ICES REPORT 10-35

August 2010

Variational Multiscale Analysis: A New Link Between Flux Correction, Total Variation, and Constrained Optimization

by

John A. Evans and Thomas J.R. Hughes

The Institute for Computational Engineering and Sciences
The University of Texas at Austin
Austin, Texas 78712

Reference: John A. Evans and Thomas J.R. Hughes, "Variational Multiscale Analysis: A New Link Between Flux Correction, Total Variation, and Constrained Optimization", ICES REPORT 10-35, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, August 2010.

Variational multiscale analysis: A new link between flux correction, total variation, and constrained optimization

John A. Evans[a,*] and Thomas J.R. Hughes[a]

[a] The University of Texas at Austin
Institute for Computational Engineering and Sciences
1 University Station C0200
Austin, TX 78712-0027, U.S.A.

[*] Corresponding author. E-mail address: evans@ices.utexas.edu

Abstract

In this paper, we present a new variational multiscale method that arises through the enforcement of a total variation constraint. After presenting the method in full generality with application to the repair of stabilized methods, we specialize to the case of a Fourier basis and show the resulting method is one of flux correction. This provides a new theoretical basis for the use of flux correction in numerical simulation and inspires a new class of discontinuity capturing techniques based on the concept of variational flux correction.

Key words. variational multiscale analysis, flux correction, total variation, transport, discontinuity capturing, stabilized methods

1 Introduction

Discontinuity capturing has played a prominent role in numerical simulation since the early days of computation. The practice of discontinuity capturing, however, has proven to be more of an art than a science, and *ad hoc* tools are usually employed to introduce artificial diffusion near discontinuities and layers and thus eliminate spurious oscillations.

In this paper, we take a new look at discontinuity capturing through the lens of variational multiscale analysis. The variational multiscale method [15, 16] was introduced as a framework for incorporating missing fine-scale effects into numerical problems governing coarse-scale behavior. It has provided a rationale for stabilized methods, and a platform for the development of new computational methods. Construction of the method is as follows: decompose the solution u to a partial differential equation into a sum of two components $\bar{u}$ and u', determine the fine-scale component u' analytically in terms of the coarse-scale component $\bar{u}$, and solve for $\bar{u}$ numerically. The original instantiation of the method was based on variational projection. That is, the decomposition of the solution into a sum of coarse-scale and fine-scale components is uniquely specified by identifying a projector from the space of all scales onto the coarse-scale subspace. However, variational projection is not guaranteed to preserve certain properties of the exact solution such as the maximum principle or entropy conservation. To address this issue, the variational multiscale method was extended to handle convex constraints in [8]. In this setting, the coarse-scale component is uniquely specified as the solution of a constrained optimization problem.

In order to arrive at a method which is both monotone and theoretically optimal, we present a new variational multiscale method for the convection-diffusion-reaction equation that arises through the enforcement of a total variation constraint. After presenting the method in full generality, we specialize to the case of Fourier discretizations. With a special case of projector, we show that the method reduces to one of flux correction, revealing a link between flux correction, total variation, and constrained optimization.

In Section 2, we briefly review the concept of constrained variational multiscale analysis. In Section 3, we restrict ourselves to the setting of the advection-diffusion-reaction equation and derive a new total variation bounded variational multiscale method. We then apply this new method to the repair of the Streamline-Upwind/Petrov Galerkin method. In Section 4, we specialize to the case of a Fourier basis. An in-depth analysis reveals that variational flux correction is responsible for discontinuity capturing, and a further inspection in one spatial dimension is conducted. In Section 5, we draw conclusions.

2 Constrained variational multiscale analysis

In this section, we briefly review the abstract framework introduced in [8] which allows for the direct enforcement of constraints in the variational multiscale method.

2.1 The abstract problem

Let V be a Hilbert space, endowed with a scalar product $(\cdot,\cdot)_V$ and induced norm $\|\cdot\|_V$. Let V^* be the dual of V and let ${}_{V^*}\langle\cdot,\cdot\rangle_V$ be the pairing between them. Let $\mathcal{L}: V \to V^*$ be a linear isomorphism. Given $\mathcal{F} \in V^*$, we consider the abstract problem: find $u \in V$ such that

$$\mathcal{L}u = \mathcal{F}. \tag{1}$$

The variational formulation of (1) is: find $u \in V$ such that

$${}_{V^*}\langle \mathcal{L}u, v\rangle_V = {}_{V^*}\langle \mathcal{F}, v\rangle_V, \quad \forall v \in V. \tag{2}$$

The solution u can be formally expressed as $u = \mathcal{G}\mathcal{F}$, where $\mathcal{G}: V^* \to V$ is the Green's (or solution) operator. That is, $\mathcal{G} = \mathcal{L}^{-1}$.

2.2 The constrained VMS formulation

We are interested in finding an approximation $\bar{u}$ belonging to a closed finite-dimensional subspace $\bar{V} \subset V$ to the solution u of (1). In the variational multiscale (VMS) approach, $\bar{V}$ represents the space of computable *coarse scales*.

Define the constraint set

$$\bar{K} = \left\{ \bar{v} \in \bar{V} : \mathbf{f}(\bar{v}) = \mathbf{0}, \mathbf{g}(\bar{v}) \leq \mathbf{0} \right\} \tag{3}$$

where $\mathbf{f}: \bar{V} \to \mathbb{R}^{n_{eq}}$ and $\mathbf{g}: \bar{V} \to \mathbb{R}^{n_{ineq}}$ are continuously differentiable (in the sense of Fréchet) vector functions of the finite-dimensional space $\bar{V}$. We assume that this set is not empty. The constraint set $\bar{K}$ represents the set of admissible approximations. For the purposes of this paper, this constraint set will represent the set of *monotone* approximations. We seek an admissible approximation which will minimize the error with respect to some induced norm.

Let H be a Hilbert space such that V is continuously embedded in H. Let H be endowed with a scalar product $(\cdot,\cdot)_H$ and induced norm $\|\cdot\|_H$. The constrained optimization problem for $\bar{u}$ is as follows: find $\bar{u} \in \bar{K}$ such that

$$\frac{1}{2}\|u - \bar{u}\|_H^2 = \min_{\bar{v}\in\bar{K}} \frac{1}{2}\|u - \bar{v}\|_H^2. \tag{4}$$

The aim of the VMS approach is to obtain a finite set of equations, independent of u, which will allow us to obtain a solution $\bar{u} \in \bar{K}$ to the above constrained optimization problem.

Before proceeding, let us define some terminology. Let $\mathcal{P}$ be the projection operator from V onto $\bar{V}$ defined by

$$(\mathcal{P}u, \bar{v})_H = (u, \bar{v})_H, \quad \forall \bar{v} \in \bar{V}. \tag{5}$$

We define $\tilde{V} = \mathrm{Ker}(\mathcal{P})$, which is also a closed subspace of V, and notice that

$$V = \bar{V} \oplus \tilde{V}. \tag{6}$$

Let $\mathcal{R}$ be the Riesz operator from $\bar{V}$ onto $\bar{V}^*$ defined by

$$ {}_{\bar{V}^*}\langle \mathcal{R}\bar{u}, \bar{v} \rangle_{\bar{V}} = (\bar{u}, \bar{v})_H, \quad \forall \bar{v} \in \bar{V}. \tag{7} $$

With the above terminology defined, we can define the Karush-Kuhn-Tucker conditions which uniquely specify the solution to (4).

Theorem 1. (Karush-Kuhn-Tucker Conditions) *The solution to the constrained optimization problem given by (4) is uniquely specified by the conditions*

$$ \mathcal{P}u' = \mathcal{P}(u - \bar{u}) = \mathcal{P}u - \bar{u} = \textstyle\sum_{i=1}^{n_{eq}} \lambda_i b_i + \sum_{j=1}^{n_{ineq}} \mu_j c_j(\bar{u}) \tag{8} $$

$$ f_i(\bar{u}) = 0, \quad \text{for } i = 1, 2, \ldots, n_{eq} \tag{9} $$

$$ g_j(\bar{u}) \leq 0, \quad \text{for } j = 1, 2, \ldots, n_{ineq} \tag{10} $$

$$ \mu_j \geq 0, \quad \text{for } j = 1, 2, \ldots, n_{ineq} \tag{11} $$

$$ \mu_j g_j(\bar{u}) = 0, \quad \text{for } j = 1, 2, \ldots, n_{ineq} \tag{12} $$

where $b_i \in \bar{V}$ is defined as

$$ b_i = \mathcal{R}^{-1} D f_i \tag{13} $$

and $c_j : \bar{V} \to \bar{V}$ is the nonlinear operator defined by

$$ c_j(\bar{v}) = \mathcal{R}^{-1} D g_j(\bar{v}), \quad \forall \bar{v} \in \bar{V}. \tag{14} $$

Using the Karush-Kuhn-Tucker conditions along with the variational projector $\mathcal{P}$, we can split the original variational problem in coarse-scale and fine-scale sub-problems, solve the fine-scale sub-problem for u', and then finally insert u' back into the coarse-scale sub-problem to arrive at a finite-dimensional problem for the desired numerical solution $\bar{u}$. This process is detailed in [8]. The resulting finite-dimensional problem takes the following form.

(†)

Find $\bar{u} \in \bar{V}, \lambda_i \in \mathbb{R}$ $(i = 1, \ldots, n_{eq})$, $\mu_j \in \mathbb{R}$ $(j = 1, \ldots, n_{ineq})$ such that

$$ {}_{V^*}\left\langle \mathcal{L}(\bar{u} + \hat{u}') - \mathcal{L}\mathcal{G}'\mathcal{L}(\bar{u} + \hat{u}'), \bar{v} \right\rangle_V = {}_{V^*}\left\langle \mathcal{F} - \mathcal{L}\mathcal{G}'\mathcal{F}, \bar{v} \right\rangle_V, \quad \forall \bar{v} \in \bar{V} \tag{15} $$

$$ f_i(\bar{u}) = 0, \quad \text{for } i = 1, 2, \ldots, n_{eq} \tag{16} $$

$$ g_j(\bar{u}) \leq 0, \quad \text{for } j = 1, 2, \ldots, n_{ineq} \tag{17} $$

$$ \mu_j \geq 0, \quad \text{for } j = 1, 2, \ldots, n_{ineq} \tag{18} $$

$$ \mu_j g_j(\bar{u}) = 0, \quad \text{for } j = 1, 2, \ldots, n_{ineq} \tag{19} $$

where $\hat{u}' \in \bar{V}$ is defined as

$$ \hat{u}' = \sum_{i=1}^{n_{eq}} \lambda_i b_i + \sum_{i=1}^{n_{ineq}} \mu_j c_j(\bar{u}), \tag{20} $$

and the fine-scale Green's operator $\mathcal{G}' : V^* \to \tilde{V}$ is defined by

$$ \mathcal{G}' = \mathcal{G} - \mathcal{G}\mathcal{P}^T \left(\mathcal{P}\mathcal{G}\mathcal{P}^T\right)^{-1} \mathcal{P}\mathcal{G}. \tag{21} $$

We have the following theorem characterizing the solution and Lagrange multipliers of the above variational formulation.

Theorem 2. *Problem* (†) *has a solution* $\bar{u} \in \bar{V}, \lambda_i \in \mathbb{R}$ $(i = 1, \ldots, n_{eq})$, $\mu_j \in \mathbb{R}$ $(j = 1, \ldots, n_{ineq})$. *Further,* $\bar{u}$ *is uniquely determined and is the unique solution to (4) with* $u = \mathcal{G}\mathcal{F}$, *and the Lagrange multipliers* λ_i, μ_j *are unique if* $\bar{u}$ *is a regular point of* $\bar{K}$.

The variational formulation (†) defines a class of VMS methods which allow for the direct enforcement of equality and inequality constraints. It consists of a finite set of equations for the coarse-scale solution $\bar{u}$ and Lagrange multipliers associated with the constraints. Theorem 2 guarantees that this formulation is well-posed under fairly modest conditions.

3 The advection-diffusion-reaction problem

In this section, we specialize the abstract framework presented in the previous section to the advection-diffusion-reaction equation. For simplicity, we consider only the case of periodic boundary conditions. A variational multiscale method is developed in which a total variation diminishing constraint is enforced.

3.1 Problem description: The periodic problem

Let $\Omega = (0,1)^d$ where d refers to the number of spatial dimensions. We consider the advection-diffusion-reaction problem with periodic boundary conditions. The strong formulation for the problem is as follows.

$$\text{(S)} \begin{cases} \text{Find } u : \overline{\Omega} \to \mathbb{R} \text{ such that } u \text{ is periodic and} \\ \qquad \sigma u - \kappa \Delta u + \beta \cdot \nabla u = f \text{ in } \Omega, \qquad (22) \\ \text{where } f : \Omega \to \mathbb{R} \text{ is the force term, } \sigma > 0 \text{ is the source parameter, } \kappa > 0 \text{ is the} \\ \text{scalar diffusivity, and } \beta : \overline{\Omega} \to \mathbb{R}^2 \text{ is the advection velocity, for which we assume} \\ \nabla \cdot \beta = 0. \end{cases}$$

Remark. We assume throughout the paper that the above equation is the result of a finite difference time discretization of an unsteady advection-diffusion equation. In this case, σ is proportional to Δt^{-1}, the inverse of the time-step, and f contains terms with history information (e.g., the solution at previous time-steps).

We assume that the specified data is sufficiently smooth such that (S) is well-defined. Define the Sobolev space

$$H^1_{per}(\Omega) = \left\{ u \in H^1(\Omega) : u \text{ is periodic } \right\} \qquad (23)$$

and its dual $H^{-1}_{per}(\Omega) = \left(H^1_{per}(\Omega)\right)^*$. The variational formulation for the periodic advection-diffusion-reaction equation is then the following.

$$
\text{(V)} \begin{cases} \text{Find } u \in H^1_{per}(\Omega) \text{ such that} \\ \qquad B(u,v) = F(v), \ \ \forall v \in H^1_{per}(\Omega) \qquad (24) \\ \text{where} \\ \qquad B(u,v) = (\sigma u, v)_{L^2(\Omega)} + (\kappa \nabla u - \beta u, \nabla v)_{L^2(\Omega)}, \qquad (25) \\ \qquad F(v) = {}_{H^{-1}_{per}(\Omega)}\langle f, v \rangle_{H^1_{per}(\Omega)} . \qquad (26) \end{cases}
$$

We see the above formulation fits within the framework of Section 2 with $V = H^1_{per}(\Omega)$, $V^* = H^{-1}_{per}(\Omega)$, $\mathcal{L} : H^1_{per}(\Omega) \to H^{-1}_{per}(\Omega)$ defined by

$$
{}_{H^{-1}_{per}(\Omega)}\langle \mathcal{L}u, v \rangle_{H^1_{per}(\Omega)} = B(u,v), \ \ \forall u, v \in H^1_{per}(\Omega), \qquad (27)
$$

or equivalently,

$$
\mathcal{L} = \sigma - \kappa\Delta + \beta \cdot \nabla, \qquad (28)
$$

and $\mathcal{F} \in H^{-1}_{per}(\Omega)$ defined by

$$
{}_{H^{-1}_{per}(\Omega)}\langle \mathcal{F}, v \rangle_{H^1_{per}(\Omega)} = F(v), \ \ \forall v \in H^1_0(\Omega). \qquad (29)
$$

The invertibility of $\mathcal{L}$ is due to the coercivity of $B(\cdot,\cdot)$. In this context, it is convenient to represent the Green's operator $\mathcal{G}$ through the Green's function $g : \Omega \times \Omega \to \mathbb{R}$ such that

$$
\mathcal{G}r(y) = \int_\Omega g(x,y) r(x) dx. \qquad (30)
$$

for all $r \in H^{-1}_{per}(\Omega)$. We have that g is periodic and $\mathcal{L}^* g(\cdot, y) = \delta(\cdot - y)$, where δ is the Dirac mass at the origin and $\mathcal{L}^*$ denotes the dual of $\mathcal{L}$. We immediately see that the solution of (V) is

$$
\begin{aligned} u(y) &= \int_\Omega g(x,y) \mathcal{F}(x) dx \\ &= \int_\Omega g(x,y) f(x) dx. \end{aligned} \qquad (31)
$$

Note that here and in what follows, integrals are to be understood in the sense of distributions.

3.2 The total variation constraint

We assume that the exact solution to (S) satisfies a total variation constraint of the form

$$\mathrm{TV}(u) \leq C \tag{32}$$

where

$$\mathrm{TV}(u) = \int_{\Omega} |\nabla u(x)| dx. \tag{33}$$

Note that when (S) is obtained through time discretization of the homogeneous unsteady advection-diffusion equation subject to a constant velocity field, we may take C to be the total variation of the solution at the preceding time-step. This choice reflects the total variation diminishing (TVD) property of the underlying time-dependent partial differential equation. When the velocity field is not constant, we may derive a modified total variation estimate (see, for example, the total variation estimate given by Theorem 2.3 of [1]). Inspired by the above discussion, we choose to impose an identical total variation constraint of the form

$$\mathrm{TV}(\bar{u}) \leq C \tag{34}$$

on the discrete numerical solution $\bar{u} \in \bar{V}$. We will refer to this constraint as the total variation bounded (TVB) property.

There has been a long history of ensuring satisfaction of a TVD or TVB property in the numerical solution of conservation laws in order to arrive at a monotone approximation. In his seminal paper [10], Harten proved that a one-dimensional explicit and conservative finite difference scheme is monotonicity preserving if and only if it is TVD. Harten derived a set of algebraic conditons, referred to as the TVD conditions, which guarantee that such a numerical scheme is TVD. In [10] and [11], Harten developed finite difference methods which control the total variation in a nonlinear way in order to prevent the creation of spurious extrema. These methods are inherently based on the notion of flux corrected transport (FCT), first introduced by Boris and Book in the early 1970s [2, 3]. The FCT and TVD schemes have since inspired an entire generation of high-resolution schemes, including the Monotone Upstream-centered Schemes for Conservation Laws (MUSCL) scheme [21], the Essentially Non-Oscillatory (ENO) and Weighted Essentially Non-Oscillatory (WENO) schemes [12, 19], and the Piecewise Parabolic Method (PPM) scheme [7]. Also of interest are the Finite Element Method Flux Corrected Transport (FEM-FCT) schemes which are based on the concept of algebraic flux correction [20]. The FEM-FCT schemes, however, are largely limited to linear finite element discretizations, though a recent extension has been made to quadratic elements [18].

3.3 The TVB VMS method

Let $\bar{V} \subset H^1_{per}(\Omega)$ be a finite-dimensional space. The space $\bar{V}$ is simultaneously the space of weighting functions and the set of trial solutions. Let us introduce the fine-scale Green's function $g' : \Omega \times \Omega \to \mathbb{R}$, which represents the fine-scale Green's operator

$\mathcal{G}'$ and satisfies

$$\mathcal{G}'r(y) = \int_\Omega g'(x,y)r(x)dx \tag{35}$$

for all $r \in H^{-1}_{per}(\Omega)$. Since the bilinear form $B(\cdot,\cdot)$ is coercive, the fine-scale Green's function is well-defined. The unconstrained VMS method can be stated as follows.

Unconstrained VMS method:

Find $\bar{u} \in \bar{V}$ such that

$$\begin{aligned}\int_\Omega \sigma\bar{u}(x)\bar{v}(x)dx + \int_\Omega (\kappa\nabla\bar{u}(x) - \beta(x)\bar{u}(x)) \cdot \nabla\bar{v}(x)dx - \\ \int_\Omega\int_\Omega \mathcal{L}\bar{u}(x)g'(x,y)\mathcal{L}^*\bar{v}(y)dydx \\ = \int_\Omega f(x)\bar{v}(x)dx - \int_\Omega\int_\Omega f(x)g'(x,y)\mathcal{L}^*\bar{v}(y)dydx\end{aligned} \tag{36}$$

for all $\bar{v} \in \bar{V}$.

Before defining the TVB variational multiscale method, let us first state the Karush-Kuhn-Tucker conditions associated with the TVB constrained optimization problem.

TVB Karush-Kuhn-Tucker Conditions:

$$(\bar{u} - \mathcal{P}u, \bar{v})_H + \mu\ {}_{\bar{V}^*}\langle D\mathrm{TV}(\bar{u}), \bar{v}\rangle_{\bar{V}} = 0 \tag{37}$$

$$\mathrm{TV}(\bar{u}) \leq C \tag{38}$$

$$\mu \geq 0 \tag{39}$$

$$\mu\,(\mathrm{TV}(\bar{u}) - C) = 0 \tag{40}$$

Note that given $\mathcal{P}u$, the solution to the TVB constrained optimization problem can be obtained uniquely from the above four conditions. Since $\mathcal{P}u$ is given by the unconstrained VMS method, one can utilize the Karush-Kuhn-Tucker conditions to post-process the unconstrained solution to arrive at the desired constrained solution. This is the procedure that is advocated in [8]. Alternatively, one can directly utilize the full TVB VMS formulation. While it is not advisable to use this form in practice, the full formulation can be utilized to understand constrained VMS analysis at a theoretical level. For example, the formulation will be utilized in Section 4 to understand the mechanisms responsible for discontinuity capturing in the context of a Fourier basis.

The TVB VMS formulation can be stated as follows.

TVB VMS method:

Find $\bar{u} \in \bar{V}, \mu \in \mathbb{R}$ such that

$$\begin{aligned}\int_\Omega \sigma \bar{u}(x)\bar{v}(x)dx + \int_\Omega (\kappa \nabla \bar{u}(x) - \beta(x)\bar{u}(x)) \cdot \nabla \bar{v}(x)dx - \\ \int_\Omega \int_\Omega \mathcal{L}\bar{u}(x) g'(x,y) \mathcal{L}^* \bar{v}(y) dy dx + \\ \int_\Omega \sigma \hat{u}'(x)\bar{v}(x)dx + \int_\Omega \left(\kappa \nabla \hat{u}'(x) - \beta(x)\hat{u}'(x)\right) \cdot \nabla \bar{v}(x)dx - \\ \int_\Omega \int_\Omega \mathcal{L}\hat{u}'(x) g'(x,y) \mathcal{L}^* \bar{v}(y) dy dx \\ = \int_\Omega f(x)\bar{v}(x)dx - \int_\Omega \int_\Omega f(x) g'(x,y) \mathcal{L}^* \bar{v}(y) dy dx \end{aligned} \tag{41}$$

for all $\bar{v} \in \bar{V}$,

$$\mathrm{TV}(\bar{u}) \leq C, \tag{42}$$

$$\mu \geq 0, \tag{43}$$

$$\mu \left(\mathrm{TV}(\bar{u}) - C\right) = 0, \tag{44}$$

and

$$\hat{u}'(x) = \mu B(\bar{u}) \tag{45}$$

where $B : \bar{V} \to \bar{V}$ satisfies

$$(B(\bar{w}), \bar{v})_H = {}_{\bar{V}^*}\langle D\mathrm{TV}(\bar{w}), \bar{v}\rangle_{\bar{V}} = \int_\Omega \frac{\nabla \bar{w} \cdot \nabla \bar{v}}{|\nabla \bar{w}|} dx, \ \ \forall \bar{w}, \bar{v} \in \bar{V}. \tag{46}$$

Let us state a few properties of the TVB VMS method for the particular case when the projector $\mathcal{P}$ is induced by the L^2-norm. This choice of projector can be interpreted as inspired by the method of *orthogonal subscales* of Codina [5]. For this case, the Karush-Kuhn-Tucker conditions for our constrained optimization problem imply that

$$\int_\Omega \bar{u}\bar{v}dx + \mu \int_\Omega \frac{\nabla \bar{u} \cdot \nabla \bar{v}}{|\nabla \bar{u}|} dx = \int_\Omega \mathcal{P}u\bar{v}dx \tag{47}$$

for every $\bar{v} \in \bar{V}$. Inserting $\bar{v} \equiv 1$ into the above equation, we obtain

$$\int_\Omega \bar{u}dx = \int_\Omega \mathcal{P}udx. \tag{48}$$

Hence, our method is conservative with respect to the unconstrained VMS method. Plugging $\bar{v} \equiv \bar{u}$ into (47), we obtain, with a simple application of the Cauchy-Schwarz inequality,

$$||\bar{u}||_{L^2(\Omega)} + \mu \frac{||\nabla \bar{u}||_{L^1(\Omega)}}{||\bar{u}||_{L^2(\Omega)}} \leq ||\mathcal{P}u||_{L^2(\Omega)}. \tag{49}$$

Since $\mu \geq 0$, this indicates that the TVB VMS method results in a numerical solution with less energy than the solution of the unconstrained VMS method. Thus, as expected, the TVB VMS is more dissipative than the unconstrained VMS method.

Remark The TVD constraint is not continuously differentiable in the sense of Fréchet. This presents problems both theoretically and numerically. In practice, the TVD constraint must be replaced with a regularized convex constraint. One possible candidate is

$$\int_\Omega \frac{|\nabla \bar{u}|^2}{|\nabla \bar{u}| + \epsilon} dx + 3\epsilon \leq C \tag{50}$$

where $0 < \epsilon \ll 1$. Since

$$\begin{aligned}
\mathrm{TV}(\bar{u}) &= \int_\Omega \frac{|\nabla \bar{u} + \epsilon|^2}{|\nabla \bar{u}| + \epsilon} dx \\
&\leq \int_\Omega \frac{|\nabla \bar{u}|^2 + 2\epsilon\left(|\nabla \bar{u}| + \epsilon\right) + \epsilon^2}{|\nabla \bar{u}| + \epsilon} dx \\
&= \int_\Omega \frac{|\nabla \bar{u}|^2}{|\nabla \bar{u}| + \epsilon} dx + \int_\Omega \left(2\epsilon + \frac{\epsilon^2}{|\nabla \bar{u}| + \epsilon}\right) dx \\
&\leq \int_\Omega \frac{|\nabla \bar{u}|^2}{|\nabla \bar{u}| + \epsilon} dx + \int_\Omega 3\epsilon dx \\
&= \int_\Omega \frac{|\nabla \bar{u}|^2}{|\nabla \bar{u}| + \epsilon} + 3\epsilon,
\end{aligned} \tag{51}$$

it follows that the original TVD constraint is satisfied if (50) is satisfied. Furthermore, the conservation and dissipation properties that were discussed above are not upset if the TVD constraint is replaced with the regularized convex constraint.

3.4 Application of TVB VMS to the Repair of SUPG

In practice, one must approximate the effect of the fine-scale Green's function appearing in VMS formulation. In the context of finite elements, a simple choice is to approximate the fine-scale Green's function locally by a constant multiple of the Dirac mass on each element, i.e.,

$$\int_\Omega \int_\Omega r_1(x) g'(x,y) r_2(y) dy dx \approx -\sum_{e=1}^{n_{el}} \tau_e \int_{\Omega_e} r_1(x) r_2(x) dx \tag{52}$$

for all $r_1, r_2 \in \bar{V}^*$ where τ_e is an element-wise constant. For linear finite elements and bilinear finite elements on a rectangular grid, this results in the Streamline-Upwind/Petrov-Galerkin (SUPG) method [4]. This is the most common method in practice for the solution of advective phenomena by the finite elements. An even stronger link between SUPG and VMS can be established using the notion of residual-free bubbles [9]. Like the unconstrained VMS method, the SUPG method suffers from

spurious oscillations near discontinuities and layers and must be supplemented by additional mechanisms to remove these oscillations.

We present two very simple numerical examples to demonstrate how the TVB VMS framework can be used to repair the SUPG method in the presence of sharp layers. Our first example involves simulating a traveling square wave satisfying the one-dimensional homogeneous advection-diffusion equation. As mentioned previously, the solution to this equation satisfies a strong TVD constraint. We discretize the unsteady advection-diffusion equation in time using the midpoint method. Then, for each time-step, we have to solve the resulting advection-diffusion-reaction equation with $u = u_{n+1}$, $\sigma = 2\Delta t^{-1}$, and $f = -2\Delta t^{-1} u_n + \kappa \Delta u_n - \beta \cdot \nabla u_n$. We then utilize SUPG to compute an approximation to the unconstrained VMS solution $\mathcal{P}u$ at each time-step. We finally post-process using the Karush-Kuhn-Tucker conditions to find the L^2 best fit to the SUPG solution satisfying a TVD constraint. For the results presented here, fine-scale quantities are not tracked in time. This quasi-static assumption can lead to unstable results for small time-steps [14]. To remedy this issue, explicit tracking can be easily accommodated [5, 6]. To compute the solution satisfying the exact TVD constraint, we compute the solutions satisfying the regularized constraint given by (50) for a sequence of $\epsilon \to 0$.

In Figure 1, we compare the unconstrained SUPG solution with the TVD SUPG solution for the choice of $\beta = 1$, $\kappa = 10^{-6}$, and initial condition

$$u_0(x) = \begin{cases} \frac{x}{\delta} & \text{if} \quad 0 \leq x \leq \delta \\ 1 & \text{if} \quad \delta \leq x \leq 0.3 - \delta \\ \frac{0.3-x}{\delta} & \text{if} \quad 0.3 - \delta \leq x \leq 0.3 \\ 0 & \text{otherwise} \end{cases} \tag{53}$$

where δ is chosen to be 10^{-3}. The solutions were obtained using piecewise linear basis functions on a 50 element mesh with a SUPG parameter equal to $\tau = h/2$. The unconstrained SUPG solution exhibits spurious oscillations and suffers from particularly large overshoots on the right side of the traveling square wave. The TVD SUPG solution, on the other hand, is monotone. Furthermore, the layers of the SUPG solution have not been smeared through the enforcement of the TVD constraint.

Our second numerical example involves the simulation of a two-dimensional square wave. Specifically, we solve the two-dimensional homogeneous advection-diffusion equation with $\beta = \left[\sqrt{2}/2, \sqrt{2}/2\right]$, $\kappa = 10^{-6}$, and

$$u_0(x) = u_0(x_1, x_2) = z(x_1) z(x_2) \tag{54}$$

where

$$z(x) = \begin{cases} \frac{x-0.1}{\delta} & \text{if} \quad 0.1 \leq x \leq 0.1 + \delta \\ 1 & \text{if} \quad 0.1 + \delta \leq x \leq 0.4 - \delta \\ \frac{0.4-x}{\delta} & \text{if} \quad 0.4 - \epsilon \leq x \leq 0.4 \\ 0 & \text{otherwise} \end{cases} \tag{55}$$

and δ is again chosen to be 10^{-3}. We discretize in time again using the midpoint method and use SUPG to approximate the effect of the fine-scale Green's function.

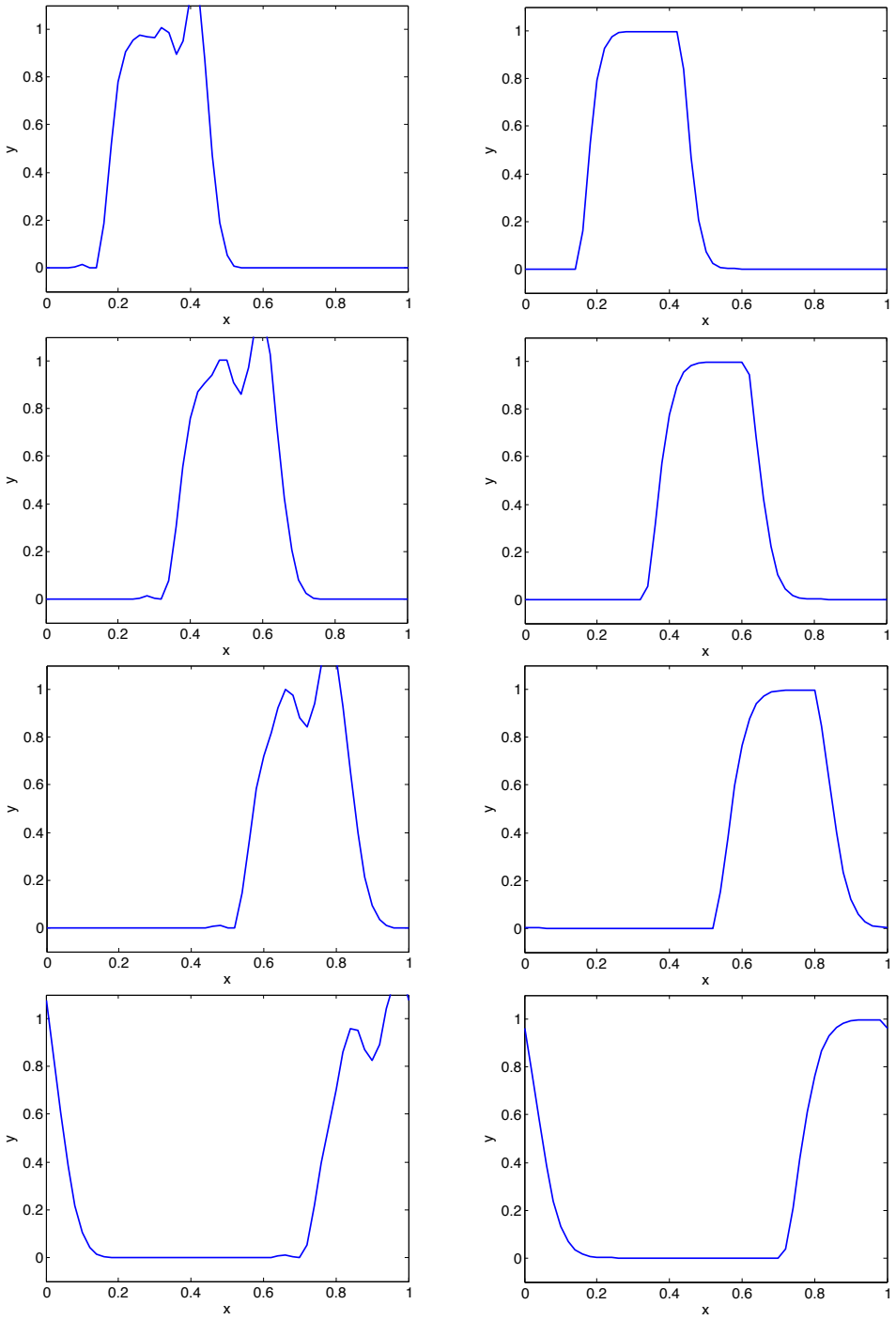

Figure 1: Snapshots of the unconstrained SUPG solution (left) and the constrained SUPG solution (right) at time-steps $10, 20, 30, 40$ (top to bottom) for the first numerical example.

In Figure 2, we compare the unconstrained SUPG solution with the TVD SUPG solution using a piecewise bilinear mesh of 50 x 50 elements and a time-step of $\Delta t = 1/50$. The SUPG parameter was chosen to be $\tau = h/2$. The unconstrained SUPG solution suffers from sharp undershoots and overshoots, especially near the downwind region of the convected square region, while the TVD SUPG solution is much better behaved. Note though that the TVD SUPG solution still suffers from a strong overshoot near the downwind corner of the convected square. This illustrates that enforcement of the TVD condition is not enough to ensure a monotone solution. One can, however, obtain a monotone solution by replacing the exact TVD constraint with a stronger total variation constraint. For example, one may choose to utilize the regularized constraint given by (50) instead with some fixed ϵ. In Figure 3, we plot such an over-constrained SUPG solution for a choice of $\epsilon = 0.01$. We see that the over-constrained solution is monotone as opposed to the TVD SUPG solution. A downfall of the over-constrained methodology is that the proper choice of ϵ for general problems is not obvious.

When β is not a constant velocity field, the exact solution to the homogeneous unsteady advection-diffusion equation no longer satisfies a TVD property. As mentioned before in Section 3.2, other total variation bounds may be derived, but these estimates are very conservative and of not much use in practice. The situation is similar for forced transport or non-periodic boundary conditions. For these cases, it may be possible to instead model the total variation bound C to arrive at a monotone numerical method. Alternatively, one may let go of the total variation constraint entirely by modeling the Lagrange multiplier μ instead of the total variation bound C, leading to a regularization formulation. Such modeling, however, is beyond the scope and purpose of this paper.

4 Specialization to a Fourier basis

To better understand the discontinuity capturing mechanisms in the TVB VMS method, we examine the method in the context of a Fourier spectral basis. Let $d = 2$ for illustrative purposes and let $\beta(x) = \beta$ be constant. We consider the case when $\bar{V}$ consists of Fourier modes. That is, we consider the space

$$\bar{V} = \left\{ \bar{v} \in \mathcal{V} : \bar{v}(x) = \bar{v}(x_1, x_2) = \sum_{\substack{k_1,k_2 \in \mathbb{I} \\ k_1^2 + k_2^2 \leq k_c^2}} \bar{v}_{k_1,k_2} e^{i2k_1\pi x_1} e^{i2k_2\pi x_2}, \;\; \bar{v}_{k_1,k_2} \in \mathbb{C}, \right\} \quad (56)$$

where k_c is a prescribed cut-off frequency. To define our coarse-scale approximation, we utilize the projector induced by the L^2-norm. In this case, it is easily shown that $\tilde{V} = \mathrm{Ker}(\mathcal{P})$ consists of Fourier modes that are above the cut-off frequency.

We now derive a simplified expression for the TVB VMS method. Let us start with

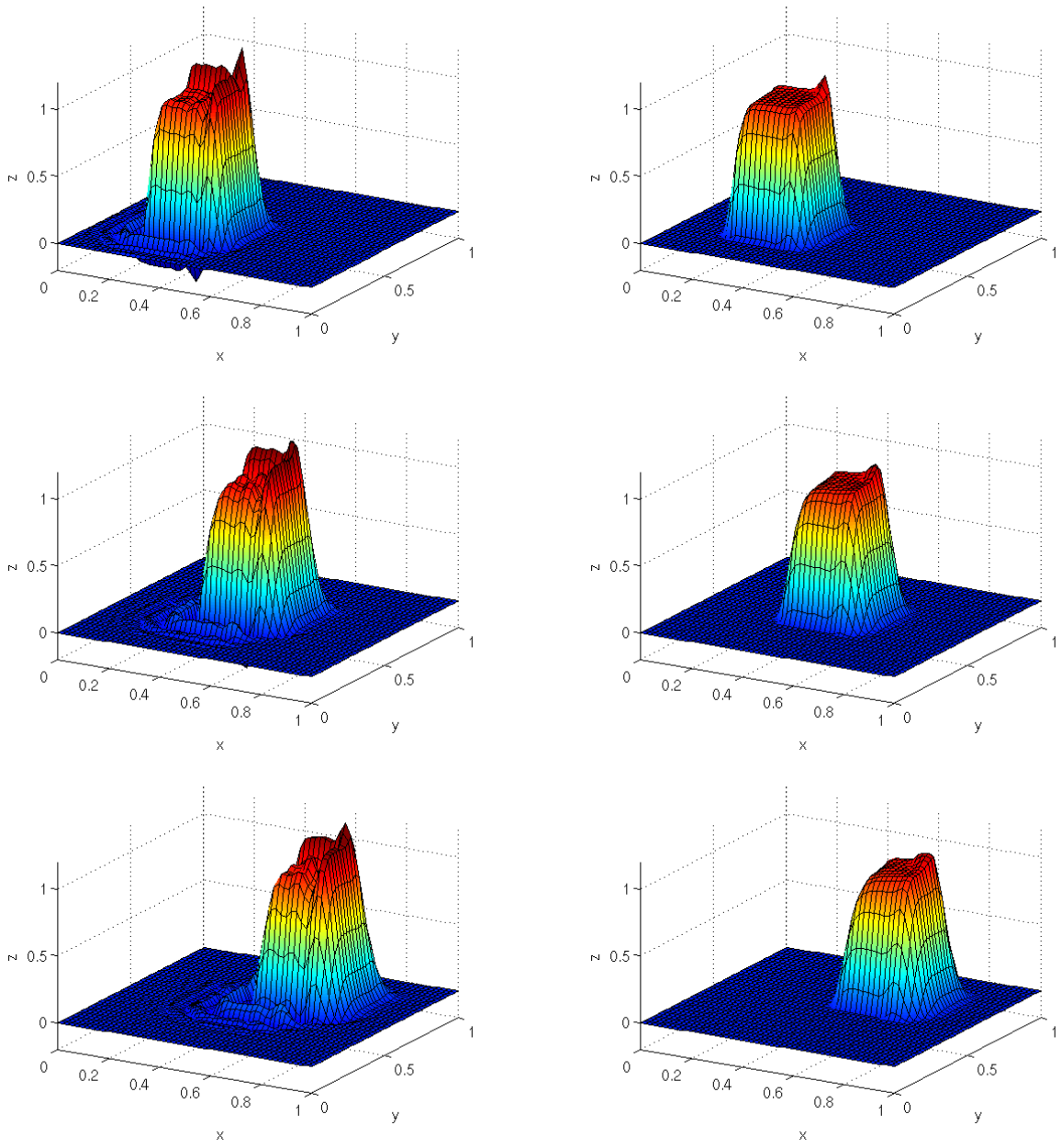

Figure 2: Snapshots of the unconstrained VMS solution (left) and the constrained VMS solution (right) at time-steps $10, 20, 30$ (top to bottom) for the second numerical example. For the constrained VMS solution, the exact TVD constraint is enforced.

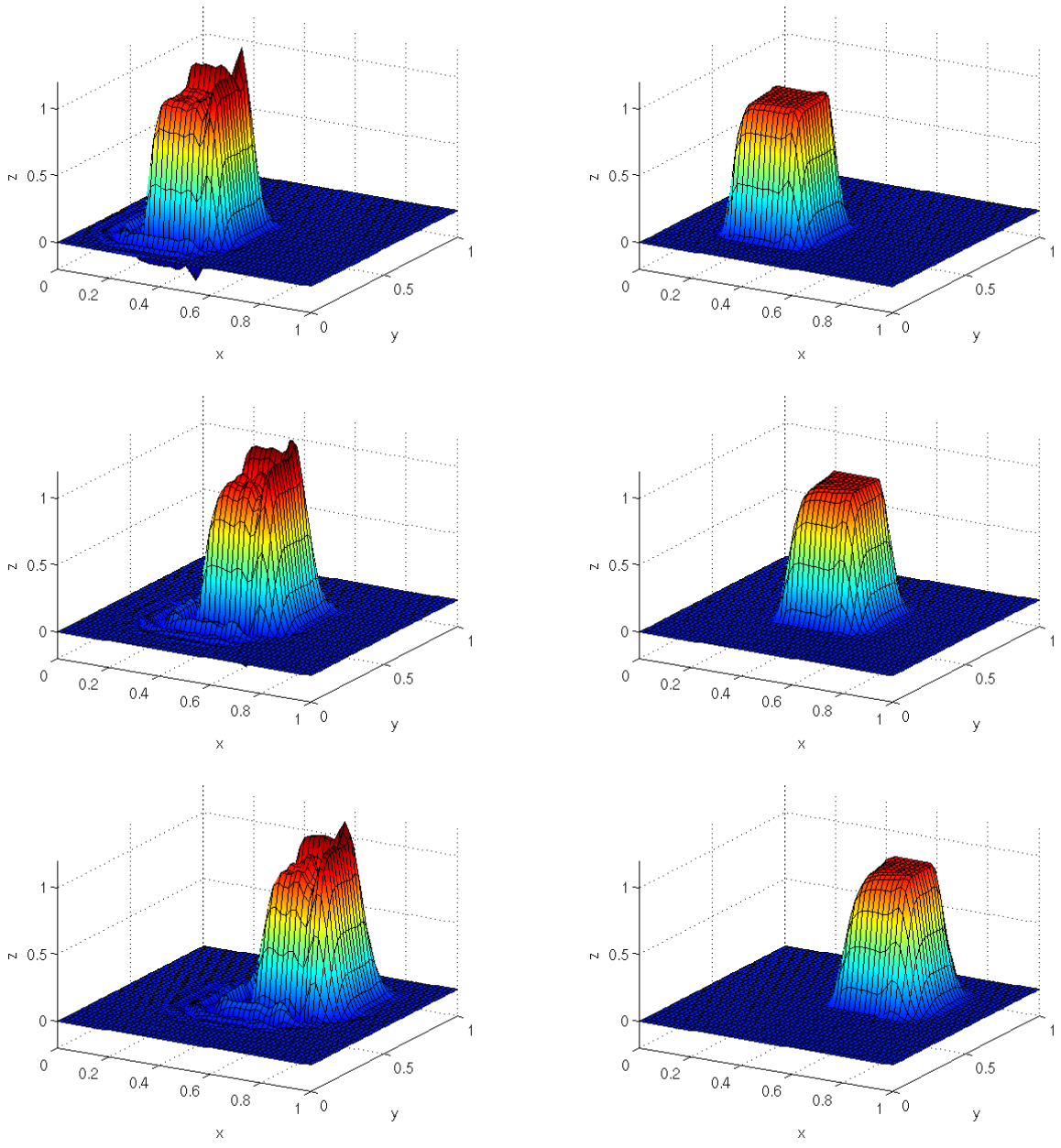

Figure 3: Snapshots of the unconstrained VMS solution (left) and the over-constrained VMS solution (right) at time-steps $10, 20, 30$ (top to bottom) for the second numerical example.

the advective terms. To begin, we note that, through integration by parts,

$$\begin{aligned} -\int_\Omega \left(\beta\cdot\nabla\bar{v}\right)\left(\bar{u}+\hat{u}'+\tilde{u}'\right)dx &= \int_\Omega \beta\cdot\nabla\bar{u}\bar{v}dx \\ &\quad -\int_\Omega \left(\beta\cdot\nabla\bar{v}\right)\left(\hat{u}'+\tilde{u}'\right)dx \end{aligned} \tag{57}$$

for all $\bar{v}\in\bar{V}_n$. Since β is constant, we have that $\beta\cdot\nabla\bar{v}\in\bar{V}_n$ and hence

$$\int_\Omega \left(\beta\cdot\nabla\bar{v}\right)\tilde{u}'dx = 0. \tag{58}$$

Further, by definition, we also have

$$\begin{aligned} \int_\Omega \left(\beta\cdot\nabla\bar{v}\right)\hat{u}'dx &= \int_\Omega \left(\beta\cdot\nabla\bar{v}\right)\mu B(\bar{u})dx \\ &= \mu\int_\Omega \frac{\nabla\bar{u}\cdot\nabla\left(\beta\cdot\nabla\bar{v}\right)}{|\nabla\bar{u}|}dx. \end{aligned} \tag{59}$$

Integrating by parts twice, we obtain

$$\int_\Omega \left(\beta\cdot\nabla\bar{v}\right)\hat{u}'dx = \mu\int_\Omega \nabla\cdot\left(\beta\nabla\cdot\left(\frac{\nabla\bar{u}}{|\nabla\bar{u}|}\right)\right)\bar{v}dx. \tag{60}$$

We now proceed onto the reaction terms. To begin, note that

$$\int_\Omega \sigma\bar{v}\tilde{u}'dx = 0. \tag{61}$$

Again by definition, we have that

$$\begin{aligned} \int_\Omega \sigma\bar{v}\hat{u}'dx &= \int_\Omega \sigma\bar{v}\mu B(\bar{u})dx \\ &= \mu\int_\Omega \frac{\nabla\bar{u}\cdot\nabla\left(\sigma\bar{v}\right)}{|\nabla\bar{u}|}dx. \end{aligned} \tag{62}$$

Integrating by parts, we obtain

$$\int_\Omega \sigma\bar{v}\hat{u}'dx = -\mu\int_\Omega \nabla\cdot\left(\sigma\left(\frac{\nabla\bar{u}}{|\nabla\bar{u}|}\right)\right)\bar{v}dx. \tag{63}$$

Finally, we deal with the diffusive terms. Since $\Delta\bar{v}\in\bar{V}$, we have that

$$\int_\Omega \kappa\nabla\bar{v}\cdot\nabla\tilde{u}'dx = -\int_\Omega \kappa\Delta\bar{v}\tilde{u}' = 0. \tag{64}$$

By definition, we have that

$$\begin{aligned} \int_\Omega \kappa\nabla\bar{v}\cdot\nabla\hat{u}'dx &= -\int_\Omega \kappa\Delta\bar{v}\mu B(\bar{u})dx \\ &= -\mu\int_\Omega \frac{\nabla\bar{u}\cdot\nabla\left(\kappa\Delta\bar{v}\right)}{|\nabla\bar{u}|}dx. \end{aligned} \tag{65}$$

Integrating by parts three times, we obtain

$$\int_\Omega \kappa \nabla \bar{v} \cdot \nabla \hat{u}' dx = \mu \int_\Omega \nabla \cdot \left(\kappa \nabla \left(\nabla \cdot \left(\frac{\nabla \bar{u}}{|\nabla \bar{u}|} \right) \right) \right) \bar{v} dx. \tag{66}$$

Combining (57), (58), (60), (61), (63), (64), and (66) with the formulation given by (41), we arrive at the following representation for the TVB VMS method.

TVB VMS method for a Fourier basis:

Find $\bar{u} \in \bar{V}$ and $\mu \in \mathbb{R}$ such that

$$\int_\Omega \left[\sigma \bar{u} + \nabla \cdot \bar{q}(\bar{u}) \right] \bar{v} dx = \int_\Omega f \bar{v}, \text{ for all } \bar{v} \in \bar{V}_n, \tag{67}$$

$$\int_\Omega |\nabla \bar{u}| dx \leq C, \tag{68}$$

$$\mu \geq 0, \tag{69}$$

$$\mu \int_\Omega \left(|\nabla \bar{u}| - C| \right) dx = 0, \tag{70}$$

where $\bar{q}$ is the flux function defined by

$$\bar{q}(\bar{u}) = \beta \left(\bar{u} - \mu \nabla \cdot \left(\frac{\nabla \bar{u}}{|\nabla \bar{u}|} \right) \right) - \sigma \mu \frac{\nabla \bar{u}}{|\nabla \bar{u}|} - \kappa \left(\nabla \bar{u} - \mu \nabla \left(\nabla \cdot \left(\frac{\nabla \bar{u}}{|\nabla \bar{u}|} \right) \right) \right). \tag{71}$$

Examining the above equations, we find the TVB VMS method is a *method of flux correction*. In contrast with methods based on modification of discrete algebraic fluxes, the TVB VMS method is one of *variational flux correction* - the *functional* flux has been modified *under* the integral sign. The standard flux $q = \beta \bar{u} - \kappa \nabla \bar{u}$ has been replaced with a modified flux $\bar{q}$ consisting of additional dispersive, diffusive, and hyper-diffusive terms. Specifically, a third-order *artificial transport* term of the form

$$-\beta \cdot \nabla \left(\mu \nabla \cdot \left(\frac{\nabla \bar{u}}{|\nabla \bar{u}|} \right) \right), \tag{72}$$

an *artificial diffusion* term of the form

$$-\mu \nabla \cdot \left(\sigma \left(\frac{\nabla \bar{u}}{|\nabla \bar{u}|} \right) \right), \tag{73}$$

and an *artificial hyper-diffusion* term of the form

$$\mu \nabla \cdot \left(\nabla \left(\kappa \nabla \cdot \left(\frac{\nabla \bar{u}}{|\nabla \bar{u}|} \right) \right) \right) \tag{74}$$

have been added to the residual of the unsteady advection-diffusion equation. Hence, from the standpoint of variational multiscale analysis, flux correction is the *natural* mechanism for obtaining the *optimal* solution satisfying a total variation constraint.

In the presence of vanishing reaction and diffusivity, the artificial transport term alone is responsible for the enforcement of the total variation constraint. This seems to be in direct contrast with classical discontinuity-capturing techniques which are based solely on the introduction of artificial diffusion. In fact, the inclusion of a third-order operator seems completely counterintuitive as excessive dispersion is the cause of unwanted oscillations in higher-order numerical schemes. However, as was shown in Section 3, the TVB VMS method *is* dissipative independent of the size of the source parameter and diffusivity terms, and a further examination of the artificial transport term in a simplified setting reveals just how it removes spurious oscillations.

Consider the one-dimensional model flux term

$$r(v) = b\left(v - \lambda \frac{d}{dx}\left(\frac{dv/dx}{|dv/dx|}\right)\right) = b\left(v - \lambda \frac{d}{dx}\left(\operatorname{sgn}\left(\frac{du}{dx}\right)\right)\right) \tag{75}$$

where $\lambda \geq 0$ and sgn is the signum operator. Note that r is the one-dimensional analogue of $\bar{q}$ in the absence of reaction and diffusion. Through an application of the chain rule, we can easily re-express r as

$$r(v) = b\left(v - 2\lambda\delta\left(\frac{du}{dx}\right)\frac{d^2u}{dx^2}\right) \tag{76}$$

where δ is the Dirac mass at the origin.

Equation (76) provides physical intuition as to how the artificial transport term removes spurious oscillations. Away from extrema, the flux term r is equivalent to bv, the standard convective flux. Alternatively, at extrema, the flux term r is modified. At maxima, the second derivative is negative, and hence the standard flux is *increased* by a non-negative multiple of $\delta(0)$, to be understood in the sense of distributions. Since overshoots are associated with maxima, we see that transport is increased near potential overshoots. As we may interpret overshoots as "accumulation of too much material" (with v being a measure of material quantity), this increased transport has the effect of moving excess material away from maxima and hence preventing any potential material accumulation that may lead to the formation of overshoots. On the other hand, the second derivative is positive at minima, and the standard flux is instead *decreased* by a non-negative multiple of $\delta(0)$. This indicates that transport is now decreased near potential undershoots. This decreased transport has the effect of retaining material near minima and hence preventing any spurious undershoots from forming. A visual overview of the behavior of the flux is provided in Figure 4. As a final note, observe that the flux is unmodified at extrema when $\lambda = 0$.

We feel that variational flux correction is an exciting new approach to discontinuity capturing which naturally extends the notion of flux corrected transport to variationally-based numerical schemes such as the finite element method and the spectral method. It is compelling that such a discontinuity capturing methodology has a firm theoretical rooting in variational multiscale analysis and constrained optimization. However, more research must be done to determine how to extend the procedures described here to arrive at effective variational flux correction schemes for general transport problems.

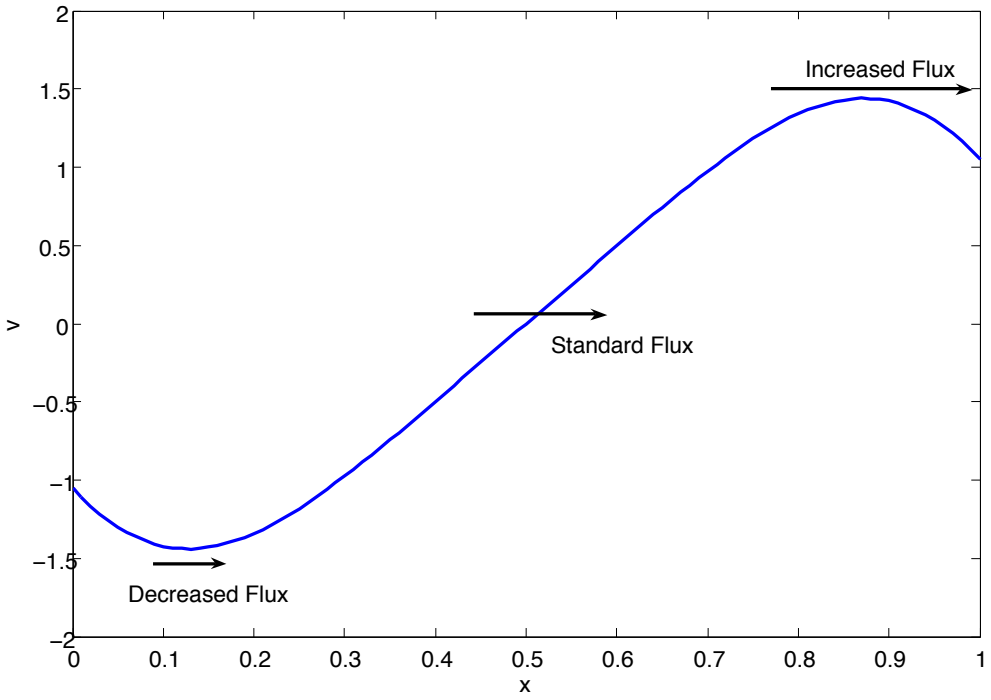

Figure 4: An illustration of the behavior of the flux function $r(v)$ at and away from extrema for $v(x) = \sin(5x - 2.5)\left(\tanh\left((2x-1)^2\right) - 1\right)$ and $b > 0$.

Remark: We would also like to point out the variational multiscale method presented here also indicates an additional link between flux limiters and *slope limiters* which instead act directly on *system states*. Notably, in the context of a Fourier basis, one can obtain the optimal TVB solution via either post-processing the unconstrained optimal solution (i.e., slope limiting) or the full TVB VMS formulation with modified flux (i.e., flux limiting). For more on slope limiting versus flux limiting, see [13].

5 Conclusions

In this paper, we presented a new variational multiscale method that arises through the enforcement of a total variation constraint. We first presented the method in full generality with application to the repair of stabilized methods. Numerical examples illustrated the effectiveness of the repair procedure. We then specialized to the case of a Fourier basis and showed the resulting method was one of flux correction, revealing a link between flux correction, total variation, and constrained optimization. We finally explored the nature of the flux correction terms in one spatial dimension.

Acknowledgements

J.A. Evans was partially supported by the Department of Energy Computational Science Graduate Fellowship, provided under grant number DE-FG02-97ER25308. T.J.R. Hughes was partially supported by the Office of Naval Research under Contract Number N00014-08-0992, by the Boeing Company under Contract Number 20010, and by the Department of Energy [National Nuclear Security Administration] under Award Number DE-FC52-08NA28615. This support is gratefully acknowledged. The authors would like to additionally thank J. Liu for reviewing a draft of this paper.

References

[1] P. Amorim, M. Ben-Artzi, and P.G. LeFloch. Hyperbolic conservation laws on manifolds: Total variation estimates and the finite volume method. *Methods and Applications of Analysis*, 12:291–324, 2005.

[2] J.P. Boris. A fluid transport algorithm that works. In *Computing as a Language of Physics*, pages 171–189. International Atomic Energy Commission, 1971.

[3] J.P. Boris and D.L. Book. Flux-corrected transport I: SHASTA, a fluid-transport algorithm that works. *Journal of Computational Physics*, 11:38–69, 1973.

[4] A.N. Brooks and T.J.R. Hughes. Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 32:199–259, 1982.

[5] R. Codina. Stabilized finite element approximation of transient incompressible flows using orthogonal subscales. *Computer Methods in Applied Mechanics and Engineering*, 191:4295–4321, 2000.

[6] R. Codina, J. Principe, O. Guasch, and S. Badia. Time dependent subscales in the stabilized finite element approximation of incompressible flow problems. *Computer Methods in Applied Mechanics and Engineering*, 196:2413–2430, 2007.

[7] P. Colella and P.R. Woodward. The piecewise parabolic method (PPM) for gas-dynamical simulations. *Journal of Computational Physics*, 54:174–201, 1984.

[8] J.A. Evans, T.J.R. Hughes, and G. Sangalli. Enforcement of constraints and maximum principles in the variational multiscale method. *Computer Methods in Applied Mechanics and Engineering*, 199:61–76, 2009.

[9] L.P. Franca, S.L. Frey, and T.J.R. Hughes. Stabilized finite element methods: I. Application to the advective-diffusive model. *Computer Methods in Applied Mechanics and Engineering*, 95:253–276, 1992.

[10] A. Harten. High resolution schemes for hyperbolic conservation laws. *Journal of Computational Physics*, 49:357–393, 1983.

[11] A. Harten. On a class of high resolution total-variation-stable finite-difference schemes. *SIAM Journal of Numerical Analysis*, 21:1–23, 1984.

[12] A. Harten, B. Engquist, S. Osher, and S.R. Chakravarthy. Uniformly higher order accurate essentially non-oscillatory schemes III. *Journal of Computational Physics*, 71:231–303, 1987.

[13] C. Hirsch. *Numerical Computation of Internal and External Flows, Volume 2, Computational Methods for Inviscid and Viscous Flows*. Wiley, New York, NY, 1990.

[14] M.C. Hsu, Y. Bazilevs, V.M. Calo, T.E. Tezduyar, , and T.J.R. Hughes. Improving stability of stabilized and multiscale formulations in flow simulations at small time steps. *Computer Methods in Applied Mechanics and Engineering*, 2009. In review.

[15] T.J.R. Hughes. Multiscale phenomena: Greens functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Computer Methods in Applied Mechanics and Engineering*, 127:387–401, 1995.

[16] T.J.R. Hughes, G.R. Feijoo, L. Mazzei, and J.-B. Quincy. The variational multiscale method - a paradigm for computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 166:3–24, 1998.

[17] T.J.R. Hughes and G. Sangalli. Variational multiscale analysis: the fine-scale Green's function, projection, optimization, localization, and stabilized methods. *SIAM Journal on Numerical Analysis*, 45:539–557, 2007.

[18] D. Kuzmin. On the design of algebraic flux correction schemes for quadratic finite elements. *Journal of Computational and Applied Mathematics*, 218:79–87, 2008.

[19] X.-D. Liu, S. Osher, and T. Chan. Weighted essentially nonoscillatory schemes. *Journal of Computational Physics*, 115:200–212, 1995.

[20] R. Löhner, K. Morgan, M. Vahdati, J.P. Boris, and D.L. Book. FEM-FCT: Combining unstructured grids with high resolution. *Communications in Applied Numerical Methods*, 7:717–729, 1988.

[21] B. van Leer. Towards the ultimate conservative difference scheme, V: A second order sequel to Godunov's method. *Journal of Computational Physics*, 32:101–136, 1979.