

ON THE ACCURACY OF THE FINITE ELEMENT METHOD PLUS TIME RELAXATION

J. CONNORS AND W. LAYTON

ABSTRACT. If \bar{u} denotes a local, spatial average of u , then $u' = u - \bar{u}$ is the associated fluctuation. Consider a time relaxation term added to the usual finite element method. The simplest case for the model advection equation $u_t + \vec{a} \cdot \nabla u = f(x, t)$ is

$$(u_{h,t} + \vec{a} \cdot \nabla u_h, v_h) + \chi(u'_h, v'_h) = (f(x, t), v_h).$$

We analyze the error in this and (more importantly) higher order extensions and show that the added time relaxation term not only suppresses excess energy in marginally resolved scales but also increases the accuracy of the resulting finite element approximation.

1. INTRODUCTION

In a landmark result published in 1973, Dupont [Du73] showed that in general the usual, continuous finite element method for first order hyperbolic equations converges suboptimally by one power of the mesh width h , even for infinitely smooth solutions, periodic boundary conditions and uniform meshes (see also Hedstrom [Hed79]). For less smooth solutions, it is also well known that the usual Galerkin finite element methods (FEM) can produce highly oscillatory approximate solutions, e.g., [C79]. Even cases (for example linear elements and cubic splines) for which optimal convergence has been proven for periodic boundary conditions on uniform meshes (e.g., Dupont [Du73], Thomée and Wendroff [TW74]), optimal convergence rates are not expected on highly non-uniform meshes. Dupont's result for smooth solutions and the “wiggles” observed in tests for less smooth solutions have motivated the development of many nonstandard Galerkin methods, stabilizations and regularizations for first order hyperbolic problems and associated convection dominated, convection-diffusion equations. Examples include the SUPG method (Hughes [BH80]), discontinuous Galerkin methods (Lesaint and Raviart [LR74]), subgrid artificial viscosity methods (e.g., Layton [L02], [L05], Guermond [Guer99], Burman and Hansbo [BH04], and Braack, Burman, John and Lube [BBJL07]).

Among these many variations on finite element methods, in complex applications there is a special interest in regularizations that are computationally inexpensive, increase accuracy and incorporate numerical realizations of important physical processes omitted in the hyperbolic model equation. This report performs a numerical

Received by the editor May 30, 2008 and, in revised form, December 13, 2008.

2010 *Mathematics Subject Classification*. Primary 65M15; Secondary 65M60.

Key words and phrases. Time relaxation, deconvolution, hyperbolic equation, finite element method.

The work of both authors was partially supported by NSF grants DMS 0508260 and 0810385.

©2009 American Mathematical Society
Reverts to public domain 28 years from publication

analysis of one such regularization which is motivated by work of Rosenau [R89] and Schochet and Tadmor [ST92] on the regularized Chapman-Enskog expansions of conservation laws. It has been extensively tested by Stolz and Adams and Kleiser in [AS02], [SA99], [SAK01a], [SAK01b], [SAK02] for compressible flows. The regularization is inexpensive and incorporates physical effects by time relaxation (see Section 1.1) to damp fluctuations in time induced by marginally resolved scales in conservation laws and convection dominated problems. This regularization is thus physically interesting; it has been proven to truncate scales [LN07] and is established in the practical computations of Stolz, Adams and Kleiser. In this report we study the complementary accuracy question:

*Does this regularization also increase the asymptotic accuracy
of the approximation as well as stabilize the
approximation of under-resolved solutions?*

To reduce the problem to a simple form, consider the advection equation: for $\Omega = (0, 1)^d$ a (periodic) box in \mathbb{R}^d , given a unit vector \vec{a} and smooth, known 1-periodic functions $u_0(x), f(x, t)$, find 1-periodic $u = u(x, t) : \Omega \times [0, T] \rightarrow \mathbb{R}$ satisfying

$$(1.1) \quad u_t + \vec{a} \cdot \nabla u = f(x, t), \quad x \in \Omega, 0 < t \leq T < \infty, \\ \text{and } u(x, 0) = u_0(x).$$

In 1 dimension the problem is, given smooth, known 1-periodic functions $u_0(x), f(x, t)$, find $u = u(x, t)$ satisfying

$$(1.2) \quad u_t + u_x = f(x, t), \quad x \in (0, 1), 0 < t \leq T < \infty, \\ u(0, t) = u(1, t), \text{ and } u(x, 0) = u_0(x).$$

The simplest example of the discretization.

The simplest example of the family of methods be considered is a small variation on the usual finite element method. To present it, let

$$X = H_{\#}^1(\Omega) := \{v \in L^2(\Omega) : \nabla v \in L^2(\Omega), v(x) = v(x + e_j), j = 1, \dots, d\},$$

and let $X_h \subset X$ denote a generic, conforming finite element space based on a mesh with representative mesh-width h and satisfying an approximation property typical of piecewise polynomials of degree k . The semi-discrete approximation begins with a chosen filter length scale (traditionally denoted δ) and a relaxation parameter χ . Let an overbar denote a discrete local averaging over radius $O(\delta)$ (defined precisely in Section 1.2). Thus, given an approximate solution u_h , its discrete average is denoted $\overline{u_h}^h$ and the associated fluctuation is $u_h' := u_h - \overline{u_h}^h$. Although our analysis is for a specific filter, it can be studied as well for many other filters. The main properties of averaging used in the analysis are $O(\delta^2)$ accuracy in $L^2(\Omega)$ (with $L^2(\Omega)$ norm denoted $\|\cdot\|$) and smoothing in the form

$$\|\phi - \overline{\phi}^h\| \leq C\delta^{2-l}\|\nabla^{2-l}\phi\|, \quad l = 0, 1, 2, \text{ and} \\ \delta^2\|\Delta^h \overline{\phi}^h\| + \delta\|\nabla \overline{\phi}^h\| + \|\overline{\phi}^h\| \leq C\|\phi\|.$$

The zeroth order example of the approximations we consider follows: find $u_h : [0, T] \rightarrow X_h$ satisfying

$$(1.3) \quad (u_{h,t} + \vec{a} \cdot \nabla u_h, v_h) + \chi(u_h', v_h') = 0, \forall v_h \in X_h, \\ u_h(x, 0) \text{ approximates } u_0 \text{ well.}$$

This is the usual Galerkin approximation plus a time relaxation/stabilization term intended to damp small fluctuations; see Section 1.1. In other studies, the time relaxation term has often been added in the simpler form $\cdots + \chi(u_h', v_h)$. The difference between the term $\cdots + \chi(u_h', v_h')$ above and the simpler form $\cdots + \chi(u_h', v_h)$ is discussed in Section 5.2.

Adding in the term $\chi(u_h', v_h')$ introduces a consistency error in the discrete equations which, using its $O(\delta^2)$ accuracy, is

$$\text{consistency error} = \sup_{v_h \in X_h} \frac{\chi(u_h', v_h')}{\|v_h\|} \simeq \sqrt{\chi \|u - \bar{u}\|^2} = C(u) \sqrt{\chi} \delta^2.$$

For an interesting example, if $\chi = O(h^{-1})$, $\delta = O(\sqrt{h})$ this consistency error term is $O(\sqrt{h})$. This suggests that the $N = 0$ case, (1.3) above, is not interesting and higher order (generalized) fluctuations are necessary to attain greater accuracy.

The higher order case.

The most important variant of (1.3), analyzed herein and introduced by Stolz, Adams and Kleiser in their computations of turbulent compressible flows [AS02], [SA99], [SAK01a], [SAK01b], [SAK02] (see also [Gue04]), is based on a higher order fluctuation model. Briefly, given a *continuous* averaging operator, denoted $\phi \rightarrow \bar{\phi}$ (see Section 3.1), a *continuous* deconvolution operator D_N is a bounded linear operator on $L^2(\Omega)$ with the property

$$(1.4) \quad \phi = D_N \bar{\phi} + O(\delta^{2N+2}) \text{ for smooth } \phi.$$

In particular,

$$\|\phi - D_N \bar{\phi}\| \leq C(N) \delta^{2N+2} \|\phi\|_{H^{2N+2}}, \text{ for } \phi \in H_{\#}^{2N+2}(\Omega).$$

In the *discrete* case, considered herein, let $\bar{\phi}^h$ and D_N^h denote discrete averaging and deconvolution operators. These act on X_h instead of X , are defined precisely in Section 2 and have properties analogous to the continuous case. The associated higher order,¹ (discrete) generalized fluctuation is

$$u_h^* := u_h - D_N^h \bar{u}_h^h.$$

The (higher order) time relaxation discretization then follows: find $u_h : [0, T] \rightarrow X_h$ satisfying

$$(1.5) \quad (u_{h,t} + \vec{a} \cdot \nabla u_h, v_h) + \chi(u_h^*, v_h^*) = 0, \forall v_h \in X_h, \\ u_h(x, 0) \in X_h \text{ approximates } u_0 \text{ well.}$$

Note that since $\phi = \bar{\phi} + O(\delta^2)$, (1.3) is the $N = 0$ case of (1.5).

The consistency error in the higher order method (1.5) is not limited (in the continuous case) since

$$\text{consistency error} \simeq \sqrt{\chi \|u - D_N \bar{u}\|^2} = C(u) \sqrt{\chi} \delta^{2N+2}.$$

¹As N increases to moderate values, $\phi \rightarrow D_N \bar{\phi}$ becomes quite close to sharp spectral cutoff.

For example, let $\chi = O(h^{-1})$, $\delta = O(\sqrt{h})$. If $N = 0$ the consistency error is $O(\sqrt{h})$ while if $N = 1$ the consistency error is already $O(h^{3/2})$.

If $\chi = 0$, (1.5) reduces to the usual FEM which converges with suboptimal rate $O(h^k)$ with continuous piecewise polynomials of degree k :

$$\begin{aligned} \sup_{0 \leq t \leq T} \|u(t) - u_h(t)\| &\leq C \|u(0) - u_h(0)\| \\ &+ C \min_{v_h: [0, T] \rightarrow X_h} \sup_{0 \leq t \leq T} \{\|u - v_h\| + \|u_t - v_{h,t}\| + \|(u - v_h)_x\|\}. \end{aligned}$$

The classic paper of Dupont [Du73] shows that in general this result is unimprovable for the usual Galerkin method: the L^2 convergence rate of $O(h^3)$ is attained for Hermite cubics.²

1.1. The genesis and use of time relaxation stabilization. Many stabilizations are used for convection dominated problems and each has its own advantages and disadvantages. The present time relaxation regularization has minimal effect on the solution's large scales. It thus has promise for longer time calculations. It also does not change the order of the equation. Thus, in more complex problems no extra boundary conditions (either explicit or implicit) are needed and no artificial boundary or interior layers are introduced in the solution or its derivatives. When an evolution equation is solved as a part of a complex application in which an efficient filtering routine is implemented, higher order time relaxation is also efficient in both computer time and programmer effort.

The time relaxation term first arose in theoretical studies of regularizations of Chapman-Enskog expansions of conservation laws in Rosenau [R89], Schochet and Tadmor [ST92]. The higher order time relaxation term was pioneered by Stolz, Adams and Kleiser in their large eddy simulations of compressible turbulence. As a stand alone regularization, it has been successful for the Euler equations for shock-entropy wave interaction and other tests ([AL99], [AS01], [AS02], [SAK01a], [SAK01b], [SAK02]), as well as aerodynamic noise prediction and control (Guenaff [Gue04]). It was observed to ensure sufficient numerical entropy dissipation for numerical solution of conservation laws; see Adams and Stolz [AS02], p.393. A mathematical foundation for its inclusion in models for turbulent flow has also been derived; see [LN07], [ELN07].

1.2. An implementation of [MLF03]. Among interesting tests and results for treatment of the nonlinear terms in flow problems, a particularly clever and efficient idea for implementation of time relaxation is given in Mathew, Lechner, Foysi, Sesterhenn and Friedrich [MLF03]. To explain their idea, consider an explicit method in time, suppress the spacial discretization and consider the following algorithm.

Algorithm 1. *Given:* $u^n (\simeq u(t_n))$
Calculate: $\tilde{u}^{n+1} = u^n - \Delta t \vec{a} \cdot \nabla u^n$
Postprocess \tilde{u}^{n+1} *by:*

$$(1.6) \quad u^{n+1} = (1 - \chi) \tilde{u}^{n+1} + \chi DG \tilde{u}^{n+1}.$$

²Dupont shows that there exist infinitely smooth solutions for which a lower estimate for the error of $O(h^3)$ holds. His proof also shows that this suboptimal rate of convergence is the generic case.

The (here explicit) time stepping $u^n \rightarrow \tilde{u}^{n+1}$ can be done by a black box code with the (here explicit) postprocessing added between steps. Eliminating the intermediate quantity \tilde{u}^{n+1} in the algorithm and rearranging terms shows that the algorithm is equivalent to

$$u^{n+1} = u^n - \Delta t \vec{a} \cdot \nabla u^n - \chi[I - DG]\tilde{u}^{n+1},$$

which includes a similar time relaxation term. The form of the time relaxation term can be further altered by various other explicit and implicit postprocessing steps replacing (1.6).

2. PRELIMINARIES: AVERAGING AND DECONVOLUTION

Averaging and deconvolution present interesting new challenges for the numerical analysis of singularly perturbed differential equations. (They are themselves interesting, discrete, elliptic-elliptic singularly perturbation problems, [RST96], [SW83].) There are surely many ways yet to be discovered to use them to increase the accuracy of approximate solutions to many problems. Let

$$C_{\#}^{\infty} := \{\phi \in C_{loc}^{\infty}(\mathbb{R}) : \frac{\partial^{|\alpha|}\phi}{\partial x^{\alpha}}(x) = \frac{\partial^{|\alpha|}\phi}{\partial x^{\alpha}}(x + e_j), \forall j = 1, \dots, d, |\alpha| \geq 0\},$$

and let $H_{\#}^k = H_{\#}^k(\Omega)$ denote the closure of $C_{\#}^{\infty}$ in the H^k norm. We let $X = H_{\#}^1(\Omega)$ and $X_h \subset X$ denote a typical, finite element subspace of X associated with a maximum mesh size h . We shall suppose that the finite element space satisfies the following approximation assumption, typical of piecewise polynomials of degree k : for all $v \in X \cap H^{l+1}(\Omega)$,

$$(2.1) \quad \inf_{v_h \in X_h} \{h|\nabla(v - v_h)| + \|v - v_h\|\} \leq Ch^{l+1}|v|_{H^{l+1}}, \text{ for } 1 \leq l \leq k.$$

2.1. Averaging by continuous and discrete differential filters. We define next the precise continuous and discrete differential filters used herein. These are related to the Yoshida regularization of semi-groups and to scale space analysis. Differential filters were introduced into flow modeling by Germano [Ger86]. The stabilization used is based on averaging by a *discrete differential filter* (Manica and Kaya-Merdan [MM06]). Let $\delta > 0$ (and typically $1 \geq \delta \geq O(h)$) be the selected averaging radius.

Definition 1 (Continuous and discrete differential filter). Given $\phi \in L^2(\Omega)$ its discrete average $G_h\phi = \bar{\phi}^h \in X_h$ is the unique solution of

$$(2.2) \quad \delta^2(\nabla \bar{\phi}^h, \nabla v_h) + (\bar{\phi}^h, v_h) = (\phi, v_h), \forall v_h \in X_h.$$

The associated fluctuation is $\phi' := \phi - \bar{\phi}^h$.

The continuous differentially filtered average $G\phi = \bar{\phi} \in X$ is the unique solution of

$$(2.3) \quad \delta^2(\nabla \bar{\phi}, \nabla v) + (\bar{\phi}, v) = (\phi, v), \forall v \in X.$$

Associated with (2.2) define the usual discrete Laplacian operator $\Delta^h : L^2(\Omega) \rightarrow X_h$ and projection $\Pi_h : L^2(\Omega) \rightarrow X_h$ by

$$\begin{aligned} (\nabla \phi, \nabla v_h) &= (-\Delta^h \phi, v_h), \forall v_h \in X_h, \text{ and} \\ (\phi, v_h) &= (\Pi_h \phi, v_h), \forall v_h \in X_h. \end{aligned}$$

With these definitions, the discrete filter (2.2) can be written $(-\delta^2 \Delta^h + \Pi_h) \bar{\phi} = (\Pi_h \phi)$ or

$$(2.4) \quad \bar{\phi}^h = G_h \phi = (-\delta^2 \Delta^h + \Pi_h)^{-1} (\Pi_h \phi),$$

and the continuous filter is $\bar{\phi} = G \phi = (-\delta^2 \Delta + I)^{-1} \phi$.

The mathematical stability and accuracy properties of continuous and discrete averaging has been extensively studied in [BIL06], [D04], [DE06], [ELN07], [L07], [LL03], [LL05], [LL06a], [LMNR06], [LMNR08], [LN07], [MM06] because they are central to large eddy simulation of turbulent flows. Next, we recall and sharpen a few useful results from [LMNR08] specialized to the periodic case.

Lemma 1 (Stability, smoothing and accuracy of averaging). *For $\phi \in X$ we have*

$$\delta^2 \|\Delta^h \bar{\phi}^h\| + \delta \|\nabla \bar{\phi}^h\| + \|\bar{\phi}^h\| \leq C \|\phi\|, \text{ and } \|\nabla \bar{\phi}^h\| \leq C \|\nabla \phi\|.$$

If $\phi \in X$ and $\Delta \phi \in L^2(\Omega)$

$$\delta^2 \|\nabla(\phi - \bar{\phi}^h)\|^2 + \|\phi - \bar{\phi}^h\|^2 \leq C \inf_{v_h \in X_h} \{\delta^2 \|\nabla(\phi - v_h)\|^2 + \|\phi - v_h\|^2\} + C \delta^4 \|\Delta \phi\|^2.$$

For all $\phi \in X$

$$(2.5) \quad \delta^2 \|\nabla(\bar{\phi} - \bar{\phi}^h)\|^2 + \|\bar{\phi} - \bar{\phi}^h\|^2 = \min_{v_h \in X_h} \{\delta^2 \|\nabla(\bar{\phi} - v_h)\|^2 + \|\bar{\phi} - v_h\|^2\}.$$

Under the approximation assumption (2.1) and for $\bar{\phi} \in X \cap H^{k+1}(\Omega)$

$$(2.6) \quad \delta^2 \|\nabla(\bar{\phi} - \bar{\phi}^h)\|^2 + \|\bar{\phi} - \bar{\phi}^h\|^2 \leq C(\delta^2 h^{2k} + h^{2k+2}) |\bar{\phi}|_{H^{k+1}}^2.$$

Further, for $\bar{\phi} \in X \cap H^{k+1}(\Omega)$

$$(2.7) \quad \begin{aligned} \|\bar{\phi} - \bar{\phi}^h\| &\leq C \frac{h}{\delta} \min_{v_h \in X_h} \{\delta^2 \|\nabla(\bar{\phi} - v_h)\|^2 + \|\bar{\phi} - v_h\|^2\}^{\frac{1}{2}} \\ &\leq C(h^{k+1} + \delta^{-1} h^{k+2}) |\bar{\phi}|_{H^{k+1}}, \end{aligned}$$

$$(2.8) \quad \begin{aligned} \|\bar{\phi} - \bar{\phi}^h\|_{H^{-1}(0,1)} &\leq C h \min_{v_h \in X_h} \{\delta^2 \|\nabla(\bar{\phi} - v_h)\|^2 + \|\bar{\phi} - v_h\|^2\}^{\frac{1}{2}} \\ &\leq C(\delta h^{k+1} + h^{k+2}) |\bar{\phi}|_{H^{k+1}}, \text{ for } k \geq 1, \end{aligned}$$

$$(2.9) \quad \begin{aligned} \|\bar{\phi} - \bar{\phi}^h\|_{H^{-1}(0,1)} &\leq C h \left(\frac{h}{\delta}\right) \min_{v_h \in X_h} \{\delta^2 \|\nabla(\bar{\phi} - v_h)\|^2 + \|\bar{\phi} - v_h\|^2\}^{\frac{1}{2}} \\ &\leq C(h^{k+2}(1 + \frac{h}{\delta})) |\bar{\phi}|_{H^{k+1}}, \text{ for } k \geq 2. \end{aligned}$$

Proof. The first two claims were proven in Lemma 2.11 and 2.12 in [LMNR08] in 2 and 3 dimensions but the proof is dimension independent. The third and fourth claim will follow from normal finite element error analysis (e.g., [SW83]), accounting for the dependence upon δ . We give a brief proof next. Given ϕ , the equations for $\bar{\phi}$ and $\bar{\phi}^h$ are, respectively,

$$(2.10) \quad \begin{aligned} \delta^2 (\nabla \bar{\phi}^h, \nabla v_h) + (\bar{\phi}^h, v_h) &= (\phi, v_h), \forall v_h \in X_h, \\ \delta^2 (\nabla \bar{\phi}, \nabla v) + (\bar{\phi}, v) &= (\phi, v), \forall v \in X. \end{aligned}$$

In other words, $\bar{\phi}^h$ is the usual Galerkin approximation of a symmetric and coercive problem so the third claim (2.5) holds (see, e.g., [SW83] for more details about estimates for elliptic-elliptic singular perturbation problems). The fourth (2.6) follows from the third plus the approximation assumption. The fifth follows

from a classical duality argument for (2.10) as follows. Let Ψ be the unique 1-periodic solution of

$$(2.11) \quad -\delta^2 \Delta \Psi + \Psi = \bar{\phi} - \bar{\phi}^h, \text{ on } \Omega \text{ and } \Psi(x) = \Psi(x + e_j), j = 1, \dots, d.$$

The solution Ψ has the following regularity, for example [L07],

$$\delta^2 \|\Delta \Psi\| + \delta \|\nabla \Psi\| + \|\Psi\| \leq C \|\bar{\phi} - \bar{\phi}^h\|.$$

Subtracting the continuous and discrete equations in (2.10) above gives a standard Galerkin orthogonality condition

$$\delta^2 (\nabla(\bar{\phi} - \bar{\phi}^h), \nabla v_h) + (\bar{\phi} - \bar{\phi}^h, v_h) = 0, \forall v_h \in X_h.$$

The variational formulation of the dual problem (2.11) is

$$\delta^2 (\nabla \Psi, \nabla v) + (\Psi, v) = (\bar{\phi} - \bar{\phi}^h, v), \forall v \in X.$$

Setting $v = \bar{\phi} - \bar{\phi}^h$ and using the Galerkin orthogonality gives, $\forall v_h \in X_h$,

$$\begin{aligned} \|\bar{\phi} - \bar{\phi}^h\|^2 &= \delta^2 (\nabla(\Psi - v_h), \nabla(\bar{\phi} - \bar{\phi}^h)) + (\Psi - v_h, \bar{\phi} - \bar{\phi}^h) \\ &\leq [\delta^2 \|\nabla(\Psi - v_h)\|^2 + \|\Psi - v_h\|^2]^{\frac{1}{2}} [\delta^2 \|\nabla(\bar{\phi} - \bar{\phi}^h)\|^2 + \|\bar{\phi} - \bar{\phi}^h\|^2]^{\frac{1}{2}}. \end{aligned}$$

Choosing v_h appropriately and using the approximation assumption (2.1) and the regularity of Ψ yields

$$\|\bar{\phi} - \bar{\phi}^h\|^2 \leq C[(\frac{h}{\delta})^2 + (\frac{h}{\delta})^4]^{\frac{1}{2}} \|\bar{\phi} - \bar{\phi}^h\| [\delta^2 \|\nabla(\bar{\phi} - \bar{\phi}^h)\|^2 + \|\bar{\phi} - \bar{\phi}^h\|^2]^{\frac{1}{2}},$$

so that, using the approximation assumption gives

$$\begin{aligned} \|\bar{\phi} - \bar{\phi}^h\| &\leq C \frac{h}{\delta} [\delta^2 \|\nabla(\bar{\phi} - \bar{\phi}^h)\|^2 + \|\bar{\phi} - \bar{\phi}^h\|^2]^{\frac{1}{2}} \\ &\leq Ch^{k+1} [1 + \frac{h}{\delta}] \|\bar{\phi}\|_{H^{k+1}}, \end{aligned}$$

which is the claimed L^2 error bound.

The H^{-1} estimate will follow similarly by duality. Indeed, let $\beta \in H_{\#}^1(\Omega)$ be fixed but arbitrary and let Ψ be the 1-periodic solution of

$$(2.12) \quad -\delta^2 \Delta \Psi + \Psi = \beta, \text{ on } \Omega \text{ and } \Psi(x) = \Psi(x + e_j), j = 1, \dots, d.$$

It is known, e.g., [L07] (and easily proven by Fourier series), that the solution Ψ satisfies

$$\delta^2 \|\nabla^3 \Psi\| + \delta \|\nabla^2 \Psi\| + \|\nabla \Psi\| \leq C \|\nabla \beta\|.$$

The variational formulation of the dual problem (2.12) is that

$$\delta^2 (\Psi_x, v_x) + (\Psi, v) = (\beta, v), \forall v \in X.$$

Setting $v = \bar{\phi} - \bar{\phi}^h$ and using the Galerkin orthogonality (2.10) gives, $\forall v_h \in X_h$,

$$\begin{aligned} (\beta, \bar{\phi} - \bar{\phi}^h) &= \delta^2 (\nabla(\Psi - v_h), \nabla(\bar{\phi} - \bar{\phi}^h)) + (\Psi - v_h, \bar{\phi} - \bar{\phi}^h) \\ &\leq [\delta^2 \|\nabla(\Psi - v_h)\|^2 + \|\Psi - v_h\|^2]^{\frac{1}{2}} [\delta^2 \|\nabla(\bar{\phi} - \bar{\phi}^h)\|^2 + \|\bar{\phi} - \bar{\phi}^h\|^2]^{\frac{1}{2}}. \end{aligned}$$

Using the approximation assumption and the regularity of Ψ yields

$$\begin{aligned} (\beta, \bar{\phi} - \bar{\phi}^h) &\leq Ch \|\nabla \beta\| [\delta^2 \|\nabla(\bar{\phi} - \bar{\phi}^h)\|^2 + \|\bar{\phi} - \bar{\phi}^h\|^2]^{\frac{1}{2}} \text{ if } k \geq 1 \text{ and} \\ (\beta, \bar{\phi} - \bar{\phi}^h) &\leq Ch (\frac{h}{\delta}) \|\nabla \beta\| [\delta^2 \|\nabla(\bar{\phi} - \bar{\phi}^h)\|^2 + \|\bar{\phi} - \bar{\phi}^h\|^2]^{\frac{1}{2}} \text{ if } k \geq 2. \end{aligned}$$

The last two results follow by dividing by $||\nabla\beta||$ and taking the supremum of $\beta \in H_{\#}^1(\Omega)$. \square

2.2. Deconvolution. The deconvolution problem, central in image processing, e.g., [BB98], is

given $\bar{\phi}$ (+noise), find ϕ (approximately).

One of the most basic deconvolution methods is the van Cittert algorithm, [vC31], [BB98]. For continuous deconvolution, it is equivalent to N steps of first order Richardson iteration for the equivalent problem

given \bar{u} solve $u = u + \{\bar{u} - Gu\}$ for u .

Algorithm 2 (van Cittert approximate deconvolution). *Set $v_0 = \bar{u}$ and fix N for $n = 1, 2, \dots, N-1$, perform $v_{n+1} = v_n + \{\bar{u} - Gv_n\}$. Define $D_N \bar{u} := v_N$.*

The discrete van Cittert deconvolution operator is defined by substituting G_h for G and \bar{u}^h for \bar{u} in the above algorithm. The discrete van Cittert deconvolution operator will be denoted by D_N^h . By eliminating the intermediate steps, the N^{th} van Cittert deconvolution operators D_N and D_N^h are given explicitly by

$$(2.13) \quad D_N \phi := \sum_{n=0}^N (I - G)^n \phi \text{ and } D_N^h \phi := \sum_{n=0}^N (\Pi_h - G_h)^n \phi.$$

The van Cittert operator acts like an extrapolation in scale space from resolved to unresolved scales. For example, the approximate deconvolution operators corresponding to $N = 0, 1, 2$ are

$$\begin{aligned} D_0 \bar{u} &= \bar{u}, \\ D_1 \bar{u} &= 2\bar{u} - \bar{\bar{u}}, \\ D_2 \bar{u} &= 3\bar{u} - 3\bar{\bar{u}} + \bar{\bar{\bar{u}}}. \end{aligned}$$

Definition 2 (Deconvolution error). Given ϕ , the deconvolution error is, respectively, in the continuous or discrete cases given by

$$\text{deconvolution error} = \phi - D_N \bar{\phi} \text{ or } = \phi - D_N^h \bar{\phi}^h.$$

The deconvolution error plays a fundamental role in the consistency error inherent in algorithms based on deconvolution methods. Although there remain many open questions about even these simplest deconvolution operators, a theory is beginning to develop. Next we summarize a few points, sharpening the results, as necessary, for the hyperbolic problem.

Proposition 1 (Stability and accuracy of deconvolution). *Let D_N and D_N^h be given by the van Cittert algorithm. Then D_N and $D_N^h : L^2(\Omega) \rightarrow L^2(\Omega)$ are bounded, self-adjoint positive operators as are $I - D_N$ and $I - D_N^h$. Further,*

$$\begin{aligned} ||D_N \bar{v}|| &\leq C(N) ||v||, \forall v \in X, \\ ||D_N^h \bar{v}^h|| &\leq C(N) ||v||, \forall v \in X, \\ ||\nabla D_N^h \bar{v}^h|| &\leq C(N) ||\nabla v||, \forall v \in X. \end{aligned}$$

In the case of differential filters $A := -\delta^2 \Delta + 1$, $A_h := (-\delta^2 \Delta^h + \Pi_h) \Pi_h$, $G = A^{-1}$, and $G_h = (A_h)^{-1} \Pi_h$. Then

$$(2.14) \quad \phi - D_N \bar{\phi} = \delta^{2N+2} (-\Delta)^{N+1} (A)^{-(N+1)} \phi, \forall \phi \in L^2(\Omega),$$

$$(2.15) \quad \phi_h - D_N^h \bar{\phi}^h = \delta^{2N+2} (-\Delta^h)^{N+1} (A_h)^{-(N+1)} \phi, \forall \phi_h \in X_h.$$

Proof. The stability claims are in Lemma 2.11 of [LMNR08]. The first accuracy results (2.14) was proven by Stoltz and Adams [SA99] and independently by Dunca [D04] and Dunca and Epshteyn [DE06], see also [BIL06]. The second (2.15) is in Lemma 2.10 in [LMNR08]. \square

Lemma 2 (Smoothing property). *For any $\phi_h \in X_h$,*

$$\delta^2 \|\Delta^h (D_N^h \bar{\phi}^h)\| + \delta \|\nabla (D_N^h \bar{\phi}^h)\| + \|D_N^h \bar{\phi}^h\| \leq C(N) \|\phi_h\|.$$

Proof. Consider the van Cittert algorithm. We have at its initiation $\phi_0 = \bar{\phi}_h^h$ which satisfies

$$(2.16) \quad \delta^2 (\nabla \bar{\phi}_h^h, \nabla v_h) + (\bar{\phi}_h^h, v_h) = (\phi_h, v_h), \forall v_h \in X_h.$$

Setting $\phi_h = v_h$ and using various inequalities gives for $\phi_0 = \bar{\phi}_h^h$

$$(2.17) \quad \|\phi_0\| \stackrel{\text{def}}{=} \{\delta^4 \|\Delta^h (D_N^h \phi_0)\|^2 + \delta^2 \|\nabla (D_N^h \phi_0)\|^2 + \|D_N^h \phi_0\|^2\}^{\frac{1}{2}} \leq C \|\phi_h\|.$$

For the first step of van Cittert,

$$\phi_1 = \phi_0 + \{\bar{\phi}_0^h - \bar{\phi}_h^h\}.$$

Thus, by the triangle inequality and the last estimate,

$$(2.18) \quad \begin{aligned} \|\phi_1\| &\leq \|\phi_0\| + \|\bar{\phi}_0^h\| + \|\bar{\phi}_h^h\| \\ &\leq C \|\phi_h\| + \|\bar{\phi}_0^h\| \end{aligned}$$

If the above estimate (2.17) is applied with ϕ_h replaced by ϕ_0 on the right-hand side, we have

$$\|\bar{\phi}_0^h\| \leq C \|\phi_0\| = C \|\bar{\phi}_h^h\| \leq C \|\phi_h\|,$$

again by (2.17). Thus,

$$\|\phi_1\| \leq C \|\phi_h\|.$$

For the general case we proceed by induction. \square

One question regards boundedness of the right-hand side of (2.14) and (2.15). Some partial answers, summarized below, are in [L07] and [LMNR08].

Proposition 2 (Deconvolution error estimates). *For all $\phi \in L^2(\Omega)$,*

$$\|\phi - D_N^h \bar{\phi}^h\| \leq C(N) \|\phi - \bar{\phi}^h\|,$$

while if $\phi \in X$ and $\Delta \phi \in L^2(\Omega)$

$$\|\phi - D_N^h \bar{\phi}^h\| \leq C(N) \inf_{v_h \in X_h} \{\delta^2 \|\nabla(\phi - \bar{\phi}^h)\|^2 + \|\phi - \bar{\phi}^h\|^2\}^{\frac{1}{2}} + C(N) \delta^2 \|\Delta \phi\|.$$

Further, under the approximation assumption (2.1) and for $\phi \in X \cap H^{2N+2}(\Omega) \cap H^{k+1}(\Omega)$,

$$\|\phi - D_N^h \bar{\phi}^h\| \leq C(N) (\delta h^k + h^{k+1}) \|\phi\|_{k+1} + C \delta^{2N+2} \|\phi\|_{2N+2}.$$

Under the same conditions,

$$\|\phi - D_N^h \bar{\phi}^h\| \leq C(N)(h^{k+1} + \delta^{-1}h^{k+2})\|\phi\|_{k+1} + C\delta^{2N+2}|\phi|_{2N+2}.$$

Proof. The first two inequalities are proven in Lemmas 2.12 and 2.13 in [LMNR08]. The third combines Lemma 2.14 and Remark 2.14 in [LMNR08] with [L07]. The remainder is a sharpening of Lemma 2.14 in [LMNR08]. Since the proof of the shortened result has same structure as that of Lemma 2.14 in [LMNR08], we shall outline the proof where it is identical to that of Lemma 2.14 in [LMNR08] and give the details where it deviates.

To begin, we rewrite $\phi - D_N^h \bar{\phi}^h = (I - D_N^h G_h)\phi$ as

$$(2.19) \quad (I - D_N^h G^h)\phi = (I - D_N G)\phi + (D_N - D_N^h)G\phi + D_N^h(G - G^h)\phi.$$

This is a small but critical reordering of the decomposition of the corresponding one in the proof Lemma 2.14 in [LMNR08]. We know $\|D_N^h\| \leq C(N)$, so

$$\|D_N^h(G - G^h)\phi\| \leq C(N)\|\bar{\phi} - \bar{\phi}^h\| \leq Ch^{k+1}(1 + \frac{h}{\delta})|\bar{\phi}|_{k+1}.$$

The bound of the right-hand side has been sharpened in Lemma 1 above from the corresponding one in [LMNR08]. The first term in (2.19) is bounded in Proposition 2 giving

$$\|(I - D_N G)\phi\| \leq C\delta^{2N+2}\|\Delta^{N+1}\phi\| \leq \delta^{2N+2}|\phi|_{2N+2}.$$

Consider the term $(D_N - D_N^h)G\phi$. Adding and subtracting terms in this order yields $G\phi$ (rather than $G_h\phi$) which is at least as smooth as ϕ . Thus, we estimate $(D_N - D_N^h)G\phi$ using the argument in [LMNR08] (Lemma 2.14). Indeed, with $\psi = G\phi$,

$$(D_N - D_N^h)G\phi = \sum_{n=0}^N [(I - G)^n - (I - G_h)^n]\psi + \sum_{n=0}^N [(\Pi_h - G)^n - (I - G_h)^n]\psi.$$

The $n = 0$ term vanishes while the $n = 1$ term is $\bar{\psi} - \bar{\psi}^h$, bounded in Lemma 1 above. The first term on the right-hand side is the critical one. Thus, consider

$$\sum_{n=0}^N [(I - G)^n - (I - G_h)^n]\psi.$$

For $n = 2$ we have, adding and subtracting $(I - G_h)(I - G)$ (instead of in the other order $(I - G)(I - G_h)$ used in [LMNR08]), gives

$$(I - G)^2\psi - (I - G_h)^2\psi = (G_h - G)(I - G)\psi + (I - G_h)(G_h - G)\psi.$$

It is not difficult to show, using the spectral mapping theorem, that both G_h and $I - G_h$ are SPD (symmetric positive definite) and $\|I - G_h\|_{L(L^2 \rightarrow L^2)} \leq 1$. Thus,

$$\begin{aligned} \|(I - G)^2\psi - (I - G_h)^2\psi\| &\leq \|(G_h - G)(I - G)\psi\| + \|(I - G_h)(G_h - G)\psi\| \\ &\leq C(h^{k+1} + \delta^{-1}h^{k+2})\{|(I - G)\bar{\phi}|_{k+1} + |\bar{\phi}|_{k+1}\}. \end{aligned}$$

Since $|\bar{\phi}|_{k+1} \leq C\|\phi\|_{k+1}$, $|\bar{\phi}|_{k+1} \leq C\|\phi\|_{k+1}$, we have

$$\|(I - G)^2\psi - (I - G_h)^2\psi\| \leq C(h^{k+1} + \delta^{-1}h^{k+2})\|\phi\|_{k+1},$$

and the result holds for $n = 2$. To complete the proof we continue by induction for $3 \leq n \leq N$ as in [LMNR08] only adding and subtracting terms in the same order as the above $n = 2$ case. \square

The error in discrete deconvolution is bounded by the error in the best approximation in X_h . If filtering is itself inexpensive to perform, then the van Cittert algorithm is economical in both computer time and programmer effort because it only requires repeated filtering. Various other deconvolution methods are also being developed for similar purposes such as Tikhonov [S07], [MS07] and optimized van Cittert [LS07], and could also be considered. It is also possible to define means and fluctuations by projections into hierarchical finite element spaces. This idea leads to more methods that are similar in motivation but whose analysis would be different in detail.

3. A QUASI-STATIC PROJECTION

The reason for the improvement in accuracy for the time relaxation discretization is captured already in the analysis of an equilibrium projection in this section. The projection

$$Q : X \rightarrow X_h$$

is defined after some necessary notation as follows. When $N = 0$, $D_0^h = \Pi_h$ and $(I - D_0^h G_h)u = u - \bar{u}^h$ is the fluctuation about the mean (normally denoted u'). Analogously, for $N > 0$, $u - D_N^h G_h u = u - D_N^h \bar{u}^h$ represents a higher order, generalized fluctuation (for example [LMNR06], [LN07]) that we will denote by u^* .

Definition 3 (Higher order fluctuations). The generalized (higher order) fluctuation, denoted u^* , is

$$u^* := u - D_N^h G_h u.$$

Given $\chi > 0$ and $\delta > 0$ and for $w \in X$, the projection $Q : X \rightarrow X_h$ by $w_h := Qw \in X_h$ is (unique) solution of the finite dimensional linear problem,

$$(3.1) \quad (\vec{a} \cdot \nabla(w - w_h) + (w - w_h), v_h) + \chi((w - w_h)^*, v_h^*) = 0, \forall v_h \in X_h.$$

Lemma 3. Let $\chi > 0$, $\delta \geq 0$. (3.1) has a unique solution. $Q : X \rightarrow X_h$ is a well-defined projection operator.

Proof. The equations (3.1) defining Q reduce to a linear system for the projection so existence is implied by triviality of $\ker(Q)$. To verify triviality, let $w = 0$, and set $v_h = w_h$. This gives

$$(\vec{a} \cdot \nabla w_h, w_h) + \|w_h\|^2 + \chi\|w_h^*\|^2 = 0.$$

Under periodic boundary conditions $(\vec{a} \cdot \nabla w_h, w_h) = 0$ and thus $w_h = 0$. That $Q^2 = Q$ follows similarly. \square

If $\chi = 0$, (3.1) reduces to the usual finite element method for a two point boundary value problem. Finite element error analysis shows that $w - Qw$ satisfies

$$\|w - Qw\| \leq C \inf_{v_h \in X_h} \{ \|\nabla(w - v_h)\| + \|w - v_h\| \}.$$

This estimate also leads to an asymptotic rate of convergence suboptimal by one power of h , sharp for some elements, [L83] (as for hyperbolic problems, Dupont [Du73]). For other, special elements on uniform meshes and for smoother solutions, this extra power of h can be recovered by a cancellation argument, e.g., Axelsson and Gustafsson [AG79]. On the other hand, if, as is commonly manifested as oscillations at the smallest resolved scale, this loss of accuracy comes in the behavior of the error at the smallest resolved scales, it is plausible that the extra control of small

scales in the time relaxation discretization for $\chi > 0$ will lead to an increase in accuracy as well as stability.

Theorem 1 (Projection error). *Let D_N^h be the discrete van Cittert deconvolution operator. Let $w \in X$ and let Q be the projection (3.1). Then, the error in the projection satisfies*

$$\begin{aligned} \|w - Qw\|^2 + \chi \|(w - Qw)^*\|^2 &\leq C(N) \inf_{v_h \in X_h} \{(1 + \delta^{-2})\|w - v_h\|^2 \\ &\quad + \chi^{-1}\|\nabla(w - v_h)\|^2 + \chi\|(w - v_h)^*\|^2\}. \end{aligned}$$

Proof. Let $\tilde{w} \in X_h$ be arbitrary (for the moment), and write

$$w - Qw = (w - \tilde{w}) - (Qw - \tilde{w}) = \eta - \phi_h,$$

where $\eta := w - \tilde{w}$ and $\phi_h := Qw - \tilde{w} \in X_h$. The projection equations can be rewritten as

$$(\vec{a} \cdot \nabla \phi_h + \phi_h, v_h) + \chi(\phi_h^*, v_h^*) = (\vec{a} \cdot \nabla \eta + \eta, v_h) + \chi(\eta^*, v_h^*) = 0, \forall v_h \in X_h.$$

Setting $v_h = \phi_h$ and using $(\vec{a} \cdot \nabla \phi_h, \phi_h) = 0$ and the usual inequalities gives

$$(3.2) \quad \|\phi_h\|^2 + \chi\|\phi_h^*\|^2 \leq 2(\vec{a} \cdot \nabla \eta, \phi_h) + \|\eta\|^2 + \chi\|\eta^*\|^2.$$

The key term is $(\vec{a} \cdot \nabla \eta, \phi_h)$. The idea is to separate scales in this term and treat the large and small scales differently. The small scales are controlled by the stabilization in the time relaxation term and the large scales are treated as a consistency error term. (This idea has been used for other stabilizations in, for example, [L02], [L05], and Guermond [Gue04].) We treat the $N = 0$ case and the general case separately to make the ideas clear.

The deconvolution order $N = 0$ case.

In the $N = 0$ case this term is split into means and fluctuations and is integrated by parts:

$$N = 0 : (\vec{a} \cdot \nabla \eta, \phi_h) = (\vec{a} \cdot \nabla \eta, \overline{\phi_h^h} + \phi_h') = -(\eta, \vec{a} \cdot \nabla \overline{\phi_h^h}) + (\vec{a} \cdot \nabla \eta, \phi_h').$$

Inserting this in the right-hand side of (3.2), recalling that $|\vec{a}| = 1$, and using standard inequalities gives

$$\|\phi_h\|^2 + \chi\|\phi_h'\|^2 \leq \|\eta\|^2 + \chi\|\eta'\|^2 + \frac{\chi}{2}\|\phi_h'\|^2 + \frac{C}{2\chi}\|\nabla \eta\|^2 + 2\|\eta\|\|\nabla \overline{\phi_h^h}\|.$$

We use the a priori bound $\|\nabla \overline{\phi_h^h}\| \leq \frac{1}{2\delta}\|\phi_h\|$ from Lemma 1 (tracking the constant through its proof) and $2\|\eta\|\|\nabla \overline{\phi_h^h}\| \leq \|\eta\|\frac{1}{\delta}\|\phi_h\| \leq \frac{1}{2}\|\phi_h\|^2 + \frac{\delta^{-2}}{2}\|\eta\|^2$. This yields

$$\|\phi_h\|^2 + \chi\|\phi_h'\|^2 \leq (2 + \delta^{-2})\|\eta\|^2 + 2\chi\|\eta'\|^2 + C\chi^{-1}\|\nabla \eta\|^2 + 2\|\eta\|.$$

Thus, by the triangle inequality,

$$\begin{aligned} \|w - Qw\|^2 + \chi\|(w - Qw)'\|^2 \\ \leq C(N) \inf_{v_h \in X_h} \{(1 + \delta^{-2})\|w - v_h\|^2 + \chi^{-1}\|\nabla(w - v_h)\|^2 + \chi\|(w - v_h)'\|^2\}. \end{aligned}$$

The case of higher deconvolution order $N \geq 1$.

The proof for $N \geq 1$ follows the above $N = 0$ case only the splitting uses generalized fluctuations. Indeed, split

$$(\vec{a} \cdot \nabla \eta, \phi_h) = (\vec{a} \cdot \nabla \eta, D_N^h(\overline{\phi_h^h}) + \phi_h^*) = -(\eta, \vec{a} \cdot \nabla D_N^h(\overline{\phi_h^h})) + (\vec{a} \cdot \nabla \eta, \phi_h^*).$$

Both terms on the right-hand side are handled analogously to the $N = 0$ case with the second term treated identically. The first term on the right-hand side $(\eta, \vec{a} \cdot \nabla(D_N^h(\overline{\phi_h^h})))$ is treated like the term $(\eta, \vec{a} \cdot \nabla \overline{\phi_h^h})$ of the $N = 0$ case using the a priori estimate from Lemma 2 that $\|\nabla D_N^h(\overline{\phi_h^h})\| \leq C\delta^{-1}\|\phi_h\|$. The remainder of the proof follows exactly the $N = 0$ case. \square

Corollary 1. *Let X_h satisfy the approximation assumption (2.1). Let $w \in X$ be smooth and let Q be the projection (3.1). Then, the error in the projection satisfies*

$$\|w - Qw\|^2 + \chi\|(w - Qw)^*\|^2 \leq C(N, w)\{(1 + \delta^{-2})h^{2k+2} + \chi^{-1}h^{2k} + \chi h^{2k+2}\}.$$

Thus if $\delta = O(h^{\frac{1}{2}})$, $\chi = O(h^{-1})$

$$\begin{aligned} \|w - Qw\| &\leq C(N, w)h^{k+\frac{1}{2}}, \\ \|(w - Qw)^*\| &\leq C(N, w)h^{k+1}. \end{aligned}$$

Proof. Use $\|(w - v_h)^*\| \leq C\|w - v_h\|$ and the approximation assumption (2.1). \square

4. NUMERICAL ANALYSIS OF THE TIME DEPENDENT PROBLEM

This section proves the following error estimate for the method (1.5) which (roughly speaking) states that the error in the method consists of the error in the equilibrium projection plus a consistency error term. With proper choice of χ , δ and N , we obtain the improved rate of convergence of $O(h^{k+\frac{1}{2}})$ which seems to be typical of stabilized methods.

Theorem 2 (Convergence of the method with time relaxation). *Let Q be the equilibrium projection (3.1). Let X_h satisfy (2.1). For $0 \leq t \leq T < \infty$ the error in the method (1.5) satisfies*

$$\begin{aligned} \sup_{[0, T]} \|u - u_h\|^2 + \int_0^T \chi\|(u - u_h)^*\|^2 dt &\leq C(T, N)\{\|u(0) - u_h(0)\|^2 + \sup_{[0, T]} \|u - Qu\|^2 \\ &\quad + \int_0^T [\chi\|(u - Qu)^*\|^2 + \|u_t - Qu_t\|^2 + \chi\|u^*(t)\|^2] dt\}. \end{aligned}$$

Remark 1. The consistency error $\int_0^T \chi\|u^*(t)\|_*^2 dt$ is directly related to the deconvolution error and will be bounded using Proposition 2. Since $(Qu)_t = Q(u_t)$, the remaining terms will be estimated using Theorem 1.

Proof. Subtraction shows that the error $e(t) := u(t) - u_h(t)$ satisfies

$$(e_t + \vec{a} \cdot \nabla e, v_h) + \chi(e^*, v_h^*) = \chi(u^*, v_h^*), \forall v_h \in X_h,$$

which is driven by the methods consistency error on the right-hand side. As usual, split the error as

$$e = \eta - \phi_h, \eta = u - Qu, \phi_h = u_h - Qu.$$

Rearranging the error equation, we then have, for any $v_h \in X_h$,

$$(4.1) \quad (\phi_{h,t} + \vec{a} \cdot \nabla \phi_h, v_h) + \chi(\phi_h^*, v_h^*) = (\eta_t - \eta, v_h) + \chi(u^*, v_h^*) + [(\vec{a} \cdot \nabla \eta + \eta, v_h) + \chi(\eta^*, v_h^*)].$$

By the definition of the projection operator Q , the bracketed term on the right-hand side vanishes. Setting $v_h = \phi_h$ gives

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\phi_h\|^2 + \chi \|\phi_h^*\|^2 &= (\eta_t - \eta, \phi_h) + \chi(u^*, \phi_h^*) \\ &\leq \frac{1}{2} [\|\eta_t\|^2 + \|\eta\|^2] + \frac{\chi}{2} \|u^*\|^2 + \frac{\chi}{2} \|\phi_h^*\|^2 + \|\phi_h\|^2. \end{aligned}$$

Equivalently,

$$\frac{d}{dt} \|\phi_h\|^2 + \chi \|\phi_h^*\|_*^2 \leq [\|\eta_t\|^2 + \|\eta\|^2] + \chi \|u^*\|_*^2 + 2 \|\phi_h\|^2.$$

Gronwall's inequality implies that for $0 \leq t \leq T < \infty$,

$$\|\phi_h(t)\|^2 + \int_0^t \chi \|\phi_h^*\|^2 dt' \leq \|\phi_h(0)\|^2 + C(T) \int_0^t [\|\eta_t(t')\|^2 + \|\eta(t')\|^2 + \chi \|u^*(t')\|^2] dt'.$$

The trial inequality then yields the claimed result. \square

Using the error estimate for the projection in Theorem 1 gives the following.

Corollary 2. *For $0 \leq t \leq T < \infty$, the error in the method (1.5) satisfies*

$$\begin{aligned} \sup_{[0,T]} \|u - u_h\|^2 + \int_0^T \chi \|(u - u_h)^*\|^2 dt &\leq C(T, N) \{ \|u(0) - u_h(0)\|^2 \\ &+ \inf_{v_h \in C^1(0,T;X_h)} \{ \sup_{[0,T]} [(1 + \delta^{-2}) \sup_{[0,T]} \|u - v_h\|^2 + \chi^{-1} \|\nabla(u - v_h)\|^2 + \chi \|(u - v_h)^*\|^2] \\ &+ \int_0^T (1 + \delta^{-2}) \|(u - v_h)_t\|^2 + \chi^{-1} \|\nabla(u - v_h)_t\|^2 + \chi \|(u - v_h)_t^*\|^2 dt \} \\ &+ \int_0^T \chi \|u^*(t)\|^2 dt \}. \end{aligned}$$

Proof. This follows from Theorems 1 and 2. \square

By Proposition 2 we can estimate the consistency error which is the square root of the last term in the above error estimate. Indeed, under the approximation assumption (2.1),

$$\begin{aligned} \int_0^T \chi \|u^*(t)\|^2 dt &= \int_0^T \chi \|(I - D_N^h G_h)u(t)\|^2 dt \\ (4.2) \quad &\leq \int_0^T C \chi h^{2k+2} (1 + (\frac{h}{\delta})^2) \|u(t)\|_{k+1}^2 + C \chi \delta^{4N+4} |u(t)|_{2N+2}^2 dt \\ &\leq C(u) \chi \{ h^{2k+2} (1 + (\frac{h}{\delta})^2) + \delta^{4N+4} \}. \end{aligned}$$

Rates of convergence then follow.

Corollary 3. *Suppose the assumptions of Theorem 2 hold, u is smooth, and the approximation assumption (2.1) holds. Then,*

$$\begin{aligned} \sup_{[0,T]} \|u - u_h\|^2 + \int_0^T \chi \|(u - u_h)^*\|^2 dt &\leq C(T, N, u) \{ \|u(0) - u_h(0)\|^2 \\ &+ (1 + \delta^{-2}) h^{2k+2} + \chi^{-1} h^{2k} + \chi h^{2k+2} + \chi h^{2k+2} (\frac{h}{\delta})^2 + \chi \delta^{4N+4} \}. \end{aligned}$$

Proof. This is immediate. \square

Taking square roots and collecting the leading terms (when $1 \geq \delta \geq h, \chi \geq 1$) in the above error estimate gives

$$\text{error}_{[0,T]} \lesssim \text{error}(0) + \{\delta^{-1}h^{k+1} + \chi^{-\frac{1}{2}}h^k + \chi^{\frac{1}{2}}h^{k+1} + \chi^{\frac{1}{2}}h^{k+1}(\frac{h}{\delta}) + \chi^{\frac{1}{2}}\delta^{2N+2}\}.$$

The error is optimized by

$$\delta \simeq \sqrt{h}, \chi \simeq h^{-1}, \text{ and } N \geq k.$$

These choices attain the accuracy for smooth solutions

$$\sup_{[0,T]} \|u - u_h\| \lesssim \text{error}(0) + Ch^{k+\frac{1}{2}} \text{ and } \sqrt{\int_0^T \|(u - u_h)^*\|^2 dt} \lesssim \text{error}(0) + Ch^{k+1},$$

which is suboptimal by one-half power of h for the L^2 error but optimal for the error in the generalized fluctuation.

5. POSSIBLE EXTENSION TO ANOTHER TIME RELAXATION TERM

The theory of the discretization has been developed for a simplified problem: one space dimension, periodic boundary conditions, no time discretization, and a convenient form of the time relaxation term. We shall consider nontrivial boundary conditions through a computational illustration in Section 6. Time discretization of the relaxation term is less understood. In the case of implicit methods, if (1.5) is discretized in time by the (1,1)-Padé/trapezoid method, it is straightforward (although longer than the continuous time case) to prove stability, convergence and even superconvergence. Thus, there remains efficiency, which is critical in 3 dimensional problems for which (1.1) is a common simplified model. In full trapezoidal discretizations of the method (1.5), at each time step a linear system must be solved which, if assembled, is fully due to the (nonlocal) filtering terms in the deconvolution operator D_N . When an iterative method is used to solve this linear system, this filtering occurs in a residual calculation and can be implemented *without* assembly. The action of D_N in this residual requires N solves with the coefficient matrix $(-\delta^2 \Delta^h + 1)$, which has condition number $O(1 + (\frac{\delta}{h})^2)$. On the other hand, if the term $\chi(u_h^*, v_h^*) = \chi(u_h^{**}, v_h)$ is treated explicitly (as tested by Guenaff [Gue04] for $N = 0$), no special care is needed since this term involves filtering a known function $2N$ times. This suggests that for complex 3 dimensional problems, some combination of implicit (for stiff terms) and explicit (for the time relaxation term) methods would be the most efficient.

This section considers alternate forms of the time relaxation term. If the generalized fluctuation operator $v \rightarrow v^*$ is positive semi-definite (as with the van Cittert deconvolution operator $I - D_N^h G_h$), the added term in the method is often simplified to $\chi(u_h^*, v_h)$. In this case some improvement in the error over the usual Galerkin FEM can be shown. This is sketched next.

When $I - D_N^h G_h$ is symmetric and positive,

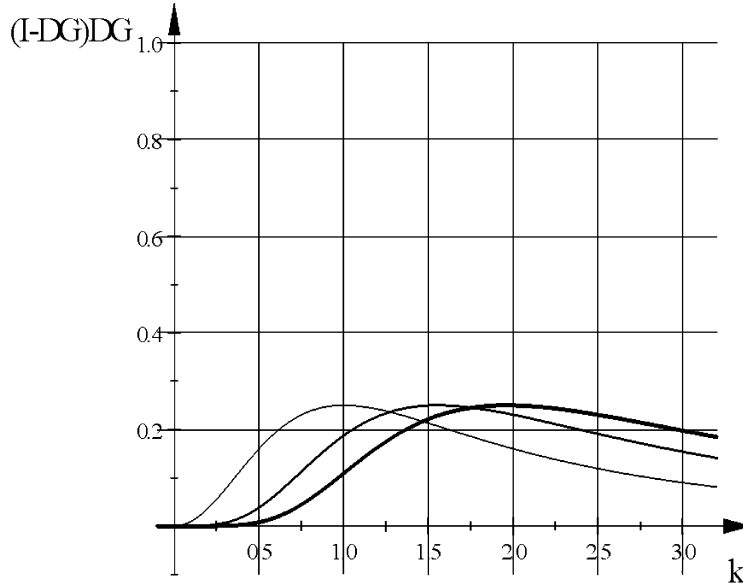
$$u, v \rightarrow ((I - D_N^h G_h)u, v)$$

defines a semi-inner product and semi-norm on $L^2(\Omega)$.

Remark 2. If the orthogonality relation $((I - D_N^h G_h)u, D_N^h G_h v) = 0$ holds, then the two forms are equivalent. In the periodic case it is easy to verify that this orthogonality fails. Indeed, if such a relation held for all h and all u and v , then it would hold in the limit as $h \rightarrow 0$. Then each Fourier mode of $(I - D_N G_h)D_N G_h$ would necessarily vanish. These (rescaling by $\delta k \leftarrow k$) are given by

$$\mathcal{F}[(I - D_N G_h)D_N G_h](k) = [1 - (\frac{k^2}{1 + k^2})^{N+1}](\frac{k^2}{1 + k^2})^{N+1}.$$

This quantity is plotted for $N = 0, 1, 2$ next in the next figure for $0 \leq k\delta \leq \pi$ (rescaled by $\delta k \leftarrow k$). From the plot we see that in the continuum limit there is no orthogonality property. There is however a strengthened CBS inequality of the form $((I - D_N G_h)u, D_N G_h v) \leq \gamma \|u\| \|v\|$, for some $\gamma < \frac{1}{2}$ and that (again in the continuum limit) there is a high order approximate orthogonality for smooth functions.



Thin line: $N = 0$; Medium line: $N = 1$; Thick line: $N = 2$

Definition 4. $(u, v)_* := ((I - D_N^h G_h)u, v)$, and $\|u\|_* := (u, u)_*^{\frac{1}{2}}$. Let $\chi > 0$ and $\delta > 0$. Define the modified projection $\tilde{Q} : X \rightarrow X_h$ as follows. Then $w_h := \tilde{Q}w \in X_h$ is the unique solution of the finite dimensional linear problem:

$$(5.1) \quad (\vec{a} \cdot \nabla(w - w_h) + (w - w_h), v_h) + \chi((w - w_h)^*, v_h) = 0, \forall v_h \in X_h.$$

Following the analysis in Section 3, we have the following projection error estimate.

Theorem 3 (The modified projection's error). *Consider the modified projection (5.1). Then, the error in the projection satisfies*

$$\|w - \tilde{Q}w\|^2 + \chi \|w - \tilde{Q}w\|_*^2 \leq C(N) \inf_{v_h \in X_h} \{ (1 + \delta^{-2}) \|w - v_h\|^2 + \chi^{-1} \|\nabla(w - v_h)\|^2 + \chi \|w - v_h\|_*^2 \}.$$

Proof. Let $\tilde{w} \in X_h$ be arbitrary (for the moment) and write

$$w - \tilde{Q}w = (w - \tilde{w}) - (\tilde{Q}w - \tilde{w}) = \eta - \phi_h,$$

where $\eta := w - \tilde{w}$ and $\phi_h := Qw - \tilde{w} \in X_h$. The projection equations can be rewritten as

$$(\vec{a} \cdot \nabla \phi_h + \phi_h, v_h) + \chi(\phi_h^*, v_h) = (\vec{a} \cdot \nabla \eta + \eta, v_h) + \chi(\eta^*, v_h) = 0, \forall v_h \in X_h.$$

Setting $v_h = \phi_h$, using $(\vec{a} \cdot \nabla \phi_h, \phi_h) = 0$, $(\eta^*, \phi_h) = (\eta, \phi_h)_* \leq \|\eta\|_* \|\phi_h\|_*$ and the usual inequalities gives

$$(5.2) \quad \|\phi_h\|^2 + \chi \|\phi_h\|_*^2 \leq 2(\vec{a} \cdot \nabla \eta, \phi_h) + \|\eta\|^2 + \chi \|\eta\|_*^2.$$

The key term is $(\vec{a} \cdot \nabla \eta, \phi_h)$. This term is split into means and fluctuations and integrated by parts. Indeed, split

$$(\vec{a} \cdot \nabla \eta, \phi_h) = (\vec{a} \cdot \nabla \eta, D_N^h(\overline{\phi_h^h}) + \phi_h^*) = -(\eta, \vec{a} \cdot \nabla D_N^h(\overline{\phi_h^h})) + (\vec{a} \cdot \nabla \eta, \phi_h^*).$$

Inserting this in the right-hand side of (5.2) and using standard inequalities gives

$$\|\phi_h\|^2 + \chi \|\phi_h\|_*^2 \leq \|\eta\|^2 + \chi \|\eta\|_*^2 + \frac{\chi}{2} \|\phi_h\|_*^2 + \frac{C}{2\chi} \|\nabla \eta\|^2 + 2\|\eta\| \|\nabla D_N^h \overline{\phi_h^h}\|.$$

Using the a priori bound $\|\nabla D_N^h(\overline{\phi_h^h})\| \leq C\delta^{-1} \|\phi_h\|$ yields

$$\|\phi_h\|^2 + \chi \|\phi_h\|_*^2 \leq (2 + \delta^{-2}) \|\eta\|^2 + 2\chi \|\eta\|_*^2 + C\chi^{-1} \|\nabla \eta\|^2.$$

Finally, the triangle inequality completes the proof. \square

Consider the modified method: given $\chi > 0$, find $u_h : [0, T] \rightarrow X_h$ satisfying

$$(5.3) \quad (u_{h,t} + \vec{a} \cdot \nabla u_h, v_h) + \chi(u_h^*, v_h) = 0, \forall v_h \in X_h, \\ u_h(x, 0) \text{ approximates } u_0(x) \text{ well.}$$

Adapting the error analysis in Theorem 2 yields the following result.

Theorem 4. *Let \tilde{Q} be the modified equilibrium projection. For $0 \leq t \leq T < \infty$ the error in the method (5.3) satisfies*

$$\sup_{[0, T]} \|u - u_h\|^2 + \int_0^T \chi \|u - u_h\|_*^2 dt \leq C(T, N) \{ \|u(0) - u_h(0)\|^2 + \sup_{[0, T]} \|u - \tilde{Q}u\|^2 \\ + \int_0^T [\chi \|u - \tilde{Q}u\|_*^2 + \|u_t - \tilde{Q}u_t\|^2 + \chi^2 \|u^*(t)\|^2] dt \}.$$

Proof. The error $e(t) := u(t) - u_h(t)$ satisfies

$$(e_t + \vec{a} \cdot \nabla e, v_h) + \chi(e^*, v_h) = \chi(u^*, v_h), \forall v_h \in X_h.$$

Split the error as $e = \eta - \phi_h$, $\eta = u - \tilde{Q}u$, $\phi_h = u_h - \tilde{Q}u$. Rearranging the error equation, gives, for any $v_h \in X_h$,

$$(\phi_{h,t} + \vec{a} \cdot \nabla \phi_h, v_h) + \chi(\phi_h^*, v_h) = (\eta_t - \eta, v_h) + \chi(u^*, v_h) + [(\vec{a} \cdot \nabla \eta + \eta, v_h) + \chi(\eta^*, v_h)].$$

Due to \tilde{Q} the bracketed term vanishes. Setting $v_h = \phi_h$ gives

$$\frac{1}{2} \frac{d}{dt} \|\phi_h\|^2 + \chi \|\phi_h\|_*^2 = (\eta_t - \eta, \phi_h) + \chi(u^*, \phi_h) \\ \leq \frac{1}{2} [\|\eta_t\|^2 + \|\eta\|^2] + \frac{1}{2} \chi^2 \|u^*\|^2 + C \|\phi_h\|^2.$$

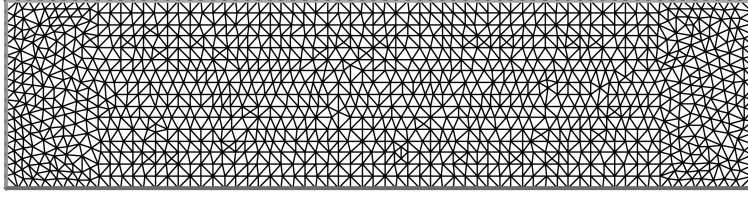


FIGURE 1. One Delaunay mesh used

In the critical step we use the inequality $\chi(u^*, \phi_h) \leq \frac{1}{2}\chi^2\|u^*\|^2 + \frac{1}{2}\|\phi_h\|^2$. The remainder of the proof follows the same as that of Theorem 2. \square

The consistency error is bounded similarly with χ replaced by χ^2 as

$$\int_0^T \chi^2\|u^*(t)\|^2 dt \leq C(u)\chi^2\{h^{2k+2}(1 + (\frac{h}{\delta})^2) + \delta^{2N+2}\}.$$

This gives, with the optimal choices $\delta \simeq \sqrt{h}$, $\chi \simeq h^{-\frac{2}{3}}$ and $N \geq k+1$, that the accuracy for smooth solutions is predicted to be somewhat less than the previous case:

$$\text{error}(t) \lesssim \text{error}(0) + Ch^{k+\frac{1}{3}}.$$

Remark 3. We believe this estimate might be improvable. For example, one can try instead $\chi(u^*, \phi_h) = \chi(u, \phi_h)_* \leq \frac{1}{2}\chi\|u\|_*^2 + \frac{1}{2}\chi\|\phi_h\|_*^2$. Another possibility is that the consistency error estimate might be improvable through estimates in negative Sobolev norms because, for example,

$$\begin{aligned} \int_0^T \chi\|u\|_*^2 dt &= \int_0^T \chi(u, u)_* dt = \int_0^T \chi(u^*, u) dt \\ &\leq \sqrt{\int_0^T \chi\|u\|_{H^1(\Omega)}^2 dt} \sqrt{\int_0^T \chi\|u^*\|_{H^{-1}(\Omega)}^2 dt}. \end{aligned}$$

This is an open problem.

6. FOUR COMPUTATIONAL ILLUSTRATIONS

There have been several simplifying assumptions in the formulation of the model hyperbolic problem, including that the problem is linear and has periodic boundary conditions. Nonlinear conservation laws require special techniques that are beyond this report so the tests will focus upon the behavior of the method when the other simplification is relaxed. We let the spacial domain be a rectangle in 2 dimensional. It is well known that many methods can have special (usually favorable) properties on uniform meshes and on meshes that are aligned exactly with the convection direction. To remove this effect, we always use meshes generated by a Delaunay algorithm. A typical example is plotted in Figure 1.

The 2 dimensional calculations were done using FreeFEM++, [HePi] and the 1 dimensional calculations in Section 6.1 were done with a MatLab program. With unstructured meshes, we can select the convecting velocity to be $\vec{a} = (1, 0)$. Thus we have the following test problem with right-hand side chosen so that the true

solution is

$$\begin{aligned} u_{\text{true}}(x, y, t) &= \sin(4\pi y) \sin(\pi x) \sin(t), \text{ so that} \\ f(x, y, t) &= \sin(4\pi y) (\sin(\pi x) \cos(t) + \cos(\pi x) \sin(t)), \end{aligned}$$

given by find $u = u(x, y, t)$ satisfying

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} &= f(x, y, t), \text{ on } \Omega = (0, 1) \times (0, \frac{1}{4}), \text{ and } t > 0, \\ u(x, y, 0) &= 0, \text{ on } \Omega, \ u(0, y, t) = 0 \text{ for } t > 0, 0 < y < \frac{1}{4}. \end{aligned}$$

In the first table we give the error in the usual finite element method using quadratics on triangles with maximum triangle diameter given as h for the Delaunay mesh generated. The rates were calculated in the table in the standard way using the error at two successive h 's and supposing $\text{error}(h) = Ch^a$ and then solving for the exponent a .

	$h =$	L^2 error	rate
	4.62798e-2	2.41113e-4	—
(6.1)	2.54801e-2	6.66330e-5	2.155
	1.30216e-2	1.64784e-5	2.081
	6.52674e-3	3.83958e-6	2.110
	3.44127e-3	9.62077e-7	2.162
	Usual FEM errors		

A line of best fit through a log-log plot of errors gives convergence rate 2.116 for the usual FEM. This is consistent with an $O(h^2)$ error estimate for quadratics for the usual FEM. Next we add a time relaxation term to this test problem to test if, with the scaling predicted by the theory (in the simplified context) the accuracy does increase.

The fact that the theory is in a simplified context is potentially important. The averaging operator is the solution of a second order problem and the PDE is only first order and thus the PDE does not have enough boundary conditions for the averaging operator. It can be argued that many convection dominated problems contain small amounts of diffusion and that including this and the accompanying boundary conditions would completely resolve this issue. We, however, wanted to see how the method could perform for the pure hyperbolic limit and test the limitations of the error analysis. We also note that the question of boundary conditions for finite element methods for even 2×2 systems contains extra subtleties; [L83b] and Gunzburger [G77].

As a first (reasonable) guess for the extra boundary conditions needed, we exploited the fact that in our chosen time stepping method we were always averaging a known function. Thus, given a function ϕ vanishing at the inflow boundary $x = 0$, we calculated its average $\phi \rightarrow \bar{\phi}^h$ by the usual FEM approximation of

$$\begin{aligned} -\delta^2 \Delta \bar{\phi} + \bar{\phi} &= \phi, \text{ in } \Omega, \\ \bar{\phi}(0, y) &= \phi(= 0), \text{ on the inflow boundary} \\ \bar{\phi} &= \phi, \text{ on the rest of the boundary.} \end{aligned}$$

Other definitions are possible and can be explored. The parameter values tested were $\delta = 0.1\sqrt{h}$, $\chi = 1/h$, and $N = 2$ deconvolution steps. These choices agree with the theoretical predictions of values that increase accuracy with quadratic elements. The following errors were observed. A line of best fit to the log-log plot of errors gives convergence rate 2.668, consistent with the theoretical prediction of 2.5. A similar test (omitted for space) that was performed also showed no convergence rate improvement over the usual FEM when using $N = 0$ and $N = 1$ deconvolution orders, also as predicted.

	$h =$	L^2 error	rate
	4.62798e-2	1.50848e-4	—
(6.2)	2.54801e-2	2.08445e-5	3.316
	1.30216e-2	3.10657e-6	2.836
	6.52674e-3	5.53691e-7	2.497
	3.44127e-3	9.59315e-8	2.739
	Time Relax FEM errors		

We note that although the exact solution is chosen to be exactly zero on the boundary, the approximate solution is zero only on the inflow boundary and only approximately zero on the rest of the boundary. One can ask if the extra boundary conditions in the averaging operator played any role in the errors. This can also be easily tested using the knowledge that the true solution is exactly zero on $\partial\Omega$. To do this we repeated the test but redefined the averaging operator $\phi \rightarrow \bar{\phi}^h$ by the usual FEM approximation of

$$-\delta^2 \Delta \bar{\phi} + \bar{\phi} = \phi, \text{ in } \Omega, \quad \bar{\phi} = 0, \text{ on } \partial\Omega.$$

With the same parameters we observed the following errors.

	$h =$	L^2 error	rate
	4.62798e-2	1.13634e-4	—
(6.3)	2.54801e-2	1.87568e-5	3.018
	1.30216e-2	2.55018e-6	2.972
	6.52674e-3	3.20682e-7	3.002
	3.44127e-3	4.28390e-8	3.145
	Time Relax FEM errors using extra BCs		

A line of best fit to log-log plot of errors gives convergence rate 3.034. This convergence rate is beyond the theory.

The second test is a problem with a nonsmooth solution. For the same domain, test problem, meshes, algorithmic parameters (δ, χ, N) , pick the body forces to be identically zero, $f(x, t) \equiv 0$. We choose a discontinuous boundary condition, for $j = 1, 2$,

$$u(0, y, t) = 0, 0 \leq x \leq \frac{1}{8},$$

$$u(0, y, t) = 1, \frac{1}{8} < x \leq \frac{1}{4}.$$

The initial condition was

$$u(x, y, 0) = 0, 0 \leq x \leq \frac{1}{8},$$

$$u(x, y, 0) = e^{-x}, \frac{1}{8} < x \leq \frac{1}{4}.$$

The exact solution is a discontinuity that moves across the domain so that after $t = 1$ it reaches a steady discontinuous step. The usual FEM on a general unstructured mesh without any upwinding or limiters is particularly unsuitable for problems like this one. Thus, we do not expect excellent solution quality from any variation of the methods studied. We test this problem to see if the time relaxation term gives any improvement at all for nonsmooth solutions. We give next approximate solution plots at $t = 1$ beginning with the (expected bad) usual FEM in Figure 2.

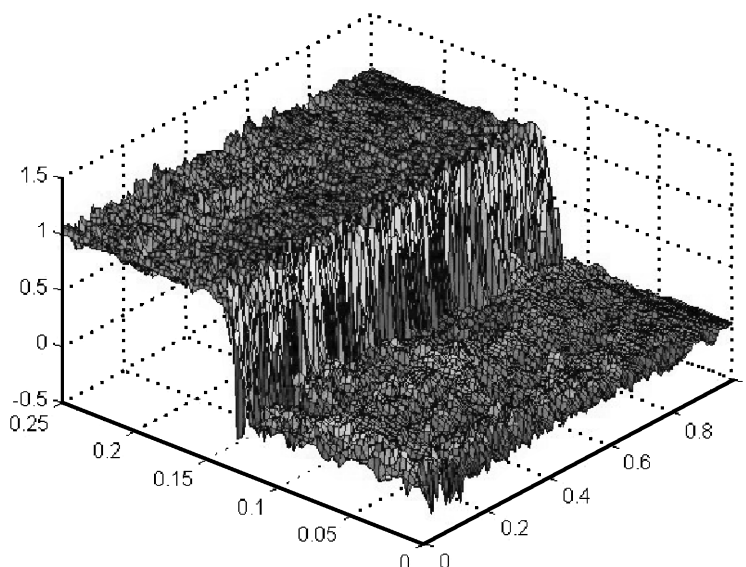


FIGURE 2. Usual FEM at $t = 1$

The behavior is actually worse than the above seems to indicate. This is revealed when one zooms in to the solution in Figure 3.

Next we give the approximate solution obtained using the FEM+time relaxation for the same data and algorithmic parameters in Figure 4. Oscillations still occur. (The aim of time relaxation is not to prevent all oscillations but rather to damp those that do occur.) However, compared with Figure 2, the oscillations are much smaller than without time relaxation and the larger ones are confined to a neighborhood of the discontinuity. Zooming in to the worst subregion in Figure 5, compared to Figure 3, is consistent with the above description.

By computing the solution for various values of δ , we found the best value of the averaging radius for this problem to be to $\delta = 0.03\sqrt{h}$; see Figure 6, below.

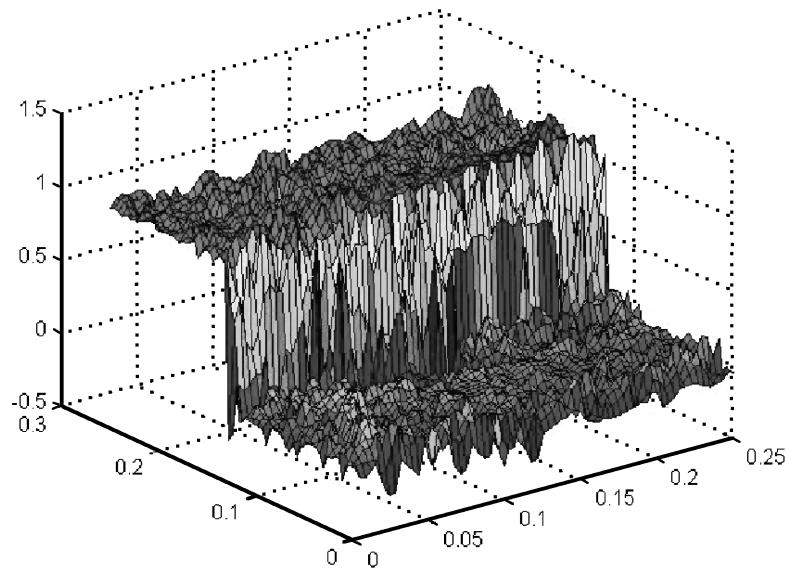


FIGURE 3. A zoom of the usual FEM solution

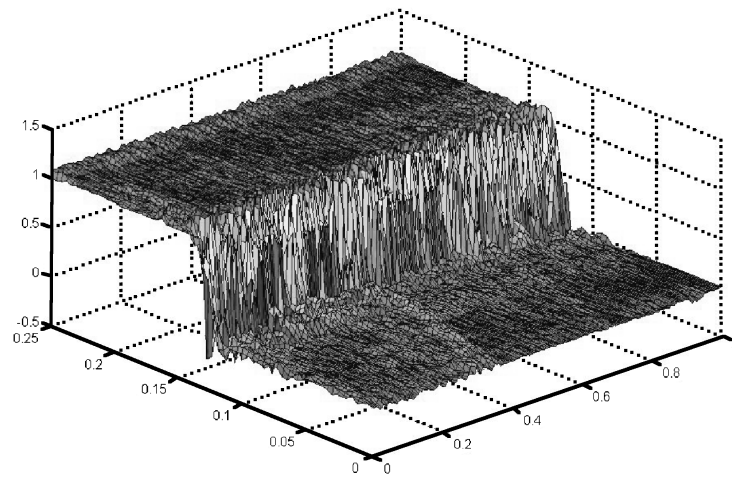


FIGURE 4. FEM+time relaxation solution

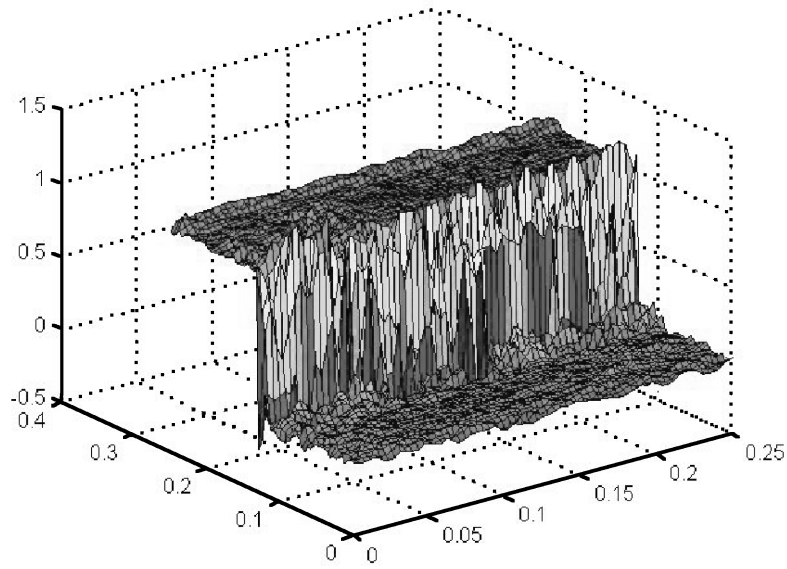


FIGURE 5. Zoom of the FEM+time relaxation solution

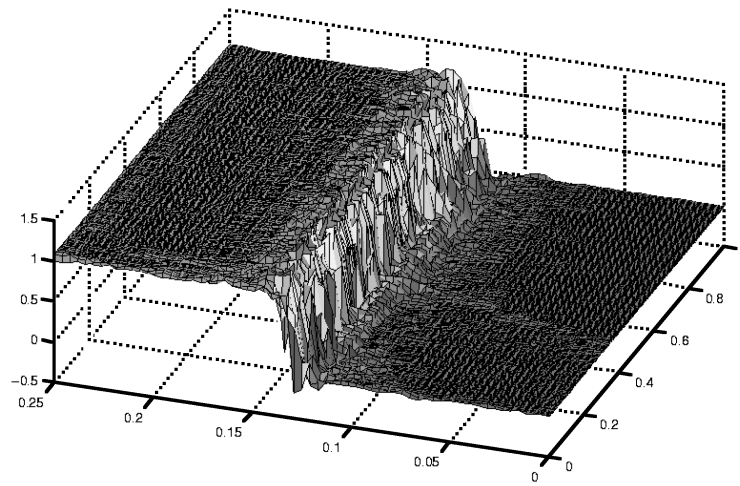


FIGURE 6. Time relaxation with $\delta = 0.03\sqrt{h}$

6.1. Tests in one space dimension: the one dimensional case is special.

Hyperbolic problems in one space dimension are special for several reasons, and there are cases known in the literature where faster convergence is observed (or proven) in one space dimension than in higher dimensions. This section examines in two tests if the rates of convergence for the time relaxation FEM (that have both been proven and observed in higher dimensions) might be potentially improvable under the special features of one space dimension. For the one dimensional tests, take $\Omega = (0, 1)$. Both uniform and nonuniform meshes are considered to address the possibility of special properties (e.g., superconvergence and leading order error cancellation) often observed with uniform meshes. Nonuniform meshes were generated by choosing a maximum mesh size $h \ll 1$, then assigning grid points starting at $x_0 = 0$ followed by $x_1 = h$, then alternately adding $x_{i+1} = x_i + h/2$, $x_{i+2} = x_{i+1} + h$ until the final grid point $x_N = 1$.

Calculations were performed using a MatLab program and continuous piecewise quadratic elements. Choosing a true solution $u(x, t) = \sin(\pi x) \sin(t)$, we calculate the right-hand side function $f(x, t) = \sin(\pi x) \cos(t) + \pi \cos(\pi x) \sin(t)$. We then have the problem of approximating $u(x, t)$, satisfying

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} &= f(x, t), \quad (x, t) \in (0, 1) \times (0, 1), \\ u(x, 0) &= 0, \quad x \in (0, 1) \\ u(0, t) &= 0, \quad t \in (0, 1). \end{aligned}$$

With this choice of u, f the following error tables were generated illustrating accuracy and convergence rates for the finite element method with and without time relaxation. The maximum size of a mesh element is denoted by h . Time discretization was performed using the trapezoidal rule, and a time step size of $\Delta t = 1/2000$ was determined to be sufficiently small for all tests for the time discretization error to be negligible. Convergence rates, the exponent p in error $\simeq C h^p$, were calculated using the maximum L^2 spacial error over all time steps. Parameter values $\delta = 0.3 \cdot h^{1/2}$ and $\chi = \frac{2.0}{h}$ were chosen, consistent with the theory for the time relaxed FEM.

The convergence rates in Table 1 for both regular and time relaxed FEM are unexpected. For 1 dimensional problems with periodic boundary conditions, the regular FEM with continuous piecewise quadratic elements on a uniform mesh has a provably optimal convergence rate $O(h^3)$; see [AG79]. Thus, the shift from periodic to Dirichlet boundary conditions at the inflow boundary points to an open problem in the theory. The observed optimal convergence rate of the time relaxed FEM in 1 dimension is better than the predicted rate $O(h^{2.5})$. Table 2 suggests that this is not a result of the mesh being uniform, but may be a special property of the method in 1 dimension (see the 2 dimension computations above). The addition of the time relaxation term results in an order of magnitude improvement in accuracy and rate of convergence for the 1 dimensional problem. Proving this observed improvement (from $O(h^{2.5})$ to $O(h^3)$) is an open problem.

TABLE 1. Maximum L^2 errors and convergence rates, uniform mesh.

h	Regular FEM		Time relaxed FEM	
	Max. Error $L^2(\Omega)$	Rate	Max. Error $L^2(\Omega)$	Rate
1/10	4.485128e-4	—	3.308497e-4	—
1/20	1.120078e-4	2.002	4.994900e-5	2.728
1/40	2.799802e-5	2.000	6.537910e-6	2.934
1/80	7.002847e-6	1.999	8.298419e-7	2.978
1/160	1.754548e-6	1.997	1.046213e-7	2.988

TABLE 2. Maximum L^2 errors and convergence rates, nonuniform mesh.

h	Regular FEM		Time relaxed FEM	
	Max. Error $L^2(\Omega)$	Rate	Max. Error $L^2(\Omega)$	Rate
1/10	3.599619e-4	—	2.779811e-4	—
1/20	8.275052e-5	2.121	4.001149e-5	2.797
1/40	2.135513e-5	1.954	5.390944e-6	2.892
1/80	5.215845e-6	2.034	6.895017e-7	2.967
1/160	1.322736e-6	1.979	8.730697e-8	2.981

The time relaxation method is also expected to improve the quality of approximations of nonsmooth solutions. This was observed in 2 dimensions and will now be tested in 1 dimension where the figures are clearer to interpret. Consider solving the following modified problem:

$$\begin{aligned}
 \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} &= 0, (x, t) \in (0, 1) \times (0, 1), \\
 u(x, 0) &= 1, x \in (0, 0.5], \\
 u(x, 0) &= 0, x \in (0.5, 1), \\
 u(0, t) &= 1, t \in (0, 1).
 \end{aligned}$$

The true solution is a step function with discontinuity moving toward $x = 1$ at a constant rate $\frac{dx}{dt} = 1$. Computations were performed using a uniform mesh with mesh size $h = 1/40$. For the time relaxation method, parameter values $\delta = 0.05(h^{1/2})$ and $\chi = \frac{2.0}{h}$ were chosen. The next figure shows approximations generated using both methods at four times. Both solutions initially behave the same and the discontinuity induces error propagation over time. The time relaxation method can be seen to dissipate spurious oscillations, as intended, and quickly settles into a more accurate representation of the true solution.

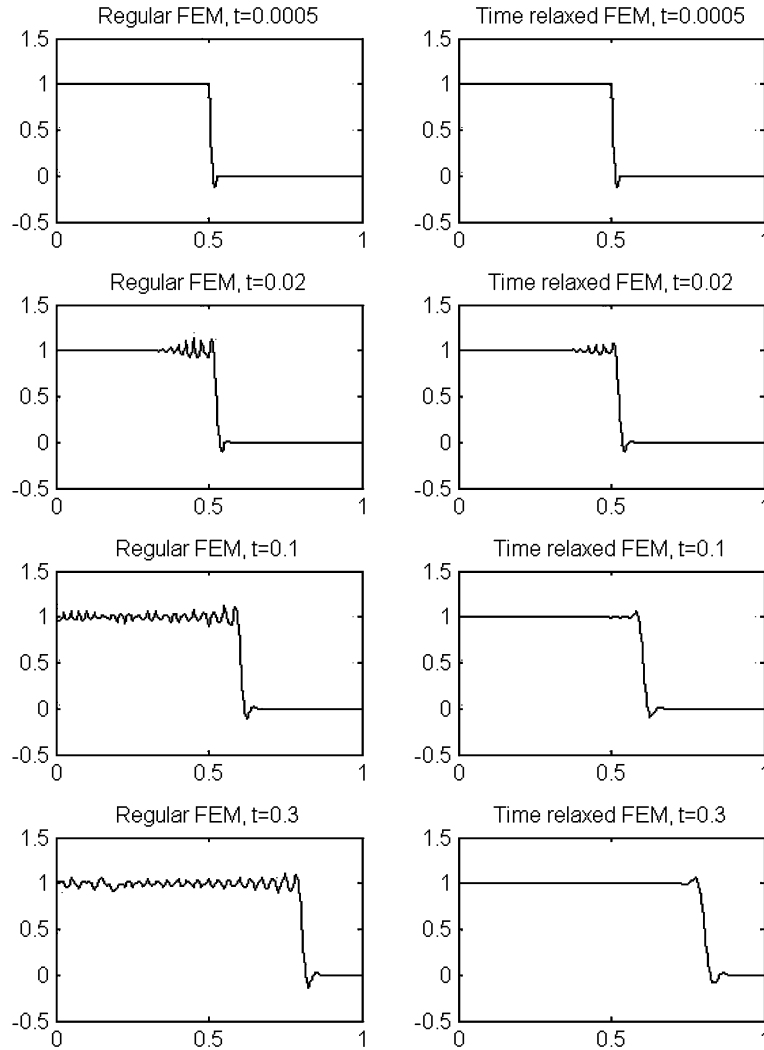


FIGURE 7. Comparison of regular, time relaxed FEM approximations.

7. CONCLUSIONS

The theory of the regularized Chapman-Enskog expansions of conservation laws in Rosenau [R89] and Schochet and Tadmor [ST92] suggest scaling $\chi \simeq \delta^{-1}$. Numerical analysis of the error in the method (for smooth solutions) suggests that there is a difference in the discrete case and instead $\chi \simeq \delta^{-2}$. If one views the mesh-width h as the induced filter width instead, we recover the scaling $\chi \simeq (\text{filter width})^{-1}$. On the other hand, in simulations of turbulent compressible flow, it is common practice to take $\delta = O(h)$ (for example, $\delta = 3h$) to try to squeeze maximum information from a given resolution. The numerical analysis herein suggests this might be overreaching and predicts $\delta \simeq \sqrt{h}$ as more accurate on the large scales. This is confirmed by the initial and very limited tests herein.

The last table of the 2 dimension tests also suggests that greater accuracy than proven herein might be hiding in the method (possibly for special elements) with a better averaging or specification of extra boundary conditions. There are many (and obvious) possibilities testable. There are also cases where the $O(h^{k+\frac{1}{2}})$ error estimate can be provably improved to $O(h^{k+1})$. When the finite element space consists of even order splines on a uniform mesh, super convergence of the time relaxation discretization has been proven [L07b] and also implies optimality in L^2 . More generally, on a uniform or near the uniform mesh, the cancellation argument of Dupont [Du73] can be adapted provided a finite element basis exists which is symmetric about the node associated with each basis function. This includes continuous, piecewise linears and cubic splines (Dupont [Du73]), continuous quadratics and cubics (Axelsson and Gustafsson [AG79]), but not C^1 Hermite cubics (Dupont [Du73], Hedstrom [Hed79]). The tests in 1 dimension strongly suggest that proven rates of convergence, when restricted to one space dimension, might be improvable by $\frac{1}{2}$ power of h . This is an open problem.

The convergence result herein extends immediately to multidimension Friedrichs systems with nonperiodic boundary conditions. Some extensions to convection dominated, convection diffusion problems are possible but possibly more delicate due to boundary and interior layers. Extension to nonlinear conservation laws is more delicate and depends on structure of the nonlinearity in the specific conservation law. It should be done in connection with limiters. Showing that the L^2 accuracy for slightly viscous Navier-Stokes equations is greater than that proven [ELN07] is also an interesting and important open problem.

7.1. Averaging and deconvolution operators. To introduce the time relaxation discretization, the treatment of boundary conditions by the differential filter must be specified. With second order differential filters extra boundary conditions, beyond those of the first order continuous problem, must be supplied for the differential filter. This difficulty does not occur when solving convection diffusion equations or other (linear or nonlinear) second order problems. The simplest idea of specifying $\bar{\phi} = \phi$ on $\partial\Omega$ when filtering a known function ϕ was used herein and seemed to work.

The differential filter used herein is natural for finite element methods, second order problems and the well developed tools of finite element error analysis. However, it is also only *approximately* local. For hyperbolic problems, it is quite possible that a purely local averaging is preferable. Small but global averaging effects couple the approximate solution away from the step to the behavior at the discontinuity. Thus, it might contribute to the small background oscillations seen away from the discontinuity in Figures 4 and 3. This needs to be tested.

Since deconvolution is a well known and important ill posed problem, there are very many deconvolution operators available for testing. Bertero and Boccacci [BB98] give many examples and we note the very interesting construction of Geurts [Geu97].

REFERENCES

- [AL99] N. A. Adams and A. Leonard, *Deconvolution of subgrid scales for the simulation of shock-turbulence interaction*, p. 201 in: *Direct and Large Eddy Simulation III*, (eds.: P. Voke, N. D. Sandham and L. Kleiser), Kluwer, Dordrecht, 1999.
- [AS01] N. A. Adams and S. Stolz, *Deconvolution methods for subgrid-scale approximation in large eddy simulation*, *Modern Simulation Strategies for Turbulent Flow*, R. T. Edwards, 2001.
- [AS02] N. A. Adams and S. Stolz, *A subgrid-scale deconvolution approach for shock capturing*, J. C. P., 178 (2002), 391-426. MR1899182 (2003b:76100)
- [AG79] A. O. H. Axelsson and J. Gustafsson, *Quasioptimal finite element approximation of first order hyperbolic and convection-diffusion equations*, pp. 273-281 in: *Analytical and Numerical Approaches to Asymptotic Problems in Analysis*, (eds: A. O. H. Axelsson, L. S. Frank, and A. van der Sluis) North Holland, Amsterdam, 1980.
- [BIL06] L. C. Berselli, T. Iliescu and W. Layton, *Mathematics of Large Eddy Simulation of Turbulent Flows*, Springer, Berlin, 2006 MR2185509 (2006h:76071)
- [BB98] M. Bertero and B. Boccacci, *Introduction to Inverse Problems in Imaging*, IOP Publishing Ltd., 1998. MR1640759 (99k:78027)
- [BBJL07] M. Braack, E. Burman, V. John and G. Lube, *Stabilized FEMs for the generalized Oseen problem*, CMAME 196(2007), 853-866. MR2278180 (2007i:76065)
- [BH80] A. Brooks and T. J. R. Hughes, *Streamline upwind Petrov-Galerkin methods for advection-dominated flows*, in: 3rd Int. Conf. on FEM in Fluid Flow, 1980.
- [BH04] E. Burman and P. Hansbo, *Edge stabilizations for Galerkin approximations of convection-diffusion-reaction problems*, CMAME 193(2004), 1437-1453. MR2068903 (2005d:65186)
- [C79] G. F. Carey, *An analysis of stability and oscillations in convection-diffusion computations*, 63-71 in: *FEM for Convection Dominated Flows*, T. J. R. Hughes, ed., ASME, vol. 34, 1979. MR571683 (81i:65072)
- [D04] A. Dunca, *Space averaged Navier-Stokes equations in the presence of walls*, Ph.D. Thesis, University of Pittsburgh, 2004.
- [DE06] A. Dunca and Y. Epshteyn, *On the Stolz-Adams de-convolution model for the large eddy simulation of turbulent flows*, SIAM J. Math. Anal. 37(2006), 1890-1902. MR2213398 (2006k:76066)
- [Du73] T. Dupont, *Galerkin methods for first order hyperbolic: an example*, SINUM 10(1973), 890-899. MR0349046 (50:1540)
- [ELN07] V. Ervin, W. Layton and M. Neda, *Numerical analysis of a higher order time relaxation model of fluids*, Int. J. Numer. Anal. and Modeling, 4(2007), 648-670. MR2344062 (2008j:76035)
- [Gue04] R. Guenanff, *Non-stationary coupling of Navier-Stokes/Euler for the generation and radiation of aerodynamic noises*, PhD thesis: Dept. of Mathematics, Université Rennes 1, Rennes, France, 2004.
- [Guer99] J.-L. Guermond, *Subgrid stabilization of Galerkin approximations of monotone operators*, C. R. Acad. Sci. Paris, Série I, 328 7 (1999) 617-622. MR1680045 (99k:65049)
- [Ger86] M. Germano, *Differential filters of elliptic type*, Phys. Fluids, 29(1986), 1757-1758. MR845232 (87h:76075b)
- [Geu97] B. J. Geurts, *Inverse modeling for large eddy simulation*, Phys. Fluids, 9(1997), 3585.
- [G77] M. Gunzburger, *On the stability of Galerkin methods for Initial-Boundary Value Problems for hyperbolic systems*, Math. Comp. 31 (1977) 661-675. MR0436624 (55:9567)
- [HePi] F. Hecht and O. Pironneau, *FreeFEM++*, webpage: <http://www.freefem.org>.
- [Hed79] G. W. Hedstrom, *The Galerkin Method Based on Hermite Cubics*, SINUM, 16(1979), 385-393. MR530476 (80i:65116)
- [H99] T. J. Hughes, *Multiscale phenomena: the Dirichlet to Neumann formulation, subgrid models, bubbles and the origin of stabilized methods*, CMAME 127(1999), 387-401.
- [L83] W. Layton, *Galerkin Methods for Two-Point Boundary Value Problems for First Order Systems*, SINUM 20 (1983), 161-171. MR687374 (84d:65059)
- [L83b] W. Layton, *Stable Galerkin Methods for Hyperbolic Systems*, SINUM 20 (1983), 221-233. MR694515 (85c:65120)

- [L07] W Layton, *A remark on regularity of an elliptic-elliptic singular perturbation problem*, technical report, <http://www.mathematics.pitt.edu/research/technical-reports.php>, 2007.
- [L07b] W. Layton, *Superconvergence of finite element discretization of time relaxation models of advection*, BIT, 47(2007) 565-576. MR2338532 (2009g:65127)
- [L05] W. Layton, *Model reduction by constraints and an induced pressure stabilization*, J. Numer. Linear Algb. and Applications, 12(2005), 547-562. MR2150167 (2006e:76083)
- [L02] W Layton, *A connection between subgrid-scale eddy viscosity and mixed methods*, Applied Math and Computing, 133[2002],147-157 MR1923189 (2003h:76102)
- [LL03] W. Layton and R. Lewandowski, *A simple and stable scale similarity model for large eddy simulation: energy balance and existence of weak solutions*, Applied Math. Letters 16(2003), 1205-1209. MR2015713 (2004j:76095)
- [LL05] W. Layton and R. Lewandowski, *Residual stress of approximate deconvolution large eddy simulation models of turbulence*. Journal of Turbulence, 46(2006), 1-21.
- [LL06a] W. Layton and R Lewandowski, *On a well posed turbulence model*, Discrete and Continuous Dynamical Systems, Series B, 6(2006) 111-128. MR2172198 (2006g:76054)
- [LMNR08] W. Layton, C. Manica, M. Neda and L. Rebholz, *Numerical analysis of a high accuracy Leray-deconvolution model of turbulence*, Numerical Methods for PDEs, 24(2008) 555-582. MR2382797 (2009b:76064)
- [LMNR06] W. Layton, C. Manica, M. Neda and L. Rebholz, *The joint Helicity-Energy cascade for homogeneous, isotropic turbulence generated by approximate deconvolution models*, Adv. and Appls. in Fluid Mechanics, 4(2008), 1-46. MR2488583
- [LS07] W. Layton and I. Stanculescu, *K41 optimized deconvolution models*, Int. J. Computing and Mathematics, 1(2007), 396-411. MR2396390 (2009a:76124)
- [LN07] W. Layton and M. Neda, *Truncation of scales by time relaxation*, JMAA 325(2007), 788-807. MR2270051 (2008c:76049)
- [LR74] P. Lesaint and P. A. Raviart, *On a finite element method for solving the neutron transport equation*, 89-112 in: *Mathematical aspects of finite elements in partial differential equations*, C. de Boor, ed., Academic Press, N.Y., 1974. MR0658142 (58:31918)
- [MM06] C. C. Manica and S. Kaya-Merdan, *Convergence Analysis of the Finite Element Method for a Fundamental Model in Turbulence*, tech. report, <http://www.math.pitt.edu/techreports.html>, 2006.
- [MS07] C. C. Manica and I. Stanculescu, *Numerical Analysis of Tikhonov Deconvolution Model*, University of Pittsburgh Technical Report, <http://www.math.pitt.edu/techreports.html>, 2007.
- [MLF03] J. Mathew, R. Lechner, H. Foysi, J. Sesterhenn and R. Friedrich, *An explicit filtering method for large eddy simulation of compressible flows*, Physics of Fluids 15(2003), 2279-2289.
- [RST96] H.-G. Roos, M. Stynes and L. Tobiska, *Numerical Methods for Singularly Perturbed Differential Equations*, Springer, Berlin, 1996. MR1477665 (99a:65134)
- [R89] Ph. Rosenau, *Extending hydrodynamics via the regularization of the Chapman-Enskog expansion*, Phys. Rev.A 40 (1989), 7193. MR1031939 (91b:82047)
- [S01] P. Sagaut, *Large eddy simulation for Incompressible flows*, Springer, Berlin, 2001. MR1815221 (2002f:76047)
- [SW83] A.H. Schatz and L. Wahlbin, *On the finite element method for a singularly perturbed reaction-diffusion equation in two and one dimension*, Math. Comp., 40(1983) 47-79. MR679434 (84c:65137)
- [ST92] S. Schochet and E. Tadmor, *The regularized Chapman-Enskog expansion for scalar conservation laws*, Arch. Rat. Mech. Anal. 119 (1992), 95. MR1176360 (93f:35191)
- [S07] I. Stanculescu, *Existence Theory of Abstract Approximate Deconvolution Models of Turbulence*, Annali dell'Università di Ferrara, 54(2008), 145-168. MR2403379 (2009d:76058)
- [SA99] S. Stolz and N. A. Adams, *On the approximate deconvolution procedure for LES*, Phys. Fluids, II(1999),1699-1701.
- [SAK01a] S. Stolz, N. A. Adams and L. Kleiser, *The approximate deconvolution model for LES of compressible flows and its application to shock-turbulent-boundary-layer interaction*, Phys. Fluids 13 (2001),2985.

- [SAK01b] S. Stolz, N. A. Adams and L. Kleiser, *An approximate deconvolution model for large eddy simulation with application to wall-bounded flows*, Phys. Fluids 13 (2001), 997.
- [SAK02] S. Stolz, N. A. Adams and L. Kleiser, *The approximate deconvolution model for compressible flows: isotropic turbulence and shock-boundary-layer interaction*, in: Advances in LES of complex flows (editors: R. Friedrich and W. Rodi) Kluwer, Dordrecht, 2002.
- [TW74] V. Thomée and B. Wendroff, *Convergence estimates for Galerkin methods for variable coefficient initial value problems*, SINUM 11(1973), 1059-1068. MR0371088 (51:7309)
- [vC31] P. van Cittert, *Zum Einfluss der Spaltbreite auf die Intensitätsverteilung in Spektrallinien II*, Zeit. für Physik 69 (1931), 298-308.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PITTSBURGH, PITTSBURGH, PENNSYLVANIA 15260

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PITTSBURGH, PITTSBURGH, PENNSYLVANIA 15260

E-mail address: `wjl@pitt.edu`