

# Minitutorial on A Posteriori Error Estimation

Donald Estep

University Interdisciplinary Research Scholar  
Editor in Chief, SIAM/ASA Journal on Uncertainty Quantification  
SIAM Fellow  
Chalmers Jubilee Professor  
Department of Statistics  
Colorado State University

Research supported by Defense Threat Reduction Agency (HDTRA1-09-1-0036),  
Department of Energy (DE-FG02-04ER25620, DE-FG02-05ER25699,  
DE-FC02-07ER54909, DE-SC0001724, DE-SC0005304, INL00120133,  
DE00000000SC9279), Engility, Inc. (BY13-050SP), Idaho National Laboratory  
(00069249, 00115474), Lawrence Livermore National Laboratory (B573139,  
B584647, B590495), National Science Foundation (DMS-0107832,  
DMS-0715135, DGE-0221595003, MSPA-CSE-0434354, ECCS-0700559,  
DMS-1065046, DMS-1016268, DMS-FRG-1065046, DMS-1228206), National  
Institutes of Health (#R01GM096192), Sandia Corporation (PO299784)

# Collaborators and Colleagues

This minitutorial describes developments in a posteriori error analysis over a period of more than twenty years

The approach is built on a foundation of functional analysis that has a broader application than just numerical analysis

My own contributions result from collaborations with a number of senior people, students, and postdocs

Significant contributions have been made by many other colleagues

I give a few references at the end and these contain further references

# A Priori and A Posteriori Analysis

Two general approaches to analyzing numerical methods:

**a priori** The goal is to derive the general convergence and accuracy properties of a numerical method for a wide class of solutions

**a posteriori** The goal is to derive a computational estimate that evaluates the accuracy of information computed from a particular numerical solution

A priori analysis is the “classic” error analysis that dominates the research literature

A priori analysis is generally useless for determining the accuracy of a particular computation

A posteriori analysis techniques yield accurate error estimates but do not describe general approximation properties well

*Functionals, Dual Spaces, and Adjoints*

*Tools for Investigating Differential Equations*

(32)

# Banach Spaces

We investigate properties of operators (maps) between spaces

We often assume the spaces in question are Banach spaces

A Banach space is a normed vector space such that Cauchy sequences converge to a limit in the space

A sequence  $\{x_n\}$  in a space  $X$  is a Cauchy sequence if for every  $\epsilon > 0$  there is an  $N$  such that  $\|x_n - x_m\| < \epsilon$  for all  $n, m > N$

This is a computable criterion for checking convergence

# What Information is to be Computed?

The starting point is the **computation of particular information**, or a **quantity of interest**, obtained from a numerical solution of a model

Considering a particular quantity of interest is important because **obtaining solutions that are accurate everywhere is often impossible**

The application should begin by answering

*What do we want to compute from the model?*

# Definition of Linear Functionals

Let  $X$  be a vector space with norm  $\| \cdot \|$

A bounded linear functional  $\ell$  is a continuous linear map from  $X$  to the reals  $\mathbb{R}$ ,  $\ell \in \mathcal{L}(X, \mathbb{R})$

A linear functional is a one dimensional “snapshot” of a vector

For  $v$  in  $\mathbb{R}^n$  fixed, the map

$$\ell(x) = v \cdot x = (x, v)$$

is a linear functional on  $\mathbb{R}^n$

The linear functional on  $\mathbb{R}^n$  given by the inner product with the basis vector  $e_i$  gives the  $i^{\text{th}}$  component of a vector

# Definition of Linear Functionals

For a continuous function  $f$  on  $[a, b]$ ,

$$\ell(f) = \int_a^b f(x) dx \quad \text{and} \quad \ell(f) = f(y) \text{ for } a \leq y \leq b$$

are linear functionals

Statistical moments like the expected value  $E(X)$  of a random variable  $X$  are linear functionals

The Fourier coefficients of a continuous function  $f$  on  $[0, 2\pi]$ ,

$$c_j = \int_0^{2\pi} f(x) e^{-ijx} dx$$

are linear functionals of  $f$

# Sets of Linear Functionals

Presumably, it is easier to compute accurate snapshots than the entire functions

We may use a set of snapshots to attempt to describe a function

In many situations, we settle for an “incomplete” set of snapshots

We are often happy with a small set of moments of a random variable

In practical applications of Fourier series, we truncate the infinite series to a finite number of terms,

$$\sum_{j=-\infty}^{\infty} c_j e^{ijx} \rightarrow \sum_{j=-J}^J c_j e^{ijx}$$

# Are There Many linear functionals?

Are there many bounded linear functionals on an arbitrary normed vector space?

The celebrated Hahn-Banach theorem provides a way to generate a large number

Hahn-Banach Theorem Let  $X$  be a Banach space and  $X_0$  a subspace of  $X$ . Suppose that  $F_0(x)$  is a bounded linear functional defined on  $X_0$ . There is a linear functional  $F$  defined on  $X$  such that  $F(x) = F_0(x)$  for  $x$  in  $X_0$  and  $\|F\| = \|F_0\|$ .

# Linear Functionals and Dual Spaces

We are interested in the structure of the set of all linear functionals

If  $X$  is a normed vector space, the vector space  $\mathcal{L}(X, \mathbb{R})$  of continuous linear functionals on  $X$  is called the dual space of  $X$ , and is denoted by  $X^*$

The dual space is a normed vector space under the **dual norm** defined for  $y \in X^*$  as

$$\|y\|_{X^*} = \sup_{\substack{x \in X \\ \|x\|_X=1}} |y(x)| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{|y(x)|}{\|x\|}$$

size = largest value of the snapshot on vectors of length 1

# Linear Functionals and Dual Spaces

Consider  $X = \mathbb{R}^n$ . Every vector  $v$  in  $\mathbb{R}^n$  is associated with a linear functional  $F_v(\cdot) = (\cdot, v)$ . This functional is bounded since  $|(x, v)| \leq \|v\| \|x\| = C\|x\|$

A classic result in linear algebra is that *all* linear functionals on  $\mathbb{R}^n$  have this form, i.e., we can make the identification  
 $(\mathbb{R}^n)^* \simeq \mathbb{R}^n$

# Linear Functionals and Dual Spaces

Recall Hölder's inequality: if  $f \in L^p(\Omega)$  and  $g \in L^q(\Omega)$  with  $p^{-1} + q^{-1} = 1$  for  $1 \leq p, q \leq \infty$ , then

$$\|fg\|_{L^1(\Omega)} \leq \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}$$

Each  $g$  in  $L^q(\Omega)$  is associated with a bounded linear functional on  $L^p(\Omega)$  when  $p^{-1} + q^{-1} = 1$  and  $1 \leq p, q \leq \infty$  by

$$F(f) = \int_{\Omega} g(x)f(x) dx$$

We can “identify”  $(L^p)^*$  with  $L^q$  for  $1 < p, q < \infty$ ,  $p^{-1} + q^{-1} = 1$

The cases  $p = 1, q = \infty$  and  $p = \infty, q = 1$  are trickier

# Duality for Hilbert Spaces

Hilbert spaces are Banach spaces with an inner product  $(\cdot, \cdot)$

$\mathbb{R}^n$  and  $L^2$  are Hilbert spaces

If  $X$  is a Hilbert space, then  $\psi \in X$  determines a bounded linear functional via the inner product

$$\ell_\psi(x) = (x, \psi), \quad x \in X$$

The Riesz Representation theorem says this is the only kind of linear functional on a Hilbert space

We can identify the dual space of a Hilbert space with itself

Linear functionals are commonly represented as inner products

# Riesz Representors for Hilbert Spaces

Some useful choices of Riesz representors  $\psi$  for functions  $f$ :

- ▶  $\psi = \chi_\omega / |\omega|$  gives the error in the average value of  $f$  over a subset  $\omega \subset \Omega$ , where  $\chi_\omega$  is the characteristic function of  $\omega$
- ▶  $\psi = \delta_c$  gives the average value  $\oint_c f(s) ds$  of  $f$  on a curve  $c$  in  $\mathbb{R}^n$ ,  $n = 2, 3$ , and  $\psi = \delta_s$  gives the average value of  $f$  over a plane surface  $s$  in  $\mathbb{R}^3$  ( $\delta$  denotes the corresponding delta function)
- ▶ We can obtain average values of derivatives using dipoles similarly
- ▶  $\psi = f / \|f\|$  gives the  $L^2$  norm of  $f$

Only some of these  $\psi$  have spatially local support

# The Bracket Notation

We “borrow” the Hilbert space notation for the general case:

If  $x$  is in  $X$  and  $y$  is in  $X^*$ , we denote the value

$$y(x) = \langle x, y \rangle$$

This is called the bracket notation

The generalized Cauchy inequality is

$$|\langle x, y \rangle| \leq \|x\|_X \|y\|_{X^*}, \quad x \in X, y \in X^*$$

# Motivation for the Adjoint Operator

Let  $X, Y$  be normed vector spaces

Assume that  $L \in \mathcal{L}(X, Y)$  is a continuous linear map

The goal is to compute a snapshot or functional value of the output

$$\ell(L(x)), \quad \text{some } x \in X$$

Some important questions:

- ▶ Can we find a way to compute the snapshot value efficiently?
- ▶ What is the error in the snapshot value if approximations are involved?
- ▶ Given a collection of snapshot values, what can we say about  $L$ ?
- ▶ Given a snapshot value, what can we say about  $x$ ?

# Definition of the Adjoint Operator

Let  $X, Y$  be normed vector spaces with dual spaces  $X^*, Y^*$

Assume that  $L \in \mathcal{L}(X, Y)$  is a continuous linear map

For each  $y^* \in Y^*$  there is an  $x^* \in X^*$  defined by

$$x^*(x) = y^*(L(x))$$

*snapshot of  $x$  in  $X$  = snapshot of image  $L(x)$  of  $x$  in  $Y$*

The adjoint map  $L^* : Y^* \rightarrow X^*$  satisfies the bilinear identity

$$\langle L(x), y^* \rangle = \langle x, L^*(y^*) \rangle, \quad x \in X, y^* \in Y^*$$

# Adjoint of a Matrix

Let  $X = \mathbb{R}^m$  and  $Y = \mathbb{R}^n$  with the standard inner product and norm

$L \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$  is associated with a  $n \times m$  matrix  $A$ :

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

and

$$y_i = \sum_{j=1}^m a_{ij}x_j, \quad 1 \leq i \leq n$$

# Adjoint of a Matrix

The bilinear identity reads

$$(Lx, y) = (x, L^*y), \quad x \in \mathbb{R}^m, y \in \mathbb{R}^n.$$

For a linear functional  $y^* = (y_1^*, \dots, y_n^*)^\top \in Y^*$

$$\begin{aligned} L^*y^*(x) &= y^*(L(x)) = \left( (y_1^*, \dots, y_n^*), \begin{pmatrix} \sum_{j=1}^m a_{1j}x_j \\ \vdots \\ \sum_{j=1}^m a_{nj}x_j \end{pmatrix} \right) \\ &= \sum_{j=1}^m y_1^* a_{1j} x_j + \cdots + \sum_{j=1}^m y_n^* a_{nj} x_j \\ &= \sum_{j=1}^m \left( \sum_{i=1}^n y_i^* a_{ij} \right) x_j \end{aligned}$$

# Adjoint of a Matrix

$L^*(y^*)$  is given by the inner product with  $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_m)^\top$  where

$$\tilde{y}_j = \sum_{i=1}^n y_i^* a_{ij}.$$

The matrix  $A^*$  of  $L^*$  is

$$A^* = \begin{pmatrix} a_{11}^* & \cdots & a_{1n}^* \\ \vdots & & \vdots \\ a_{m1}^* & \cdots & a_{mn}^* \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ \vdots & & & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix} = A^\top.$$

# Practical Point about Defining an Adjoint

We can define  $L^*$  for a linear operator  $L : X \rightarrow Y$  by considering  $L$  on a dense subset of  $X$

We define the adjoint to the restriction of  $L$  on a dense subset of  $X$

If this “restricted” adjoint is uniformly continuous on the dense subset, then the Hahn-Banach theorem implies that the restriction can be extended continuously to all of  $X$

This is useful in the weak formulation of differential equations in Sobolev spaces

# Properties of Adjoint Operators

Theorem Let  $X$ ,  $Y$ , and  $Z$  be normed linear spaces. For  $L_1, L_2 \in \mathcal{L}(X, Y)$ :

$$L_1^* \in \mathcal{L}(Y^*, X^*)$$

$$\|L_1^*\| = \|L_1\|$$

$$0^* = 0$$

$$(L_1 + L_2)^* = L_1^* + L_2^*$$

$$(\alpha L_1)^* = \alpha L_1^*, \quad \text{all } \alpha \in \mathbb{R}$$

If  $L_2 \in \mathcal{L}(X, Y)$  and  $L_1 \in \mathcal{L}(Y, Z)$ , then  $(L_1 L_2)^* \in \mathcal{L}(Z^*, X^*)$  and

$$(L_1 L_2)^* = L_2^* L_1^*.$$

# Adjoints for Differential Operators

The adjoint of the differential operator  $L$

$$(Lu, v) \rightarrow (u, L^*v)$$

is obtained by a succession of integration by parts

Boundary terms involving functions and derivatives arise from each integration by parts

We use a two step process

1. We first compute the formal adjoint by assuming that all functions have compact support and ignoring boundary terms
2. We then compute the adjoint boundary and data conditions to make the bilinear identity hold

# Formal Adjoints

Consider

$$Lu(x) = -\frac{d}{dx} \left( a(x) \frac{d}{dx} u(x) \right) + \frac{d}{dx} (b(x)u(x))$$

on  $[0, 1]$ . Integration by parts gives

$$\begin{aligned} & - \int_0^1 \frac{d}{dx} \left( a(x) \frac{d}{dx} u(x) \right) v(x) dx \\ &= \int_0^1 a(x) \frac{d}{dx} u(x) \frac{d}{dx} v(x) dx - a(x) \frac{d}{dx} u(x)v(x) \Big|_0^1 \\ &= - \int_0^1 u(x) \frac{d}{dx} \left( a(x) \frac{d}{dx} v(x) \right) dx + u(x)a(x) \frac{d}{dx} v(x) \Big|_0^1 \end{aligned}$$

# Formal Adjoints

$$\begin{aligned} & \int_0^1 \frac{d}{dx} (b(x)u(x)) v(x) dx \\ &= - \int_0^1 u(x) b(x) \frac{d}{dx} v(x) dx + b(x)u(x)v(x) \Big|_0^1, \end{aligned}$$

Neglecting boundary terms gives the formal adjoint

$$\begin{aligned} Lu(x) &= -\frac{d}{dx} \left( a(x) \frac{d}{dx} u(x) \right) + \frac{d}{dx} (b(x)u(x)) \\ \implies L^*v &= -\frac{d}{dx} \left( a(x) \frac{d}{dx} v(x) \right) - b(x) \frac{d}{dx} (v(x)) \end{aligned}$$

# Formal Adjoints

In higher space dimensions, we use the **divergence theorem**

A general linear second order differential operator  $L$  in  $\mathbb{R}^n$ :

$$L(u) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i \frac{\partial u}{\partial x_i} + cu,$$

where  $\{a_{ij}\}$ ,  $\{b_i\}$ , and  $c$  are functions of  $x_1, x_2, \dots, x_n$ . Then,

$$L^*(u) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 (a_{ij}v)}{\partial x_i \partial x_j} - \sum_{i=1}^n \frac{\partial (b_i v)}{\partial x_i} + cv$$

# Formal Adjoints

It can be verified directly that

$$vL(u) - uL^*(v) = \sum_{i=1}^n \frac{\partial p_i}{\partial x_i},$$

where

$$p_i = \sum_{j=1}^n \left( a_{ij} v \frac{\partial u}{\partial x_j} - u \frac{\partial (a_{ij} v)}{\partial x_j} \right) + b_i u v.$$

The divergence theorem yields

$$\int_{\Omega} (vL(u) - uL^*(v)) dx = \int_{\partial\Omega} p \cdot n ds = 0,$$

where  $p = (p_1, \dots, p_n)$  and  $n$  is the outward normal in  $\partial\Omega$ .

# Formal Adjoints

A typical term,

$$\begin{aligned}va_{11}\frac{\partial^2 u}{\partial x_1^2} &= va_{11}\frac{\partial}{\partial x_1}\left(\frac{\partial u}{\partial x_1}\right) = \frac{\partial}{\partial x_1}\left(va_{11}\frac{\partial u}{\partial x_1}\right) - \frac{\partial(a_{11}v)}{\partial x_1}\frac{\partial u}{\partial x_1} \\&= \frac{\partial}{\partial x_1}\left(va_{11}\frac{\partial u}{\partial x_1}\right) - \frac{\partial}{\partial x_1}\left(u\frac{\partial(a_{11}v)}{\partial x_1}\right) + u\frac{\partial^2(a_{11}v)}{\partial x_1^2}\end{aligned}$$

yielding

$$va_{11}\frac{\partial^2 u}{\partial x_1^2} - u\frac{\partial^2(a_{11}v)}{\partial x_1^2} = \frac{\partial}{\partial x_1}\left(a_{11}v\frac{\partial u}{\partial x_1} - u\frac{\partial(a_{11}v)}{\partial x_1}\right).$$

# Formal Adjoints

Important examples:

$$\text{grad}^* = -\text{div}$$

$$\text{div}^* = -\text{grad}$$

$$\text{curl}^* = \text{curl}$$

$$Lu = \sum_{|\alpha| \leq p} a_\alpha(x) D^\alpha u$$

then

$$L^*v = \sum_{|\alpha| \leq p} (-1)^{|\alpha|} D^\alpha (a_\alpha(x)v(x)).$$

# Adjoint Boundary Conditions

In the second stage, we deal with the boundary terms that arise during integration by parts

The standard adjoint boundary conditions are the **minimal** conditions required so the bilinear identity holds true

The form of the boundary conditions imposed on the differential operator is important, but not the values

We assume homogeneous boundary values for the differential operator when determining the adjoint conditions

We may need to use other boundary conditions, e.g. if the quantity of interest is determined on a boundary

# Adjoint Boundary Conditions

Consider

$$\begin{cases} s''(x) = f(x), & 0 < x < 1, \\ s(0) = s'(0) = 0 \end{cases}$$

Integrating by parts,

$$\int_0^1 (s''v - sv'') dx = (vs' - sv')|_0^1$$

The boundary conditions imply the contributions at  $x = 0$  vanish, while at  $x = 1$  we have

$$v(1)s'(1) - v'(1)s(1)$$

The adjoint boundary conditions are  $v(1) = v'(1) = 0$  since we cannot specify  $s'(1)$  or  $s(1)$

# Adjoint Boundary Conditions

Recall

$$\int_{\Omega} (u \Delta v - v \Delta u) dx = \int_{\partial\Omega} \left( u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \right) ds,$$

The Dirichlet and Neumann boundary value problems for the Laplacian are their own adjoints

# Adjoint Boundary Conditions

Let  $\Omega \subset \mathbb{R}^2$  be bounded with a smooth boundary and let  $s =$  arclength along the boundary

Consider

$$\begin{cases} -\Delta u = f, & x \in \Omega, \\ \frac{\partial u}{\partial n} + \frac{\partial u}{\partial s} = 0, & x \in \partial\Omega \end{cases}$$

Since

$$\int_{\Omega} (u \Delta v - v \Delta u) dx = \int_{\partial\Omega} \left( u \left( \frac{\partial v}{\partial n} - \frac{\partial v}{\partial s} \right) - v \left( \frac{\partial u}{\partial n} + \frac{\partial u}{\partial s} \right) \right) ds,$$

the adjoint problem is

$$\begin{cases} -\Delta v = g, & x \in \Omega, \\ \frac{\partial v}{\partial n} - \frac{\partial v}{\partial s} = 0, & x \in \partial\Omega. \end{cases}$$

# Adjoint for an Evolution Operator

For an initial value problem, we have  $\frac{d}{dt}$  and an initial condition

Now

$$\int_0^T \frac{du}{dt} v \, dt = u(t)v(t) \Big|_0^T - \int_0^T u \frac{dv}{dt} \, dt$$

The boundary term at 0 vanishes because  $u(0) = 0$

The adjoint is a final-value problem with “initial” condition  
 $v(T) = 0$

The adjoint problem has  $-\frac{dv}{dt}$  and time “runs backwards”

# Adjoint for an Evolution Operator

$$\begin{cases} Lu = \frac{du}{dt} - \Delta u = f, & x \in \Omega, 0 < t \leq T, \\ u = 0, & x \in \partial\Omega, 0 < t \leq T, \\ u = u_0, & x \in \Omega, t = 0 \end{cases}$$

$$\implies \begin{cases} L^*v = -\frac{dv}{dt} - \Delta v = \psi, & x \in \Omega, T > t \geq 0, \\ v = 0, & x \in \partial\Omega, T > t \geq 0, \\ v = 0, & x \in \Omega, t = T \end{cases}$$

A useful alternative is

$$\implies \begin{cases} L^*v = -\frac{dv}{dt} - \Delta v = 0, & x \in \Omega, T > t \geq 0, \\ v = 0, & x \in \partial\Omega, T > t \geq 0, \\ v = \psi, & x \in \Omega, t = T \end{cases}$$

## *The Usefulness of Duality and Adjoints*

(6)

# The Dual Space is Well Behaved

The dual space can be better behaved than the original normed vector space

If  $X$  is a normed vector space over  $\mathbb{R}$ , then  $X^*$  is a Banach space whether or not  $X$  is a Banach space

This can be exploited in analysis in  $X$ , e.g. “weak” convergence results

# Condition of an Operator

There is an intimate connection between the adjoint operator and the stability properties of the original operator

The singular values of a matrix  $\mathbf{L}$  are the square roots of the eigenvalues of the **square, symmetric** transformations  $\mathbf{L}^*\mathbf{L}$  or  $\mathbf{LL}^*$

This connects the condition number of a matrix  $\mathbf{L}$  to  $\mathbf{L}^*$

# Solving Problems for Operators

Given normed vector spaces  $X$  and  $Y$ , an operator  $\mathcal{L}(X, Y)$ , and  $b \in Y$ , find  $x \in X$  such that

$$Lx = b$$

A necessary condition that  $b$  is in the range of  $L$  is  $y^*(b) = 0$  for all  $y^*$  in the null space of  $L^*$

This is a sufficient condition if the range of  $L$  is closed in  $Y$

If  $A$  is an  $n \times m$  matrix, a necessary and sufficient condition for the solvability of  $Ax = b$  is  $b$  is orthogonal to all linearly independent solutions of  $A^\top y = 0$

# Solving Problems for Operators

When  $X$  is a Hilbert space and  $L \in \mathcal{L}(X, Y)$ , then the range of  $L^*$  is a subset of the orthogonal complement of the null space of  $L$

If the range of  $L^*$  is “large”, then the orthogonal complement of the null space of  $L$  must be “large” and the null space of  $L$  must be “small”

The existence of sufficiently many solutions of the homogeneous adjoint equation  $L^*\phi = 0$  implies there is at most one solution of  $Lu = b$  for a given  $b$

This is the basis for the Holmgren Uniqueness Theorem in PDEs

# Green's Functions

Suppose we wish to compute a functional  $\ell(x)$  of the solution  $x \in \mathbb{R}^n$  of a  $n \times n$  system

$$\mathbf{L}x = b$$

For a linear functional  $\ell(\cdot) = (\cdot, \psi)$ , we define the adjoint problem

$$\mathbf{L}^*\phi = \psi$$

Variational analysis yields the representation formula

$$\ell(x) = (x, \psi) = (x, \mathbf{L}^*\phi) = (\mathbf{L}x, \phi) = (b, \phi)$$

We can compute many solutions by computing one adjoint solution and taking inner products

# Green's Functions

For

$$\begin{cases} -\Delta u = f, & x \in \Omega, \\ u = 0, & x \in \partial\Omega \end{cases}$$

the Green's function solves

$$-\Delta \phi = \delta_x \quad (\text{delta function at } x)$$

This yields

$$u(x) = (u, \delta_x) = (f, \phi)$$

The **generalized Green's function** solves the adjoint problem with general functional data, rather than just  $\delta_x$

The imposition of the adjoint boundary conditions is crucial

## *Nonlinear Operators*

(13)

# Nonlinear Operators

There is no unique “natural” adjoint for a general nonlinear operator

We assume that the Banach spaces  $X$  and  $Y$  are Sobolev spaces and use  $(\cdot, \cdot)$  for the  $L^2$  inner product, and so forth

We define the adjoint for a specific kind of nonlinear operator

We assume  $f$  is a nonlinear map from  $X$  into  $Y$ , where the domain of  $f$  is a convex set

# A Perturbation Operator

Consider a “true”  $u$  and “approximation”  $U$  in the domain of  $f$  and define the “error”  $e = U - u$

$u$  and  $e$  are “unknown”

Define

$$F(e) = f(u + e) - f(u),$$

The domain of  $F$  is

$$\{v \in X \mid v + u \in \text{domain of } f\}$$

Assume the domain of  $F$  is independent of  $e$  and dense in  $X$

Note that 0 is in the domain of  $F$  and  $F(0) = 0$

# A Perturbation Operator

Two reasons to work with functions of this form:

- ▶ This is the kind of nonlinearity that arises when estimating the error of a numerical solution or studying the effects of perturbations
- ▶ Nonlinear problems typically do **not** enjoy the global solvability that characterizes linear problems, only a local solvability

## Definition 1

A common definition is based on the bilinear identity

An operator  $A^*(e)$  is an adjoint operator corresponding to  $F$  if

$$(F(e), w) = (e, A^*(e)w) \quad \text{for all } e \in \text{domain of } F, w \in \text{domain of } A^*$$

This is an adjoint operator associated with  $F$ , not the adjoint operator to  $F$

## Definition 1

Suppose that  $F$  can be represented as  $F(e) = A(e)e$ , where  $A(e)$  is a linear operator with the domain of  $F$  contained in the domain of  $A$

For a fixed  $e$  in the domain of  $F$ , define the adjoint of  $A$  satisfying

$$(A(e)w, v) = (w, A^*(e)v)$$

for all  $w \in \text{domain of } A$ ,  $v \in \text{domain of } A^*$

Substituting  $w = e$  shows this defines an adjoint of  $F$  as well

If there are several such linear operators  $A$ , then there will be several different possible adjoints.

## Definition 1

Let  $(t, x) \in \Omega = (0, 1) \times (0, 1)$ , with  $X = X^* = Y = Y^* = L^2$  denoting the space of periodic functions in  $t$  and  $x$ , with period equal to 1

Consider a periodic problem

$$F(e) = \frac{\partial e}{\partial t} + e \frac{\partial e}{\partial x} + ae = f$$

where  $a > 0$  is a constant and the domain of  $F$  is the set of continuously differentiable functions.

# Definition 1

We can write  $F(e) = A_i(e)e$  where

$$A_1(e)v = \frac{\partial v}{\partial t} + e \frac{\partial v}{\partial x} + av \implies A_1^*(e)w = -\frac{\partial w}{\partial t} - \frac{\partial(ew)}{\partial x} + aw$$

$$A_2(e)v = \frac{\partial v}{\partial t} + \left(a + \frac{\partial e}{\partial x}\right)v \implies A_2^*(e)w = -\frac{\partial w}{\partial t} + \left(a + \frac{\partial e}{\partial x}\right)w$$

$$A_3(e)v = \frac{\partial v}{\partial t} + \frac{1}{2} \frac{\partial(ev)}{\partial x} + av \implies A_3^*(e)w = -\frac{\partial w}{\partial t} - \frac{e}{2} \frac{\partial w}{\partial x} + aw$$

## Definition 2

If the nonlinearity is Frechet differentiable, we base the **second** definition of an adjoint on the integral mean value theorem

The integral mean value theorem states

$$f(U) = f(u) + \int_0^1 f'(u + se) ds \quad e$$

where  $e = U - u$  and  $f'$  is the Frechet derivative of  $f$

## Definition 2

We rewrite this as

$$F(e) = f(U) - f(u) = A(e)e$$

with

$$A(e) = \int_0^1 f'(u + se) ds$$

Note that we can apply the integral mean value theorem to  $F$ :

$$A(e) = \int_0^1 F'(se) ds$$

To be precise, we should discuss the smoothness of  $F$

## Definition 2

For a fixed  $e$ , the adjoint operator  $A^*(e)$ , defined in the usual way for the linear operator  $A(e)$ , is said to be an adjoint for  $F$

Continuing the previous example,

$$F'(e)v = \frac{\partial v}{\partial t} + e \frac{\partial v}{\partial x} + \left( a + \frac{\partial e}{\partial x} \right) v.$$

After some technical analysis of the domains of the operators involved,

$$A^*(e)w = -\frac{\partial w}{\partial t} - \frac{e}{2} \frac{\partial w}{\partial x} + aw.$$

This coincides with the third adjoint computed above

# A Point about Linearization

The linearization approach is used in many practical applications

Typically the linearization in the integral mean value theorem is approximated by linearization at a “known” point

This raises the issue of the effect of such linearization “error” on any subsequent use of the adjoint

For many analysis purposes, the issue is the accuracy of the inverse of the approximation

# A Point about Linearization

Consider  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ :

$$F(u) = \begin{pmatrix} u_1^2 + 3u_2 \\ u_1 e^{u_2} \end{pmatrix}$$

For perturbation  $\varepsilon = (\varepsilon_1, \varepsilon_2)^\top$ ,

$$\mathcal{E}(\varepsilon) = F(u + \varepsilon) - F(u) = \begin{pmatrix} 2u_1 + \varepsilon_1 & 3 \\ e^{u_2 + \varepsilon_2} & u_1 e^{u_2} \left( \frac{e^{\varepsilon_2} - 1}{\varepsilon_2} \right) \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

This yields

$$\mathcal{E}(\varepsilon)^* = \begin{pmatrix} 2u_1 + \varepsilon_1 & e^{u_2 + \varepsilon_2} \\ 3 & u_1 e^{u_2} \left( \frac{e^{\varepsilon_2} - 1}{\varepsilon_2} \right) \end{pmatrix}$$

# A Point about Linearization

For small  $\varepsilon$ ,

$$\mathcal{E}(\varepsilon)^* \approx (F'(u))^*$$

where  $F'(u)$  is the Jacobian,

$$F'(u) = \begin{pmatrix} 2u_1 & 3 \\ e^{u_2} & u_1 e^{u_2} \end{pmatrix}$$

In practical computations, we use  $F'(u)^*$

For  $|u_1|$  bounded away from  $\sqrt{3/2}$ ,  $(F'(u)^*)^{-1} \approx (\mathcal{E}(v)^*)^{-1}$  for all  $\varepsilon$  with sufficiently small norm

If  $|u_1| \approx \sqrt{3/2}$ ,  $(F'(u)^*)^{-1}$  may not be close to  $(\mathcal{E}(\varepsilon)^*)^{-1}$



## *A Posteriori Error Analysis*

(28)

# *A Posteriori* Error Analysis

Problem: Estimate the error in a quantity of interest computed using a numerical solution of a differential equation

We assume that the quantity of information can be represented as a linear functional of the solution

We use the adjoint problem associated with the linear functional

# What about Convergence Analysis?

Recall the standard *a priori* convergence result for an initial value problem

$$\begin{cases} \dot{y} = f(y), & 0 < t, \\ y(0) = y_0 \end{cases}$$

Let  $Y \approx y$  be an approximation associated with time step  $\Delta t$

A typical *a priori* (Gronwall argument) bound is

$$\|Y - y\|_{L^\infty(0,t)} \leq C e^{Lt} \Delta t^p \left\| \frac{d^{p+1}y}{dt^{p+1}} \right\|_{L^\infty(0,t)}$$

$L$  is often large in practice, e.g.  $L \sim 100 - 10000$

It is typical for an *a priori* convergence bound to be orders of magnitude larger than the error

# A Linear Algebra Problem

We compute a quantity of interest  $(u, \psi)$  from a solution of

$$\mathbf{A}u = b$$

If  $U$  is an approximate solution, we estimate the error

$$(e, \psi) = (u - U, \psi)$$

We can compute the residual

$$R = \mathbf{A}U - b$$

Using the adjoint problem  $\mathbf{A}^\top \phi = \psi$ , variational analysis gives

$$|(e, \psi)| = |(e, \mathbf{A}^\top \phi)| = |(Ae, \phi)| = |(R, \phi)|$$

We solve for  $\phi$  numerically to compute the estimate

# Finite Element Method for Elliptic Problem

The a posteriori analysis naturally applies to finite element discretizations

Variational form of the equation is formed by multiplying by a test function, integrating over space and time, and using integration by parts to reduce derivative orders

Appropriate function spaces must be chosen

The finite element approximation uses finite dimensional function spaces, e.g. piecewise polynomials

Finite volume and finite difference methods are written as finite elements + quadrature

# Finite Element Method for Elliptic Problem

Approximate  $u : \mathbb{R}^n \rightarrow \mathbb{R}$  solving

$$\begin{cases} Lu = f, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases}$$

where

$$L(D, x)u = -\nabla \cdot a(x)\nabla u + b(x) \cdot \nabla u + c(x)u(x),$$

- ▶  $\Omega \subset \mathbb{R}^n$ ,  $n = 1, 2, 3$ , is a convex polygonal domain
- ▶  $a = (a_{ij})$ , where  $a_{i,j}$  are continuous and there is a  $a_0 > 0$  such that  $v^\top a v \geq a_0$  for all  $v \in \mathbb{R}^n \setminus \{0\}$  and  $x \in \Omega$
- ▶  $b = (b_i)$  where  $b_i$  is continuous
- ▶  $c$  and  $f$  are continuous

# Finite Element Method for Elliptic Problem

The variational formulation reads

Find  $u \in H_0^1(\Omega)$  such that

$$A(u, v) = (a\nabla u, \nabla v) + (b \cdot \nabla u, v) + (cu, v) = (f, v)$$

for all  $v \in H_0^1(\Omega)$

$H_0^1(\Omega)$  is the space of  $L^2(\Omega)$  functions whose first derivatives are in  $L^2(\Omega)$

This says that the solution solves the “average” form of the problem for a large number of weights  $v$

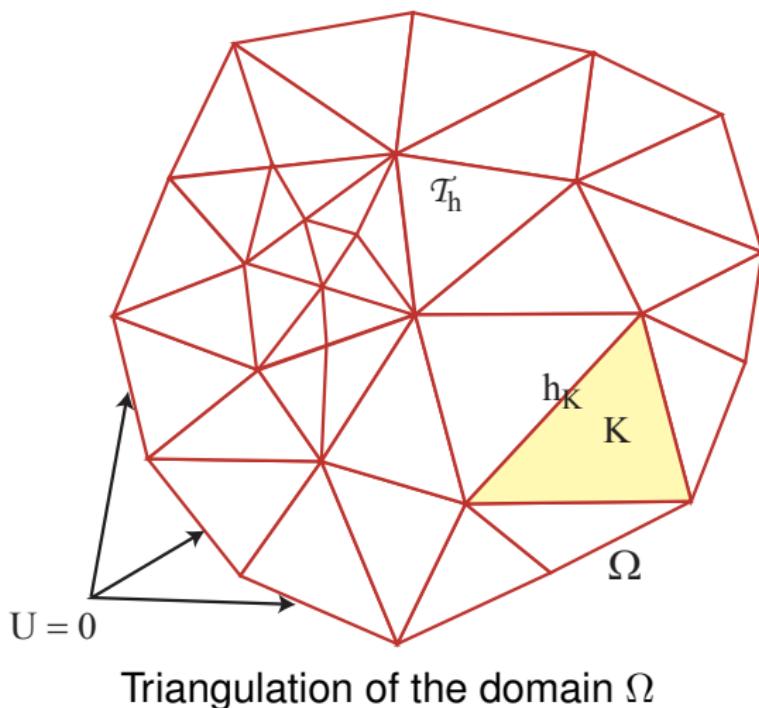
# Finite Element Method for Elliptic Problem

We construct a triangulation of  $\Omega$  into simplices, or elements, such that boundary nodes of the triangulation lie on  $\partial\Omega$

$\mathcal{T}_h$  denotes a simplex triangulation of  $\Omega$  that is locally quasi-uniform, i.e. no arbitrarily long, skinny triangles

We use the length of the longest edge  $h_K$  of  $K \in \mathcal{T}_h$  to quantify size

# Finite Element Method for Elliptic Problem



# Finite Element Method for Elliptic Problem

$V_h$  denotes the space of functions that are

- ▶ continuous on  $\Omega$
- ▶ piecewise linear with respect to  $\mathcal{T}_h$
- ▶ zero on the boundary

$V_h \subset H_0^1(\Omega)$ , and for smooth functions, the error of interpolation into  $V_h$  is  $\mathbf{O}(h^2)$  in  $\| \cdot \|$

The finite element method is:

Compute  $U \in V_h$  such that  $A(U, v) = (f, v)$  for all  $v \in V_h$

This is Galerkin orthogonality since it is equivalent to  
 $A(U - u, v) = 0$  for all  $v \in V_h$

# A Posteriori Analysis for an Elliptic Problem

We assume that quantity of interest is the functional  $(u, \psi)$

The **generalized Green's function**  $\phi$  solves the weak adjoint problem : Find  $\phi \in H_0^1(\Omega)$  such that

$$A^*(v, \phi) = (\nabla v, a \nabla \phi) - (v, \operatorname{div}(b\phi)) + (v, c\phi) = (v, \psi) \quad \text{for all } v \in H_0^1(\Omega),$$

corresponding to the adjoint problem  $L^*(D, x)\phi = \psi$

# A Posteriori Analysis for an Elliptic Problem

We now estimate the error  $e = U - u$ :

$$\begin{aligned}(e, \psi) &= (\nabla e, a \nabla \phi) - (e, \operatorname{div}(b\phi)) + (e, c\phi) \\&= (a \nabla e, \nabla \phi) + (b \cdot \nabla e, \phi) + (ce, \phi) \\&= (a \nabla u, \nabla \phi) + (b \cdot \nabla u, \phi) + (cu, \phi) \\&\quad - (a \nabla U, \nabla \phi) - (b \cdot \nabla U, \phi) - (cU, \phi) \\&= (f, \phi) - (a \nabla U, \nabla \phi) - (b \cdot \nabla U, \phi) - (cU, \phi)\end{aligned}$$

The weak residual of  $U$  is

$$R(U, v) = (f, v) - (a \nabla U, \nabla v) - (b \cdot \nabla U, v) - (cU, v), \quad v \in H_0^1(\Omega)$$

$$R(U, v) = 0 \text{ for } v \in V_h \text{ but not for general } v \in H_0^1(\Omega)$$

# *A Posteriori* Analysis for an Elliptic Problem

$\pi_h\phi$  denotes an approximation of  $\phi$  in  $V_h$

Theorem The error representation is,

$$\begin{aligned}(e, \psi) = & (f, \phi - \pi_h\phi) - (a\nabla U, \nabla(\phi - \pi_h\phi)) \\ & - (b \cdot \nabla U, \phi - \pi_h\phi) - (cU, \phi - \pi_h\phi),\end{aligned}$$

Subtracting the projection  $\phi - \pi_h\phi$  is not needed theoretically  
but useful in practice

The subtraction is required for standard adaptive error control

# *A Posteriori* Analysis for an Elliptic Problem

We approximate  $\phi$  using a higher order finite element method or on a significantly finer mesh

For a second order elliptic problem, good results are obtained using the space  $V_h^2$

The approximate generalized Green's function  $\Phi \in V_h^2$  solves

$$A^*(v, \Phi) = (\nabla v, a \nabla \Phi) - (v, \operatorname{div}(b\Phi)) + (v, c\Phi) = (v, \psi) \text{ for all } v \in V_h^2$$

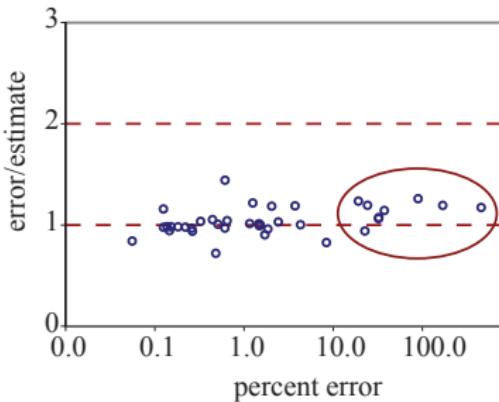
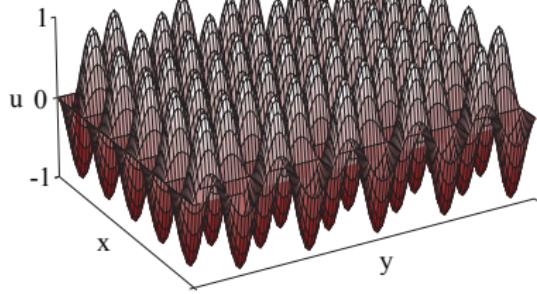
Theorem The approximate error representation is

$$\begin{aligned} (e, \psi) \approx & (f, \Phi - \pi_h \Phi) - (a \nabla U, \nabla(\Phi - \pi_h \Phi)) \\ & - (b \cdot \nabla U, \Phi - \pi_h \Phi) - (cU, \Phi - \pi_h \Phi) \end{aligned}$$

# A Posteriori Analysis for an Elliptic Problem

$$\begin{cases} -\Delta u = 200\pi^2 \sin(10\pi x) \sin(10\pi y), & (x, y) \in \Omega = [0, 1] \times [0, 1], \\ u = 0, & (x, y) \in \partial\Omega \end{cases}$$

The solution is  $u = \sin(10\pi x) \sin(10\pi y)$



# *A Posteriori* Analysis for an Initial Value Problem

We develop an estimate for

$$\begin{cases} y' = f(y), & 0 < t \leq T, \\ y(0) = y_0 \end{cases}$$

where  $f$  is smooth

Discretization of domain:

$$0 = t_0 < t_1 < \cdots < t_N = T,$$

$$I_n = (t_{n-1}, t_n], \quad k_n = t_n - t_{n-1}, \quad k|_{I_n} = k_n$$

Space of approximation:

$\mathcal{P}^q(I_n)$  = polynomials of degree  $q$  and less on  $I_n$

$W_n^q = \{w : w \in \mathcal{P}^q(I_n)\}$ ,  $W_n = \{w : w|_{I_n} \in W_n^q\}$

$\pi_k$  = interpolant/projection operator into  $\mathcal{P}^q(I_n)$  for each  $n$

# A Posteriori Analysis for an Initial Value Problem

The continuous Galerkin method of order  $q$  (CG( $q$ )): For  $n = 1, \dots, N$ ,

$$\begin{cases} \int_{I_n} (Y', v) dt = \int_{I_n} (f(Y), v) dt & \text{all } v \in W_n^{q-1} \\ Y_{n-1}^+ = Y_{n-1}^- \end{cases}$$

with  $Y_0 = y_0$ ,  $U_n^+ = \lim_{t \downarrow t_n} U(t)$ ,  $U_n^- = \lim_{t \uparrow t_n} U(t)$

There is a discontinuous Galerkin method

Quantity of interest:  $\int_0^T (y, \psi) dt$

$\psi = 1/T$  gives the average value

# *A Posteriori* Analysis for an Initial Value Problem

Linearization to define adjoint:

$$\bar{f}' = \overline{f'(y, Y)} = \int_0^1 \frac{\partial f}{\partial y}(sy + (1-s)Y) ds$$

Adjoint problem:

$$\begin{cases} -\phi' = (\bar{f}')^* \phi, & T > t \geq 0, \\ \phi(T) = 0 \end{cases}$$

time runs “backwards” but note the “-” on the time derivative

In practice,  $\bar{f}' \rightarrow f'(Y)$

# *A Posteriori* Analysis for an Initial Value Problem

Substituting

$$\int_0^T (e, \psi) dt = \sum_{i=1}^N \int_{I_n} (e, -\phi' - (\bar{f}')^* \phi) dt$$

$$\int_{I_n} (e, -\phi') dt = -(e_n^-, \phi_n) + (e_{n-1}^+, \phi_n) + \int_{I_n} (e', \phi) dt$$

$$\int_{I_n} (e, (\bar{f}')^* \phi) dt = \int_{I_n} (f(y) - f(Y), \phi) dt$$

Theorem The error representation is

$$\int_0^T (e, \psi) dt = \sum_{n=1}^N \int_{I_n} (\mathcal{R}_n, \pi_k \phi - \phi) dt$$

$$\mathcal{R}_n = Y' - f(Y) \text{ on } I_n$$

# *A Posteriori* Analysis for an Initial Value Problem

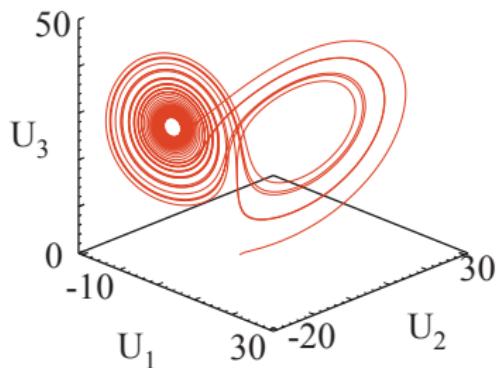
The chaotic Lorenz problem

$$\begin{cases} u'_1 = -10u_1 + 10u_2, \\ u'_2 = 28u_1 - u_2 - u_1u_3, \quad 0 < t, \\ u'_3 = -\frac{8}{3}u_3 + u_1u_2, \end{cases}$$

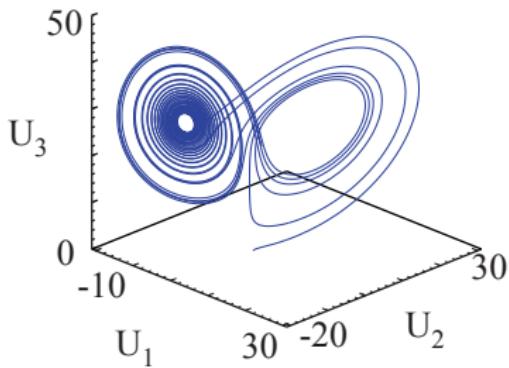
Chaotic behavior affects numerical solutions as well

# *A Posteriori* Analysis for an Initial Value Problem

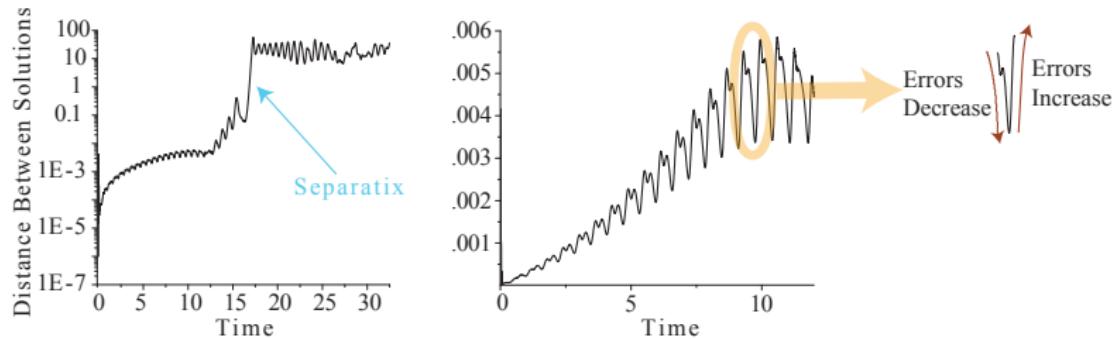
2% error on  $[0,30]$



100% error at  $t=18$



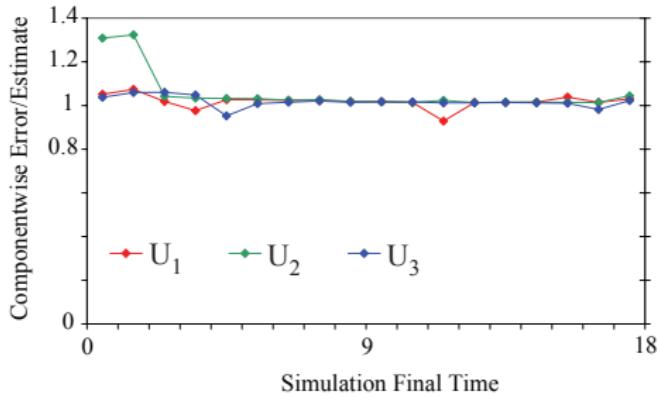
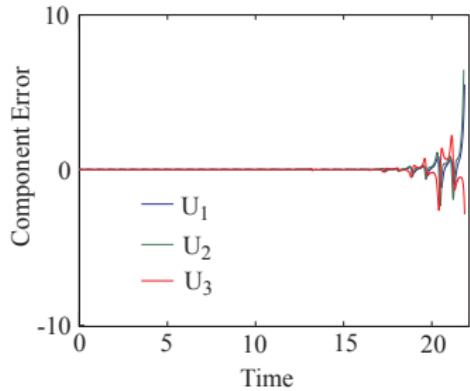
# A Posteriori Analysis for an Initial Value Problem



The distance between the two numerical solutions

The errors follow an increasing trend, but decrease as well as increase

# A Posteriori Analysis for an Initial Value Problem



The accuracy of the estimate in the pointwise error

# A Posteriori Analysis for a Reaction-Diffusion System

A system of  $D$  reaction-diffusion equations consisting of  $d$  **parabolic** equations and  $D - d$  **ordinary** equations for  $u \in \mathbb{R}^D$ :

$$\begin{cases} u'_i - \nabla \cdot (\epsilon_i(u) \nabla u_i) = f_i(u), & (x, t) \in \Omega \times \mathbb{R}^+, \quad 1 \leq i \leq D, \\ u_i(x, t) = 0, & (x, t) \in \partial\Omega \times \mathbb{R}^+, \quad 1 \leq i \leq d, \\ u(x, 0) = u_0(x), & x \in \Omega, \end{cases}$$

Assumptions:

$\epsilon_i(u) \geq \epsilon_0 > 0$  for  $1 \leq i \leq d$  and  $\epsilon_i(u) \equiv 0$  for the rest

$\Omega$  is a convex polygonal domain with boundary  $\partial\Omega$

$\epsilon, f$  have smooth second derivatives

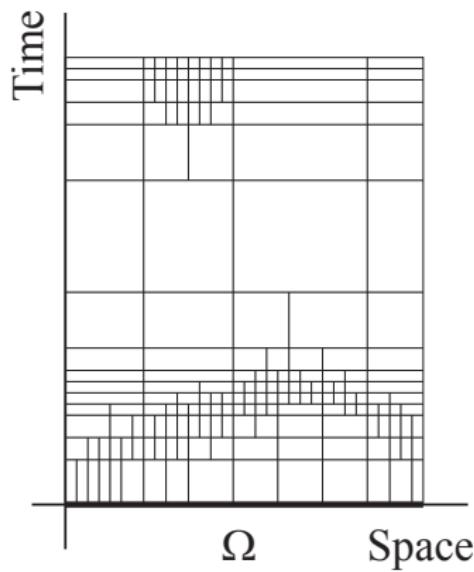
$u_i^p = u_i$  for  $1 \leq i \leq d$  and  $u_i^p = 0$  for  $d < i \leq D$

# A Posteriori Analysis for a Reaction-Diffusion System

Discretization:

$$0 = t_0 < t_1 < t_2 < \cdots < t_N = T, \quad I_n = (t_{n-1}, t_n], \quad k_n = t_n - t_{n-1}$$

$\mathcal{T}_n$  is a triangulation of  $\Omega$  for  $t \in I_n$



# A Posteriori Analysis for a Reaction-Diffusion System

Discrete space: polynomials in time and piecewise polynomials in space on each space-time “slab”  $S_n = \Omega \times I_n$

$V_n \subset (H_0^1(\Omega))^d \times (H^1(\Omega))^{D-d}$  denotes the space of piecewise linear continuous vector-valued functions  $v(x) \in \mathbb{R}^D$  defined on  $\mathcal{T}_n$ , where the first  $d$  components of  $v$  are zero on  $\partial\Omega$

Define

$$W_n^q = \left\{ w(x, t) : w(x, t) = \sum_{j=0}^q t^j v_j(x), v_j \in V_n, (x, t) \in S_n \right\}.$$

$W^q$  denotes the space of functions defined on the space-time domain  $\Omega \times \mathbb{R}^+$  such that  $v|_{S_n} \in W_n^q$  for  $n \geq 1$

# A Posteriori Analysis for a Reaction-Diffusion System

The cG(q)-cG(1) approximation  $U \in W^q$  satisfies  $U_0^- = P_0 u_0$  and for  $n \geq 1$ ,

$$\begin{cases} \int_{t_{n-1}}^{t_n} ((U'_i, v_i) + (\epsilon_i(U) \nabla U_i, \nabla v_i)) dt = \int_{t_{n-1}}^{t_n} (f_i(U), v_i) dt \\ \text{for all } v \in W_n^{q-1}, 1 \leq i \leq D, \\ U_{n-1}^+ = P_n U_{n-1}^- \end{cases}$$

$P_n$  is the  $L^2$  projection into  $V_n$

The approximation is continuous across time nodes where there is no mesh change

# A Posteriori Analysis for a Reaction-Diffusion System

Linearization for adjoint:

$$\bar{\epsilon}_i = \bar{\epsilon}_i(u, U) = \int_0^1 \epsilon_i(us + U(1-s)) ds,$$

$$\bar{\beta}_{ij} = \bar{\beta}_{ij}(u, U) = \int_0^1 \frac{\partial \epsilon_j}{\partial u_i}(us + U(1-s)) \nabla(u_i s + U_i(1-s)) ds,$$

$$\bar{f}_{ij} = \bar{f}_{ij}(u, U) = \int_0^1 \frac{\partial f_j}{\partial u_i}(us + U(1-s)) ds.$$

Adjoint problem:

$$\begin{cases} -\phi'_i - \nabla \cdot (\bar{\epsilon}_i \nabla \phi_i) + \sum_{j=1}^D \bar{\beta}_{ji} \cdot \nabla \phi_j - \sum_{j=1}^D \bar{f}_{ij} \phi_j = \psi_i, \\ \quad (x, t) \in \Omega \times (t_n, 0], \quad 1 \leq i \leq D, \\ \phi_i(x, t) = 0, \quad (x, t) \in \partial\Omega \times (t_N, 0], \quad 1 \leq i \leq d, \\ \phi(x, t_N) = 0, \quad x \in \Omega, \end{cases}$$

# *A Posteriori* Analysis for a Reaction-Diffusion System

## Theorem

$$\begin{aligned} \int_0^T (e, \psi) dt &= (u_0 - P_0 u_0, \phi(0)) \\ &+ \sum_{n=1}^N \int_{I_n} (U', \pi_k \pi_h \phi - \pi_h \phi) + (\epsilon(U) \nabla U, \nabla(\pi_k \pi_h \phi - \pi_h \phi)) \\ &\quad - (f(U), \pi_k \pi_h \phi - \pi_h \phi) dt \\ &+ \sum_{n=1}^N \int_{I_n} (U', \pi_h \phi - \phi) + (\epsilon(U) \nabla U, \nabla(\pi_h \phi - \phi)) \\ &\quad - (f(U), \pi_h \phi - \phi) dt \end{aligned}$$

## *Forward Error Propagation*

(4)

# Forward Error Propagation

We briefly describe the classic alternative for error estimation

For the initial value problem

$$y' = f(y)$$

write the equation for a numerical solution  $Y \approx y$

$$Y' = F(Y)$$

where

$$F(\cdot) \approx f(\cdot)$$

# Forward Error Propagation

Subtraction yields an equation for the error  $e = y - Y$

$$e' = f(y) - F(Y) = f(y) - f(Y) + f(Y) - F(Y)$$

Linearizing

$$e' \approx f'(Y)e + \mathcal{R}(Y)$$

with the **defect**  $\mathcal{R}(Y) = f(Y) - F(Y)$

We can solve the system

$$\begin{cases} Y' = F(Y), \\ e' = f'(Y)e + \mathcal{R}(Y) \end{cases}$$

to compute the numerical solution and an approximation to the error simultaneously

# Forward Error Propagation

Technical issues that must be addressed:

- ▶ The same issue regarding the choice of linearization point that affects the adjoint-based approach
- ▶ The defect must be interpreted in a correct mathematical sense
- ▶ The defect can be approximated in multiple ways without a clearly best choice

Computational evaluation of the two approaches:

- ▶ The forward propagation is cheaper when the goal is to estimate the error in the solution itself
- ▶ The adjoint-based method is cheaper when the goal is to estimate the error in a handful of quantity of interest functionals

# My Preference

I prefer the adjoint-based approach because

- ▶ I have rarely encountered situations in which error in the solution is important
- ▶ The dual problem provides quantitative information about stability
- ▶ It seems easier to quantify relative contributions of different sources of error
- ▶ It seems easier to adapt the method to deal with such things as multiphysics, finite iteration, quadrature, etc.

A lot of nonsense is written about relative merits of the two approaches

## *Adjoints and Stability*

(8)

# The Adjoint and Stability

The solution of the adjoint problem scales local perturbations to global effects on a solution

The adjoint problem carries stability information about the quantity of interest computed from the solution

We can use the adjoint problem to investigate stability

# Condition Numbers and Stability Factors

The classic error bound for an approximate solution  $U$  of  $\mathbf{A}u = b$  is

$$\|e\| \leq C\kappa(\mathbf{A})\|R\|, \quad R = \mathbf{A}U - b$$

The **condition number**  $\kappa(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\|$  measures stability

$$\kappa(\mathbf{A}) = \frac{1}{\text{distance from } \mathbf{A} \text{ to } \{\text{singular matrices}\}}$$

The *a posteriori* estimate  $|(e, \psi)| = |(R, \phi)|$  yields

$$|(e, \psi)| \leq \|\phi\| \|R\|$$

The **stability factor**  $\|\phi\|$  is a **weak** condition number for the quantity of interest

We can obtain  $\kappa$  from  $\|\phi\|$  by taking the sup over all  $\|\psi\| = 1$

# Condition Numbers and Stability Factors

We consider the problem of computing  $(u, e_1)$  from the solution of

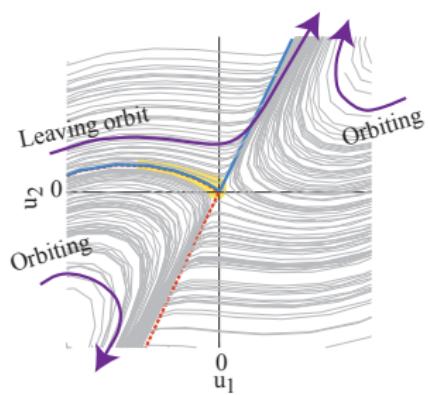
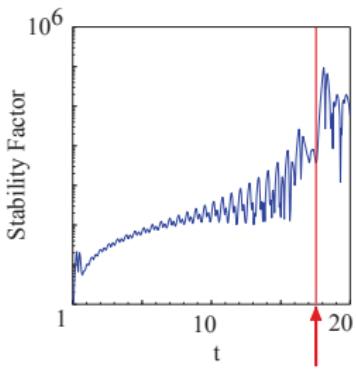
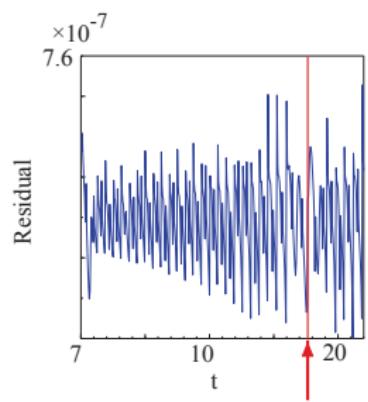
$$\mathbf{A}u = b$$

where  $\mathbf{A}$  is a random  $800 \times 800$  matrix

The condition number is  $1.7 \times 10^5$

The stability factor is 16

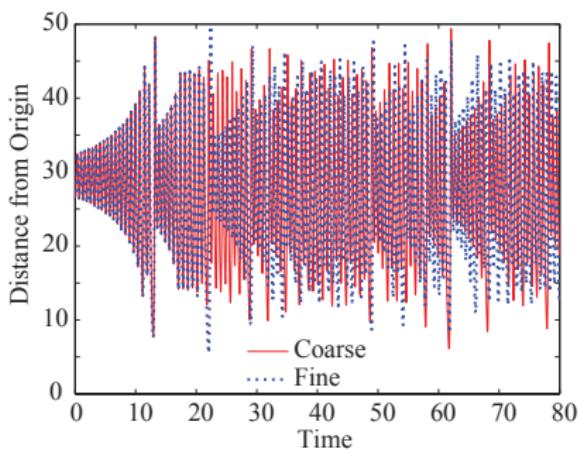
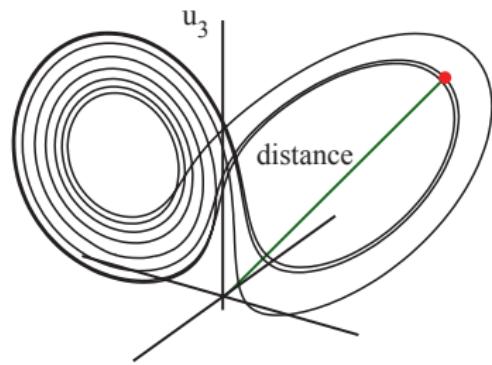
# The Condition of the Lorenz Problem



# The Quantity of Interest is Important

The rate that errors grow depends strongly on the information being computed

We consider the **average** distance from a solution to the origin over a long time interval



# The Quantity of Interest is Important

We compare to an ensemble average of 100 accurate solutions computed using time step .0001 for 15 time units

End Time	Coarse Solutions		Fine Solutions		Ensemble Ave	
	Ave	Var	Ave	Var	Ave	Var
20	27.620	52.011	27.622	51.947		
80	26.470	79.461	26.467	79.231		
320	26.3	83.7	26.3	83.0	26.3	83.7

# The Bistable Problem

A parabolic problem:

$$\begin{cases} u' - \epsilon \Delta u = u - u^3, & x \in \Omega, t > 0 \\ \text{Neumann boundary conditions} & x \in \partial\Omega, t > 0 \\ u(x, 0) = u_0(x), & x \in \Omega \end{cases}$$

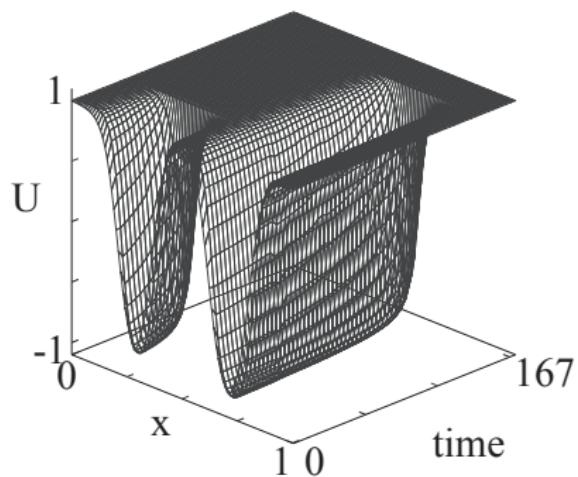
Long time dynamics depends on space dimension

- 1D Metastable behavior of long periods of nearly motionless behavior punctuated by rapid transients
- 2D Approximate motion by mean curvature

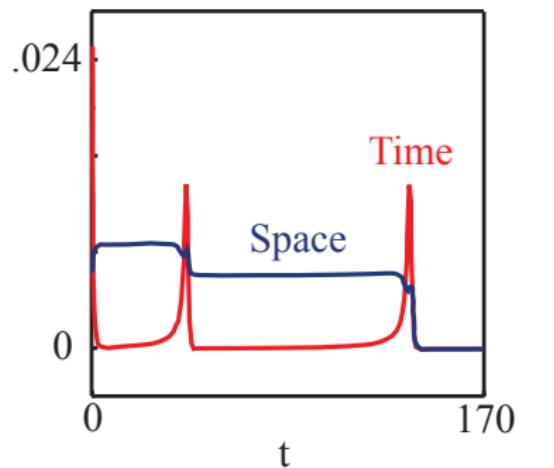
The solutions in two dimensions are much more stable with respect to perturbations

# The Bistable Problem

Example of metastable solution in one space dimension



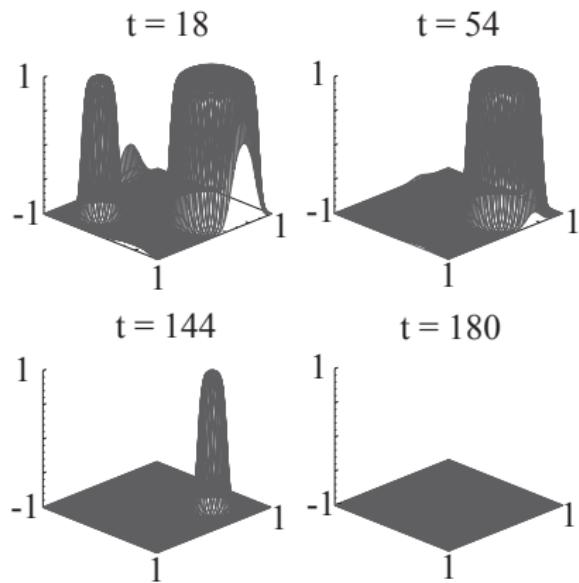
Evolution of Solution



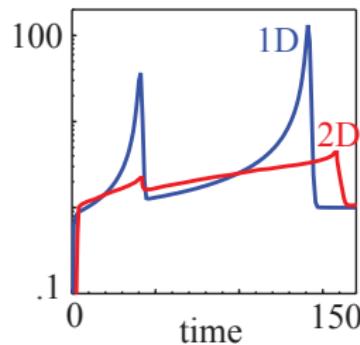
Norms of Adjoint Components

# The Bistable Problem

Example of solution in two space dimensions



Evolution of Solution



Norms of Adjoint Solutions

## *Adjoints and Adaptive Error Control*

(7)

# Idea Behind Adaptive Error Control

The possibility of accurate error estimation suggests the possibility of optimizing discretizations

Unfortunately, cancellation of errors significantly complicates the optimization problem

In fact, there is no good theory for adaptive control of error

There is good theory for adaptive control of error bounds

The standard approach is based on optimal control theory

The stability information in adjoint-based *a posteriori* error estimates is useful for this

# Bounds and Estimates

We consider the problem of computing  $(u, e_1)$  from the solution of

$$\mathbf{A}u = b$$

where  $\mathbf{A}$  is a random  $800 \times 800$  matrix

The condition number of  $\mathbf{A}$  is  $1.7 \times 10^5$

estimate of the error in the quantity of interest  $\approx 1.4 \times 10^{-14}$   
*a posteriori* error bound for the quantity of interest  $\approx 2.2 \times 10^{-10}$

The traditional error bound for the error  $\approx 5.7 \times 10^{-5}$

# Adjoints and Adaptive Error Control

An abstract *a posteriori* error estimate has the form

$$|(e, \psi)| = |(Residual, Adjoint\ Weight)|$$

Given a tolerance TOL, a given discretization is refined if

$$|(Residual, Adjoint\ Weight)| \geq TOL$$

Refinement decisions are based on a bound consisting of a sum of **element contributions**

$$|(e, \psi)| \leq \sum_{elements\ K} |(Residual, Adjoint\ Weight)_K|$$

where  $(\ , \ )_K$  is the inner product on  $K$

The element contributions in the bound do not cancel

# Adjoints and Adaptive Error Control

There is **no cancellation of errors across elements** in the bound, so optimization theory yields

Principle of Equidistribution: The optimal discretization is one in which the element contributions are equal

The **adaptive strategy** is to refine some of the elements with the largest element contributions

The adjoint weighted residual approach is different than traditional approaches because the element residuals are scaled by an adjoint weight, which measures how much error in that element affects the solution on other elements

An optimal solution is approximated iteratively starting with a coarse mesh

# Adjoints and Adaptive Error Control

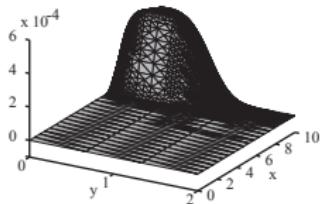
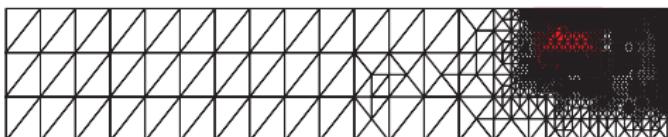
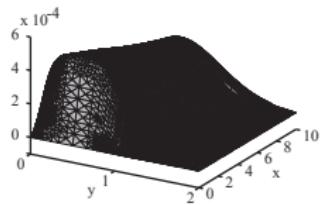
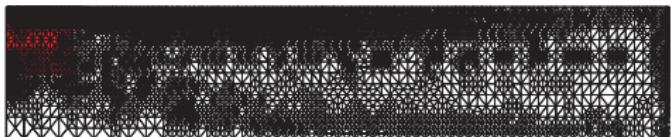
- 1: Choose error tolerance TOL and compute contribution tolerances
- 2: Choose initial coarse discretization parameters
- 3: Compute initial coarse solution
- 4: Compute error estimate for initial solution
- 5: **while** Error Estimate > TOL **do**
- 6:     Adjust discretization parameters depending on the contributions relative to the contribution tolerances
- 7:     Compute solution
- 8:     Compute error estimate for solution
- 9: **end while**

# Adjoints and Adaptive Error Control

$$\begin{cases} -\nabla \cdot ((.05 + \tanh(10(x-5)^2 + 10(y-1)^2)) \nabla u) \\ \quad + \begin{pmatrix} -100 \\ 0 \end{pmatrix} \cdot \nabla u = 1, & (x, y) \in \Omega = [0, 10] \times [0, 2], \\ u = 0, & (x, y) \in \partial\Omega \end{cases}$$



# Adjoints and Adaptive Error Control



Final meshes for an average error of 4%  
24,000 elements versus 3500 elements



## *Treatment of Multiphysics Problems*

(15)

# Multiphysics, Multiscale Systems

Multiphysics, multiscale systems couple different physical processes interacting across a wide range of scales

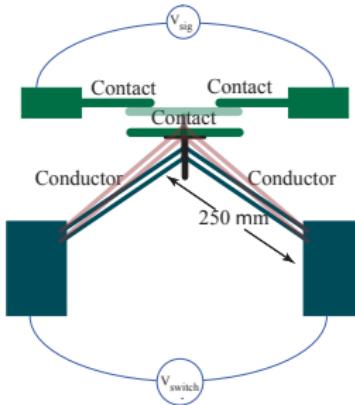
Such systems abound in science and engineering application domains

Computational modeling plays a critical role in the predictive study of such systems

Some applications where multiphysics systems arise:

- ▶ Fusion and fission reactors
- ▶ Reacting fluids and fluid-solid interactions
- ▶ Advanced materials, nano-manufacturing
- ▶ Biological systems, drug design and delivery
- ▶ Environmental, climate, ecological models
- ▶ Weather models

# Example: A MEMs Thermal Actuator



$$\nabla \cdot (\sigma(\mathbf{d}) \nabla \mathbf{V}) = 0$$

electrostatic current

$$\nabla \cdot (\kappa(\mathbf{T}) \nabla \mathbf{T}) = \sigma(\nabla \mathbf{V} \cdot \nabla \mathbf{V})$$

energy

$$\hat{\nabla} \cdot (\lambda \text{tr}(E) I + 2\mu E - \beta(\mathbf{T} - T_{ref})I) = 0$$

displacement

$$E = (\hat{\nabla} \mathbf{d} + \hat{\nabla} \mathbf{d}^\top)/2$$

# Solution of Multiscale, Multiphysics Systems

Multiscale, multiphysics systems pose severe challenges for computational solution:

- ▶ Scale effects
- ▶ Complex stability
- ▶ Coupling between physics
- ▶ Different kinds of physical descriptions
- ▶ Mixtures of discretization methods and scales
- ▶ Complex HPC discretization techniques

Significant extensions of the a posteriori theory are required

# Solution of a Coupled Elliptic System

A simplified Thermal Actuator model:

$$\begin{cases} -\nabla \cdot a_1 \nabla u_1 + \mathbf{b}_1 \cdot \nabla u_1 + c_1 u_1 = f_1(x), & x \in \Omega, \\ -\nabla \cdot a_2 \nabla u_2 + \mathbf{b}_2 \cdot \nabla u_2 + c_2 u_2 = f_2(x, u_1, Du_1), & x \in \Omega, \\ u_1 = u_2 = 0, & x \in \partial\Omega, \end{cases}$$

$\Omega$  is a bounded domain with boundary  $\partial\Omega$

The coefficients are smooth functions and  $a_1, a_2$  are bounded away from zero

We wish to compute information depending on  $u_2$

# Solution of a Coupled Elliptic System

## Algorithm

- ▶ Construct discretizations  $\mathcal{T}_{h,1}, \mathcal{T}_{h,2}$  and finite element spaces  $V_{h,1}, V_{h,2}$
- ▶ Compute a finite element solution  $U_1 \in V_{h,1}(\Omega)$  of the first equation
- ▶ Project  $U_1 \in V_{h,1}(\Omega)$  into the space  $V_{h,2}(\Omega)$
- ▶ Compute a finite element solution  $U_2 \in V_{h,2}(\Omega)$  of the second equation

The projection between the discretization spaces is a crucial step

In a fully coupled system,  $U_2$  is projected into  $V_{h,1}(\Omega)$  for the next iteration

# A Simple Thermal Actuator

Consider

$$\begin{cases} -\Delta \textcolor{brown}{u}_1 = \sin(4\pi x) \sin(\pi y), & x \in \Omega \\ -\Delta u_2 = b \cdot \nabla \textcolor{brown}{u}_1 = 0, & x \in \Omega, \\ u_1 = u_2 = 0, & x \in \partial\Omega, \end{cases} \quad b = \frac{2}{\pi} \begin{pmatrix} \sin(4\pi x) \\ 25 \sin(\pi y) \end{pmatrix}$$

where  $\Omega = [0, 1] \times [0, 1]$

We consider the quantity of interest

$$u_2(.25, .25)$$

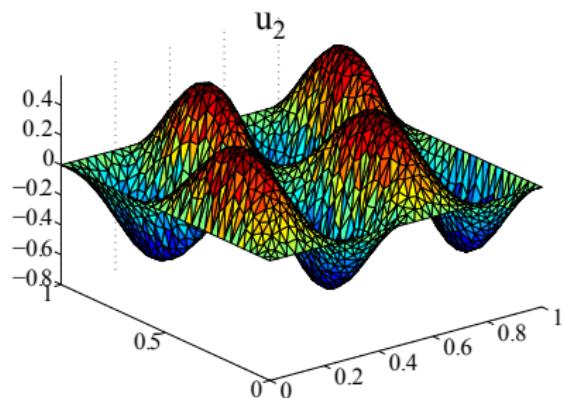
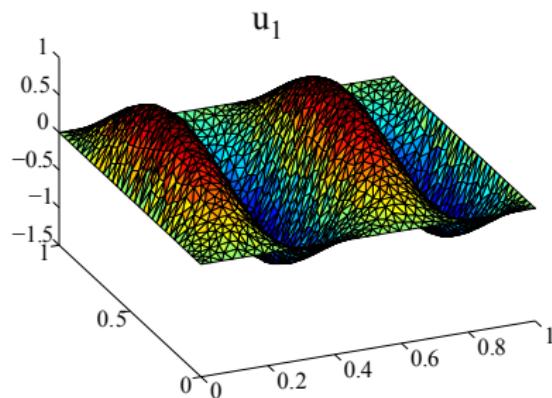
We solve for  $u_1$  first and then solve for  $u_2$  using **independent meshes**

# A Simple Thermal Actuator

Using uniform meshes, an a posteriori error estimate yields  
estimate of the error in the quantity of interest  $\approx .0042$

true error  $\approx .0048$

discrepancy in estimate  $\approx .0006$  ( $\approx 13\%$ )



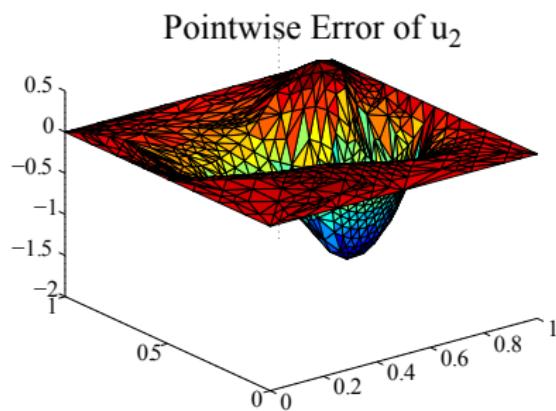
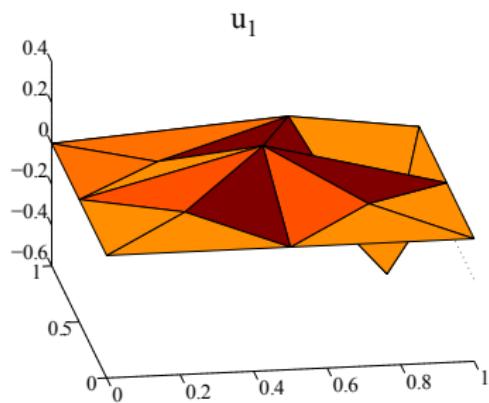
This arises from the operator decomposition

# A Simple Thermal Actuator

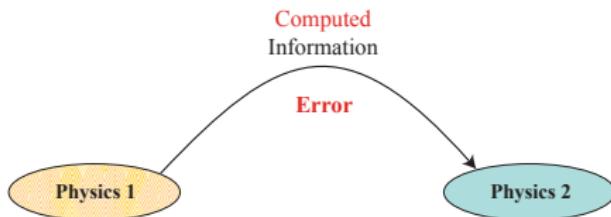
Adapting the mesh using **only** an error estimate for the second component causes the discrepancy to become alarmingly worse

estimate of the error in the quantity of interest  $\approx .0001$

true error  $\approx .2244$



# Analysis for Multiscale, Multiphysics Solutions



We define **auxiliary quantities of interest** corresponding to information passed between components

We solve auxiliary adjoint problems to estimate the error in that information

In an iterative scheme, we also estimate the “history” of errors passed from one iteration level to the next

We can also estimate the effect of processing, e.g. up and down scaling, the information

# Analysis for Multiscale, Multiphysics Solutions

Recall the “triangular” problem

$$\begin{cases} -\Delta \textcolor{red}{u}_1 = \sin(4\pi x) \sin(\pi y), & x \in \Omega \\ -\Delta u_2 = b \cdot \nabla \textcolor{red}{u}_1 = 0, & x \in \Omega, \\ u_1 = u_2 = 0, & x \in \partial\Omega, \end{cases} \quad b = \frac{2}{\pi} \begin{pmatrix} 25 \sin(4\pi x) \\ \sin(\pi x) \end{pmatrix}$$

where  $\Omega = [0, 1] \times [0, 1]$

We consider the quantity of interest

$$u_2(.25, .25)$$

We solve for  $u_1$  first and then solve for  $u_2$  using **independent meshes**

# Analysis for Multiscale, Multiphysics Solutions

$$\begin{cases} -\Delta u_1 = f_1(x), & x \in \Omega, \\ -\Delta u_2 = f_2(x, u_1, Du_1), & x \in \Omega \\ u_1 = 0, \quad u_2 = 0, & x \in \partial\Omega \end{cases}$$

We compute a **quantity of interest**  $(\psi^{(1)}, u)$  that only depends on  $u_2$ , thus

$$\psi^{(1)} = \begin{pmatrix} 0 \\ \psi_2^{(1)} \end{pmatrix}$$

We require the **linearization**  $Lf_2(w)$  of  $f_2$  with respect to  $u_1$  around the function  $w$

# Analysis for Multiscale, Multiphysics Solutions

The weak form of the **primary** adjoint problem is

$$\begin{cases} (\nabla \phi_1^{(1)}, \nabla v_1) + (Lf_2(U_1)\phi_2^{(1)}, v_1) = 0, \\ (\nabla \phi_2^{(1)}, \nabla v_2) = (\psi_2^{(1)}, v_2), \end{cases}$$

all test functions  $v_1, v_2$

This yields the first error representation

$$(\psi^{(1)}, e) = (f_2(u_1, Du_1), (I - \pi_h)\phi_2^{(1)}) - (\nabla U_2, \nabla(I - \pi_h)\phi_2^{(1)})$$

The residual depends on the unknown true solution

We write

$$f_2(u_1, Du_1) = f_2(U_2, DU_2) + (f_2(u_1, Du_1) - f_2(U_2, DU_2))$$

# Analysis for Multiscale, Multiphysics Solutions

We have a **new** linear functional of the error

$$\begin{aligned}(f_2(u_1, Du_1) - f_2(U_1, DU_1), \phi_2^{(1)}) &\approx (Lf(U_1)e_1, \phi_2^{(1)}) \\ &= (Lf(U_1)^* \phi_2^{(1)}, e_1) = (\psi^{(2)}, e)\end{aligned}$$

We pose a **secondary** adjoint problem

$$\begin{cases} (\nabla \phi_1^{(2)}, \nabla v_1) + (Lf_2(U_1) \phi_2^{(2)}, v_1) = (Lf_2(U_1)^* \phi_2^{(1)}, v_1) \\ (\nabla \phi_2^{(2)}, \nabla v_2) = 0 \end{cases}$$

all test functions  $v_1, v_2$

# Analysis for Multiscale, Multiphysics Solutions

## Theorem

$$\begin{aligned} (\psi^{(1)}, e) = & \left( f_2(U_1, DU_1), (I - \pi_h)\phi_2^{(1)} \right) - \left( \nabla U_2, \nabla(I - \pi_h)\phi_2^{(1)} \right) \\ & + \left( f_1, (I - \pi_h)\phi_1^{(2)} \right) - \left( \nabla U_1, \nabla(I - \pi_h)\phi_1^{(2)} \right) \end{aligned}$$

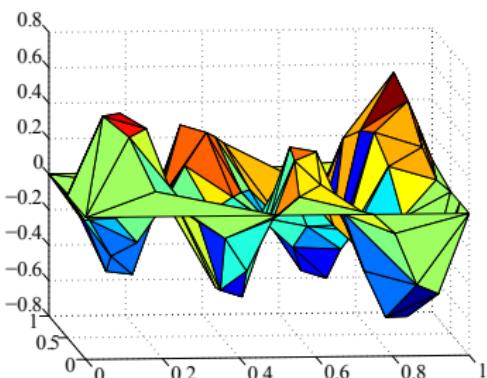
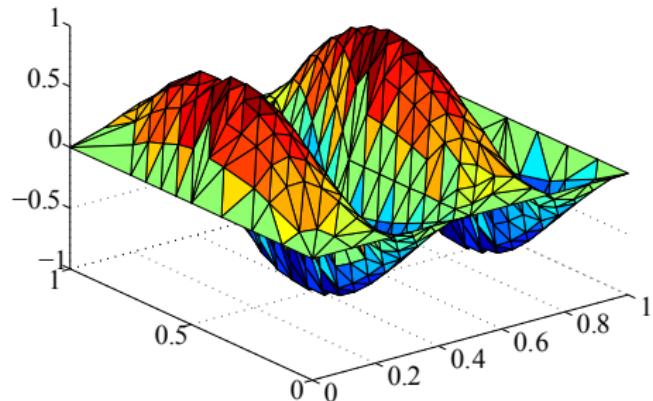
$\phi^{(1)}, \phi^{(2)}$  are the primary and secondary adjoint solutions

The new analysis estimates the error in the information passed between components

There is a tertiary adjoint problem and another term in the estimate if the different discretizations are used for the two components

# A Simple Thermal Actuator

If we adapt the meshes using all the terms in the estimate, we can drive the error below .0001



We actually refine the mesh for  $u_1$  more than the mesh for  $u_2$

## *Extending to Other Discretizations*

(16)

# Finite Elements and Other Discretizations

The a posteriori error analysis applies to wide classes of discretization

The key is to write a chosen discretization method as some kind of finite element method with modifications such as quadrature to evaluate integrals

Examples of extensions include

- ▶ Finite difference schemes
- ▶ Finite volume methods
- ▶ Explicit and IMEX time integration methods
- ▶ Operator split discretizations

Extending the analysis always involves identifying new residuals and may involve defining additional adjoint problems

# Finite Difference Schemes for an Initial Value Problem

Let  $Q_r(g, n)$  denote a quadrature formula of order  $r$  for the function  $g$  on  $I_n$

CG(q) method with quadrature: For  $n = 1, \dots, N$ ,

$$\begin{cases} \int_{I_n}(Y', v) dt = Q_r((f(Y), n), v), n \\ Y_{n-1}^+ = Y_{n-1}^- \end{cases} \quad \text{all } v \in W_n^{q-1}$$

with  $Y_0 = y_0$ ,  $U_n^+ = \lim_{t \downarrow t_n} U(t)$ ,  $U_n^- = \lim_{t \uparrow t_n} U(t)$

This yields a system of discrete equations determining  $Y$

The discrete equations yield a finite difference scheme for  $Y$

# Finite Difference Schemes for an Initial Value Problem

Consider  $q = 1$  and the trapezoidal rule quadrature

On  $I_n$ ,

$$Y = Y_n \frac{t - t_{n-1}}{k_n} + Y_{n-1} \frac{t_n - t}{k_n}$$
$$Q_1(g, n) = \frac{1}{2}(g(t_n) + g(t_{n-1}))k_n$$

The equation for the unknown coefficient  $Y_n$  is

$$Y_n - \frac{1}{2}f(Y_n)k_n = Y_{n-1} + \frac{1}{2}f(Y_{n-1})k_n$$

# Treatment of Quadrature in A Posteriori Analysis

Key observation: Galerkin orthogonality uses  $Q_r((f(Y), n), v), n$  instead of exact integration

We add and subtract  $Q_r((f(Y), n), v), n$  in the argument

Theorem The error representation is

$$\begin{aligned}\int_0^T (e, \psi) dt &= \sum_{n=1}^N \int_{I_n} (\mathcal{R}_n, \pi_k \phi - \phi) dt \\ &\quad + \sum_{n=1}^N \left( \int_{I_n} (f(Y), \phi) dt - Q_r((f(Y), n), \phi), n \right)\end{aligned}$$

$$\mathcal{R}_n = Y' - f(Y)$$

# Treatment of Quadrature in A Posteriori Analysis

Observations about the contribution to the error from quadrature

$$\left( \int_{I_n} (f(Y), \phi) dt - Q_r((f(Y), n), \phi), n \right)$$

- ▶ The stability factor is  $\phi$  not  $\pi_k \phi - \phi$ , so this contribution accumulates at a different rate in general
- ▶ Exactly evaluating this expression requires exact integration of the nonlinear term, and this has to be approximated in general

Generically, such estimates include terms that cannot be evaluated exactly

# Operator Splitting for Reaction-Diffusion Equations

Estimation of the effects of methods such as quadrature is important

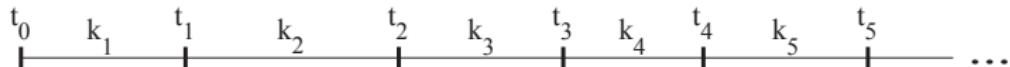
We consider operator splitting for a reaction-diffusion problem

$$\begin{cases} \frac{du}{dt} = \Delta u + F(u), & 0 < t, \\ u(0) = u_0 \end{cases}$$

The diffusion component  $\Delta u$  induces **stability** and change over **long time scales**

The reaction component  $F$  induces **instability** and change over **short time scales**

# Operator Splitting for Reaction-Diffusion Equations



On  $(t_{n-1}, t_n]$ , we numerically solve

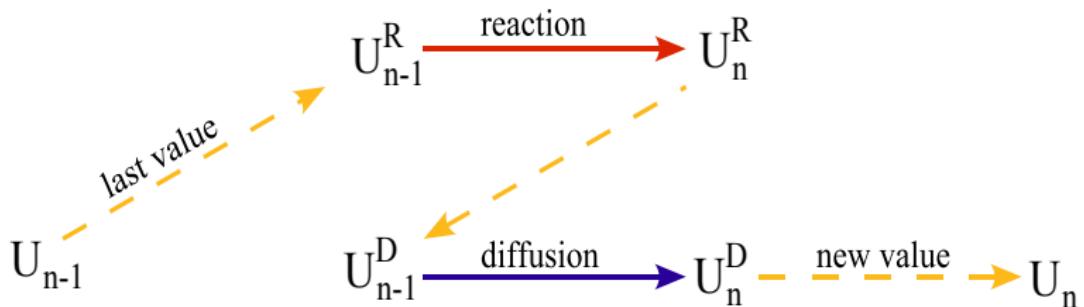
$$\begin{cases} \frac{du^R}{dt} = F(u^R), & t_{n-1} < t \leq t_n, \\ u^R(t_{n-1}) = u^D(t_{n-1}) \end{cases}$$

On  $(t_{n-1}, t_n]$ , we numerically solve

$$\begin{cases} \frac{du^D}{dt} - \Delta(u^D), & t_{n-1} < t \leq t_n, \\ u^D(t_{n-1}) = u^R(t_n) \end{cases}$$

The operator split approximation is  $u(t_n) \approx u^D(t_n)$

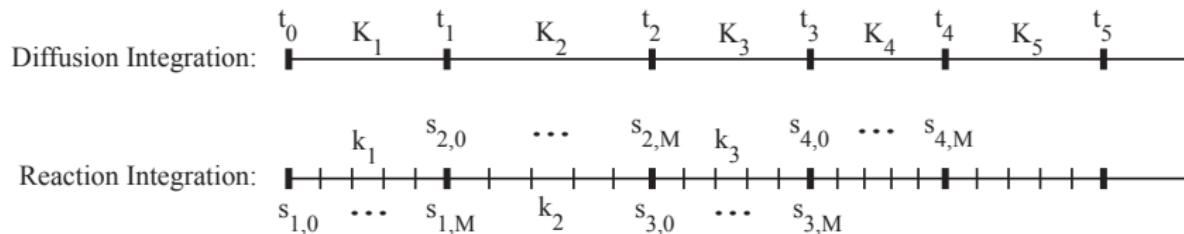
# Operator Splitting for Reaction-Diffusion Equations



# Operator Splitting for Reaction-Diffusion Equations

We discretize using a finite element in time

To account for the fast reaction, we approximate  $u^r$  using **many** time steps inside each diffusion step



# Operator Splitting for Reaction-Diffusion Equations

Decompose the error using an exact operator split solution  $u^s$ :

$$u - U = (u - u^s) + (u^s - U)$$

- ▶  $u - u^s$ : error of analytic operator splitting

This is determined by properties of adjoint operators

- ▶  $u^s - U$ : error of numerical component solves

This is determined by the numerical errors made in each component

Estimates of the numerical error in each component do not indicate the effects of splitting in general

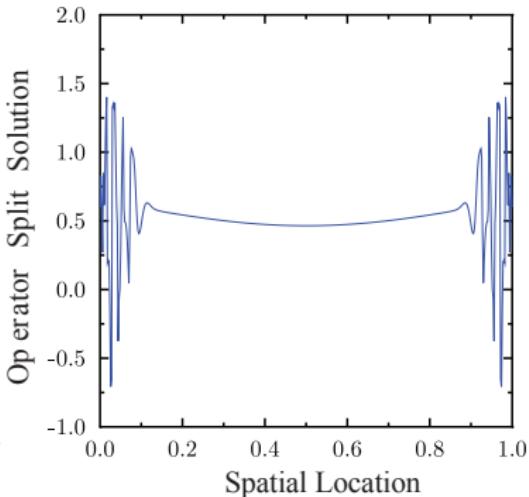
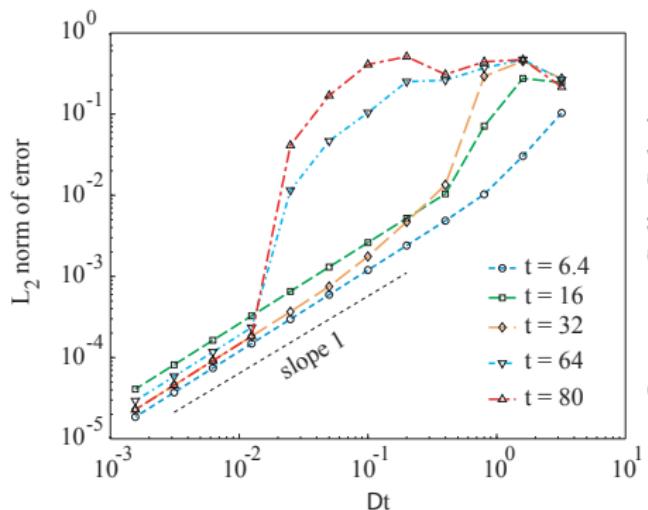
# Operator Splitting for Reaction-Diffusion Equations

## The Brusselator problem

$$\begin{cases} \frac{\partial u_i}{\partial t} - 0.025 \frac{\partial^2 u_i}{\partial x^2} = f_i(u_1, u_2) & i = 1, 2 \\ f_1(u_1, u_2) = 0.6 - 2u_1 + u_1^2 u_2 \\ f_2(u_1, u_2) = 2u_1 - u_1^2 u_2 \end{cases}$$

- ▶ Use a linear finite element method in space with 500 elements
- ▶ Use a standard first order splitting scheme
- ▶ Use Trapezoidal Rule with time step of .2 for the diffusion and Backward Euler with time step of .004 for the reaction

# Operator Splitting for Reaction-Diffusion Equations



## Instability in the Brusselator Operator Splitting

*For large times, there is a critical step size above which there is no convergence. The instability is a direct consequence of the operator splitting*

# Operator Splitting for Reaction-Diffusion Equations

Operator splitting complicates the definition of adjoint operators

- ▶ In a linear problem, the adjoint operators associated with the original problem and an operator split discretization are fundamentally different
- ▶ Additionally in a nonlinear problem, the issue of choice of linearization point becomes complicated

Estimates of the effects on adjoints cannot be entirely computable

# Operator Splitting for Reaction-Diffusion Equations

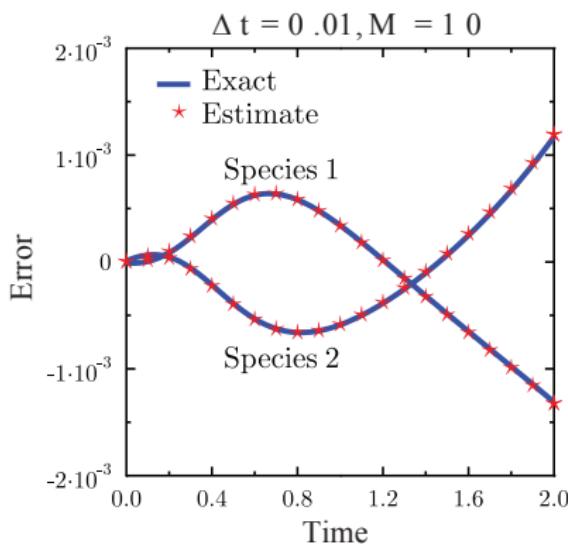
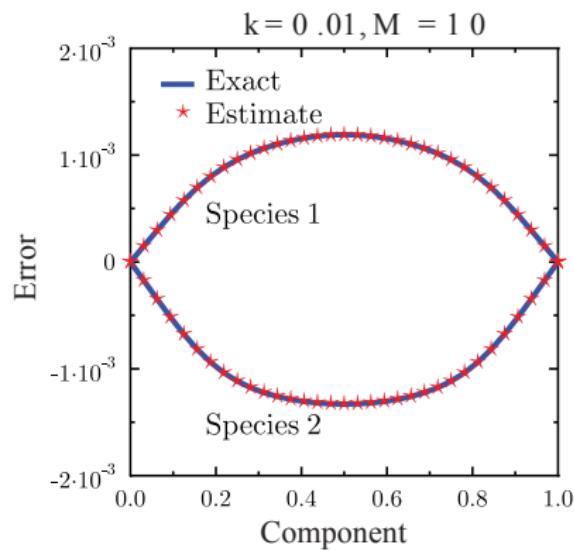
We derive a hybrid *a priori - a posteriori* estimate

$$(e(t_N), \psi) = \mathcal{Q}_1 + \mathcal{Q}_2 + \mathcal{Q}_3$$

- ▶  $\mathcal{Q}_1$  estimates the contribution of the numerical solution of each component in an standard a posteriori way
- ▶  $\mathcal{Q}_2 \approx \sum_{n=1}^N (U_{n-1}, E_{n-1})$ ,  $E \approx$  a computable contribution for the error in the adjoint arising from operator splitting
- ▶  $\mathcal{Q}_3$  is a noncomputable *a priori* expression that is provably higher order

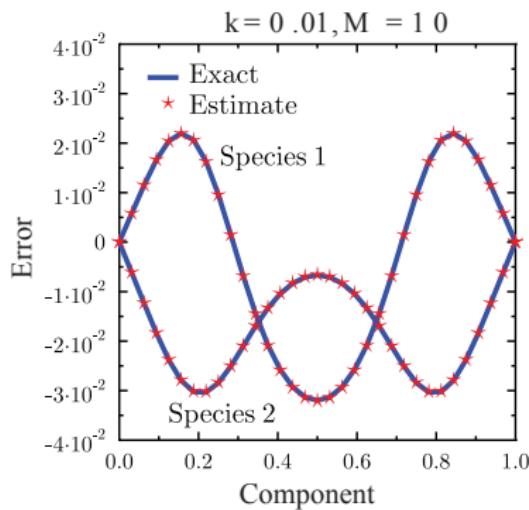
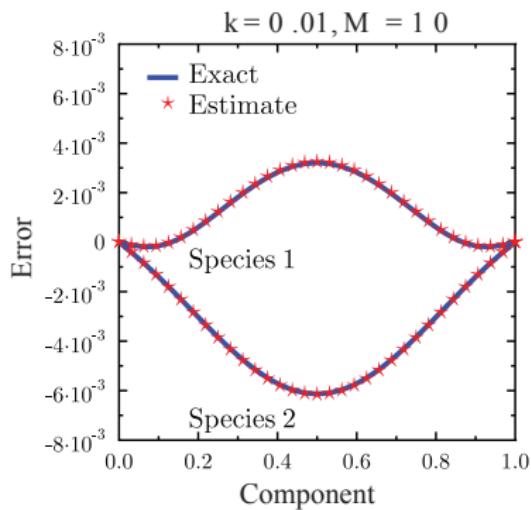
# Operator Splitting for Reaction-Diffusion Equations

Accuracy of the error estimate for the Brusselator example over  $[0, 2]$



# Operator Splitting for Reaction-Diffusion Equations

Accuracy of the error estimate for the Brusselator example at  $T = 8$  and  $T = 40$



## *Conclusion*

# Conclusion

A posteriori error analysis based on duality, variational analysis, and adjoint equations provides a functional analytic approach to accurate computational error estimation

Deriving the estimates and quantifying sources of discretization errors provides significant insight into the behavior of the numerical scheme

Consideration of the adjoint problem provides insight into the stability, e.g. accumulation, propagation, and transmission of error

Having accurate estimates not only is a key aspect of uncertainty quantification, but it also is useful for computation in several different ways

## *Limited References*

# Limited References

## Duality and adjoints:

- ▶ Linear Differential Operators, C. Lanczos, SIAM, 1996
- ▶ Applied Functional Analysis, E. Zeidler, Springer, 1995
- ▶ Adjoint Equations and Perturbation Algorithms in Nonlinear Problems, G. Marchuk, V. Agoshkov, and V. Shutyaev, CRC Press, 1996
- ▶ Real Analysis, G. Folland, J. Wiley and Sons, 1999
- ▶ Applied Functional Analysis, J. Aubin, J. Wiley and Sons, 2000
- ▶ Analysis for Applied Mathematics, W. Cheney, Springer, 2001

## General a posteriori error analysis and adaptive error control:

- ▶ Global error control for the continuous Galerkin finite element method for ordinary differential equations, D. Estep and D. French, RAIRO MMAN, 1994
- ▶ A posteriori error bounds and global error control for approximations of ordinary differential equations, D. Estep, SIAM. J. Numer. Analysis, 1995
- ▶ Introduction to adaptive methods for differential equations, K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, Acta Numerica, 1995

## Limited References

General a posteriori error analysis and adaptive error control:

- ▶ Computational Differential Equations, K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, Cambridge University Press, 1996
- ▶ Accurate parallel integration of large sparse systems of differential equations, D. Estep and R. Williams, MMMAS, 1996
- ▶ Estimating the error of numerical solutions of systems of nonlinear reaction-diffusion equations, D. Estep, M. Larson and R. Williams, Memoirs of the AMS, 2000
- ▶ An optimal control approach to a posteriori error estimation in finite element methods, R. Becker and R. Rannacher, Acta Numerica, 2001
- ▶ Adjoint methods for PDEs: a posteriori error analysis and post-processing by duality, M. Giles and E. Suli, Acta Numerica, 2002
- ▶ Accounting for stability: a posteriori estimates based on residuals and variational analysis, D. Estep, M. Holst, D. Mikulencak, Communications in Numerical Methods in Engineering, 2002
- ▶ Adaptive finite element methods for differential equations, W. Bangerth and R. Rannacher, Birkhauser, 2003

# Limited References

General a posteriori error analysis and adaptive error control:

- ▶ Generalized Green's functions and the effective domain of influence, D. Estep, M. Holst, and M. Larson, SIAM J. Sci. Comput., 2005
- ▶ A posteriori error estimation of approximate boundary fluxes, T. Wildey, S. Tavener, and D. Estep, Comm. Num. Meth. Engin., 24 (2008)
- ▶ A posteriori error analysis of a cell-centered finite volume method for semilinear elliptic problems, D. Estep, M. Pernice, D. Pham, S. Tavener, H. Wang, J. Computational and Applied Mathematics, 2009
- ▶ Bridging the Scales in Science and Engineering, Editor J. Fish, Oxford University Press, Chapter 11: Error Estimation for Multiscale Operator Decomposition for Multiphysics Problems, D. Estep, 2010
- ▶ Blockwise adaptivity for time dependent problems based on coarse scale adjoint solutions, V. Carey, D. Estep, A. Johansson, M. Larson, and S. Tavener, SIAM J. Sci. Comput., 2010
- ▶ A posteriori error analysis of two stage computation methods with application to efficient resource allocation and the Parareal Algorithm, J. H. Chaudhry, D. Estep, S. Tavener, V. Carey, and J . Sandelin, SIAM J. Numerical Analysis, 2014, submitted

# Limited References

## Our work on multiphysics problems:

- ▶ A posteriori - a priori analysis of multiscale operator splitting, D. Estep, V. Ginting, D. Ropp, J. Shadid, and S. Tavener, SINUM, 46 (2008)
- ▶ A posteriori analysis and improved accuracy for an operator decomposition solution of a conjugate heat transfer problem, D. Estep, S. Tavener, T. Wildey, SINUM, 46 (2008)
- ▶ A posteriori error analysis of multiscale operator decomposition methods for multiphysics models, D. Estep, V. Carey, V. Ginting, S. Tavener, T. Wildey, J. Physics: Conf. Ser. 125 (2008)
- ▶ A posteriori analysis and adaptive error control for multiscale operator decomposition methods for coupled elliptic systems I: One way coupled systems, V. Carey, D. Estep, and S. Tavener, SINUM 47 (2009)
- ▶ A posteriori error analysis for a transient conjugate heat transfer problem, D. Estep, S. Tavener, T. Wildey, Fin. Elem. Analysis Design 45 (2009)
- ▶ Bridging the Scales in Science and Engineering, Editor J. Fish, Oxford University Press, Chapter 11: Error Estimation for Multiscale Operator Decomposition for Multiphysics Problems, D. Estep, 2010

# Limited References

Our work on multiphysics problems:

- ▶ A posteriori error estimation and adaptive mesh refinement for a multi-discretization operator decomposition approach to fluid-solid heat transfer, D. Estep, S. Tavener, T. Wildey, JCP, 2010
- ▶ A posteriori analysis of multirate numerical methods for ordinary differential equations, SINUM, D. Estep, V. Ginting, S. Tavener, Comput. Meth. Appl. Mech. Eng., 2012
- ▶ A posteriori analysis and adaptive error control for operator decomposition solution of coupled semilinear elliptic systems, V. Carey, D. Estep, S. Tavener, Int. J. Numer. Meth. Engin., 2013
- ▶ A posteriori analysis of an iterative multi-discretization method for reaction-diffusion systems, D. Estep, V. Ginting, J. Hameed, and S. Tavener, Comp. Meth. Appl. Mech. Engin., 2013
- ▶ A-posteriori error estimates for mixed finite element and finite volume methods for problems coupled through a boundary with non-matching grids, T. Arbogast, D. Estep, B. Sheehan, and S. Tavener, IMA J. Numer. Analysis, 2013

# Limited References

## Our work on multiphysics problems:

- ▶ A posteriori error estimates for mixed finite element and finite volume methods for parabolic problems coupled through a boundary with non-matching discretizations, T. Arbogast, D. Estep, B. Sheehan, and S. Tavener, SIAM J. Uncertainty Quantification, 2015
- ▶ Adaptive finite element solution of multiscale PDE-ODE systems, A. Johansson, J. H. Chaudhry, V. Carey, D. Estep, V. Ginting, M. Larson, and S. Tavener, CMAME, 2015

## Treatment of various discretization “crimes”:

- ▶ A posteriori error analysis for a cut cell finite volume method, D. Estep, S. Tavener, M. Pernice and H. Wang, Computer Methods in Applied Mechanics and Engineering, 2011
- ▶ A posteriori error estimation for the Lax-Wendroff finite difference scheme, J. B. Collins, D. Estep, and S. Tavener, J. Comput. Appl. Math. 263C (2014)
- ▶ A posteriori error analysis for finite element methods with projection operators as applied to explicit time integration techniques, J. Collins, D. Estep and S. Tavener, BIT, 2014

## Limited References

Treatment of various discretization “crimes”:

- ▶ A posteriori error analysis of IMEX time integration schemes for advection-diffusion-reaction equations, J. Chaudry, D. Estep, V. Ginting, J. Shadid, and S. Tavener, CMAME, 2014
- ▶ A posteriori analysis for iterative solvers for non-autonomous evolution problems, J. H. Chaudry, D. Estep, V. Ginting, and S. Tavener, SIAM J. Uncertainty Quantification, 2015

Optimal control problems:

- ▶ Computational error estimation and adaptive mesh refinement for a finite element solution of launch vehicle trajectory problems, D. Estep, D. Hodges and M. Warner, SIAM Journal on Scientific Computing, 2000
- ▶ The solution of a launch vehicle trajectory problem by an adaptive finite element method, D. Estep, D. H. Hodges, M. Warner, Computer Methods in Applied Mechanics and Engineering, 2001

# Limited References

## Use of adjoints to investigate stability:

- ▶ The computability of the Lorenz system, D. Estep and C. Johnson, Mathematical Models and Methods in Applied Sciences, 1998
- ▶ Analysis of the sensitivity properties of a model of vector-borne bubonic plague, M. Buzby, D. Neckels, M. Antolin, and D. Estep, Royal Society Journal Interface, 5 (2008), 1099-1107

## A posteriori error analysis for uncertainty quantification:

- ▶ Fast and reliable methods for determining the evolution of uncertain parameters in differential equations, D. Estep and D. Neckels, J. Comput. Phys., 2006
- ▶ Fast methods for determining the evolution of uncertain parameters in reaction-diffusion equations, D. Estep and D. Neckels, Comput. Meth. Appl. Mech. Engin., 2007
- ▶ Nonparametric density estimation for randomly perturbed elliptic problems I: Computational methods, a posteriori analysis, and adaptive error control, D. Estep, A. Malqvist, and S. Tavener, SISC, 2009
- ▶ Nonparametric density estimation for randomly perturbed elliptic problems II: Applications and adaptive modeling, D. Estep, A. Malqvist, S. Tavener, Int. J. Numer. Meth. Engin., 2009

## Limited References

A posteriori error analysis for uncertainty quantification:

- ▶ A measure-theoretic computational method for inverse sensitivity problems I: Method and analysis, J. Breidt, T. Butler and D. Estep, SIAM J. Numer. Analysis, 2011
- ▶ A computational measure theoretic approach to inverse sensitivity problems II: A posteriori error analysis, T. Butler, D. Estep and J. Sandelin, SIAM J. Numerical Analysis, 2012
- ▶ Uncertainty quantification for approximate  $p$ -quantiles for physical models with uncertain inputs, D. Elfverson, D. Estep, F. Hellman and A. Malqvist, SIAM J. Uncertainty Quantification, 2014