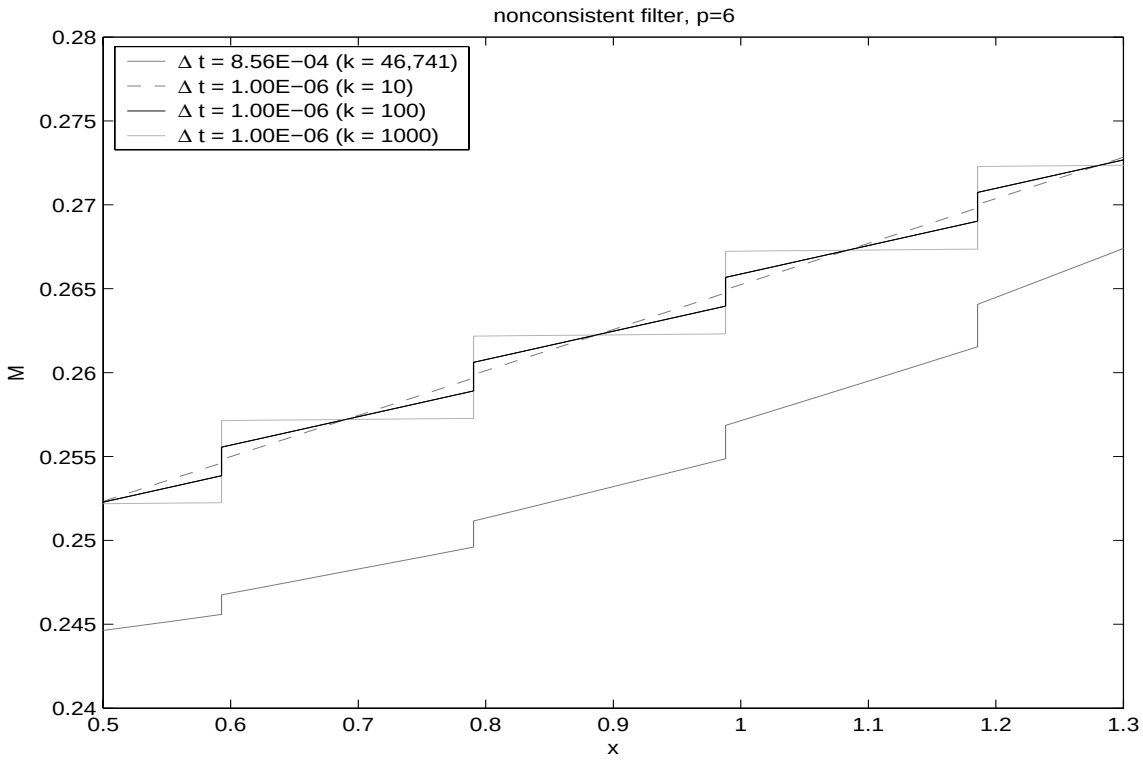

Figure 5.9: Staircase effect for nonconsistent exponential filters with filter order $p = 6$. The single gray solid curve is the ERK solution from Fig. 1.1, while the other 3 curves are the ERK test case results from Fig. 3.3.

be carried out to understand the levels of additivity of filtering errors for varying levels of $\delta = \|u_N(x, t + k\Delta t) - u_N(x, t)\|/\|u_N(x, t)\|$.

We conducted several numerical experiments that support the theory developed in Section 3. Nonconsistent exponential filters zero out all but the first couple of modes, and result in staircasing of the numerical solution, while the consistent sharp-cutoff filter only zeros out the highest mode $n = N$.

Consistent or nearly consistent filters need to be applied to prevent "additive" filtering . For smooth solutions, time-consistent filtering may be achieved by using sharp-cutoff filters, although sharp-cutoff filters are ill-advised for smooth solutions that are not adequately resolved [6, 4]. The situation is more involved for solutions with discontinuities and under-resolved smooth solutions.

Remarks:

- *For nonsmooth solutions and under-resolved smooth solutions, we suggest the development and application of time-dependent, nearly-consistent spectral filters, such as those described by (4.47) and (4.48).*
- *Filter at most once per time step to minimize filtering-in-time errors.*

In hindsight, we now understand that the ERK nozzle flow results presented in Fig. 1.1 were less accurate than the IMEX-RK results for two main reasons: (i) the time-steps taken were roughly 10 times smaller $\left(\frac{\Delta t_{ERK}}{\Delta t_{IMEX-RK}} \approx \frac{1}{10}\right)$ and therefore resulted in a more additive filtering error, and (ii) the ERK approximation was filtered roughly 10 times more since 10 times as many time steps were taken.

In the future, spatially [20] and temporally adaptive filters, for which the filter order $p = p(x,t)$ is a function of both space and time, need to be constructed for spectral and spectral element methods.

6 Appendix

We provide some background information about the numerical test case used throughout this paper.

6.1 Two-dimensional nozzle flows

Consider the two-dimensional Euler equations given in conservation form

$$\frac{\partial \mathbf{q}}{\partial t} + \nabla \cdot \mathbf{F}(\mathbf{q}) = 0, \quad t > 0. \tag{6.51}$$

The state vector $\mathbf{q}$ and the flux vector $\mathbf{F}(\mathbf{q})$ are given as

$$\mathbf{q} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ E \end{bmatrix}, \quad \mathbf{F}(\mathbf{q}) = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ (E+p)u \end{bmatrix} \hat{i} \quad + \quad \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ (E+p)v \end{bmatrix} \hat{j}, \tag{6.52}$$

where ρ is density, u and v are the Cartesian velocity components, E is the total energy, and p is the pressure. The total energy

$$E = \rho \left(T + \frac{1}{2}(u^2 + v^2) \right). \tag{6.53}$$

The pressure and temperature are related through the ideal gas law

$$p = (\gamma - 1)\rho T. \tag{6.54}$$

where T is the temperature and $\gamma = c_p/c_v$ is the ratio between the constant pressure (c_p) and constant volume (c_v) heat capacities. $\gamma = 1.4$ for air.

The converging-diverging nozzle (Fig. 1.1) has an area $A(x)$ given by

$$A(x) = \begin{cases} 1.75 - .75\cos\left((.2x - 1.0)\pi\right), & 0 \leq x \leq 5 \\ 1.25 - .25\cos\left((.2x - 1.0)\pi\right), & 5 \leq x \leq 10 \end{cases}. \tag{6.55}$$

This is a classic one-dimensional steady (steady-state), inviscid compressible flow problem that has an analytic solution [1]. A ratio between the stagnation pressure and the back pressure of .75 (back pressure/stagnation pressure) results in a choked flow with a stationary normal shock in the divergent part of the nozzle at $x \cong 7.56$. The Mach number $M = 1.0$ and the stagnation temperature $T = 300°K$ as the flow is choked.

6.2 Numerical Scheme

We follow the method of lines approach, and discretize the spatial operators using a nodal discontinuous galerkin collocation method based on [7, 13, 14].

The local approximation in each subdomain D is the Nth-degree polynomial

$$p_N(x) = \sum_{k=1}^{(N+1)(N+2)/2} p_N(x_k) L_k(x), \tag{6.56}$$

where x_k are the $(N+1)(N+2)/2$ Guass-Lobatto Legendre collocation points in each domain, $L_k(x), k \in [1, ..., (N+1)(N+2)/2]$ is the local polynomial basis, and the subdomains D are linear triangular elements. The flux is approximated as

$$F_N(p_N) = \sum_{k=1}^{(N+1)(N+2)/2} F(p_N(x_k)) L_k(x). \tag{6.57}$$

We require that the equation be satisfied on each element in the following discontinuous Galerkin way

$$\int_D (\frac{\partial p_N}{\partial t} + \nabla \cdot F_N) L_k(x) dx = \oint_{\delta D} L_k(x) \hat{n} \cdot [F_N - F_N^*] dx. \tag{6.58}$$

We use a local Lax-Friedrichs numerical flux

$$F_N^* = \frac{F_N(p_N^+) + F_N(p_N^-)}{2} - \frac{|\lambda|}{2}(p_N^+ - p_N^-), \tag{6.59}$$

where $|\lambda|$ is the maximum local eigenvalue of the flux Jacobian

$$|\lambda| = \max_{p_N^+, p_N^-} (|\mathbf{u}| + \mathbf{c}) = \max_{p_N^+, p_N^-} \left(\sqrt{u^2 + v^2} + \sqrt{\gamma p / \rho} \right). \tag{6.60}$$

p_N^+ is the local solution, while p_N^- is the solution in the neighboring element.

To integrate the resulting system of ODEs in time, we use a 4th-order low-storage explicit Runge-Kutta method [16] for all numerical experiments in this paper. For IMEX-RK results shown in Fig. 1.1, we use the 4th-order Additive Runge-Kutta scheme, ARK4(3) [17].

7 Acknowledgments

The work of AK was partly supported by NASA Graduate Student Researchers Program (GSRP) Fellowship NGT-1-01024 and by the NSF VIGRE Program. The work of MHC .... The work of JSH was partly supported by NSF Career Award DMS-0132967 and by the Alfred P. Sloan Foundation through a Sloan Research Fellowship. The authors also thank Professor W.-S. Don, Professor D. Gottlieb, Dr. S. J. Miller, R. Sendersky, and Professor C.-W. Shu for many illuminating discussions.

References

[1] J. D. Anderson, *Modern Compressible Flow*, McGraw-Hill. New York, 2002.

[2] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*, Second edition. Dover. New York, 2001.

[3] C. Canuto, M. Y. Hussaini, A. Quarteroni and T. A. Zang, *Spectral Methods in Fluid Dynamics*, Springer Series in Computational Physics. Springer-Verlag. New York, 1988.

[4] W.-S. Don, D. Gottlieb, C.-W. Shu, O. Schilling and L. Jameson, *Numerical Convergence Study of Nearly-Incompressible, Inviscid Taylor-Green Vortex Flow*, J. Sci. Comput. – to appear.

[5] M. Dubiner, *Spectral Methods on Triangles and Other Domains*, J. Sci. Comput. **6**(1993), pp. 345-390.

[6] P. F. Fischer and J. S. Mullen, *Filter-based Stabilization of Spectral Element Methods*, Comptes Rendus de l'Academie des sciences Paris, Serie I - Analyse numerique. **t. 332**(2001), pp. 265-270.

[7] F. X. Giraldo, J. S. Hesthaven and T. Warburton, *Nodal High-Order Discontinuous Galerkin Methods for the Spherical Shallow Water Equations*, J. Comput. Phys. **181**(2002), pp. 499-525.

[8] D. Gottlieb and J. S. Hesthaven, *Spectral Methods for Hyperbolic Problems*, J. Comp. Appl. Math. **128**(2001), pp. 83-131.

[9] D. Gottlieb and S. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications*, SIAM monograph, 1977.

[10] D. Gottlieb and C.-W. Shu, *On the Gibbs Phenomenon and Its Resolution*, SIAM Rev. **39**(1997), pp. 644-668.

[11] D. Gottlieb, C.-W. Shu, A. Solomonoff and H. Vandeven, *On the Gibbs Phenomenon I. Recovering Exponential Accuracy From the Fourier Partial Sum of a Nonperiodic Analytic Function*, J. Comput. Appl. Math. **43**(1992), pp. 81-98.

[12] J. S. Hesthaven and R. M. Kirby, *Filtering in Legendre Spectral Methods*, Math. Comp. (2004). – submitted.

[13] J. S. Hesthaven and T. Warburton, *Nodal High-Order Methods on Unstructured Grids I: Time-Domain Solution of Maxwell's Equations*, J. Comput. Phys. **181**(2002), pp. 186-221.

[14] J. S. Hesthaven and T. Warburton, *Discontinuous Galerkin Methods for the Time-Domain Maxwell's Equations: An Introduction*, ACES Newsletter. **vol. 19**(2004), pp. 12-30.

[15] A. Kanevsky, M. H. Carpenter, D. Gottlieb and J. S. Hesthaven, *Implicit-Explicit RKDG Methods and Applications*, (2004). – in preparation.

[16] C. A. Kennedy and M. H. Carpenter, *Low-storage, Explicit Runge-Kutta Schemes for the Compressible Navier-Stokes Equations*, Appl. Numer. Math. **35**(2000), pp. 177-219.

[17] C. A. Kennedy and M. H. Carpenter, *Additive Runge-Kutta Schemes for Convection-Diffusion-Reaction Equations*, Appl. Numer. Math. **44**(2003), pp. 139-181.

[18] H. O. Kreiss and J. Oliger, *Stability of the Fourier Method*, SIAM J. Numer. Anal. **16**(1979), pp. 421-433.

[19] A. Majda, J. McDonough and S. Osher, *The Fourier Method for Nonsmooth Initial Data*, Math. Comp. **32**(1978), pp. 1041-1081.

[20] E. Tadmor and J. Tanner, *Adaptive mollifiers - High Resolution Recovery of Piecewise Smooth Data from its Spectral Information*, Foundat. Comput. Math. **2**(2002), pp.155-189.

[21] H. Vandeven, *Family of Spectral Fitlers for Discontinuous Problems*, J. Sci. Comput. **6**(1991), pp. 159-192.