



Published in final edited form as:

SIAM J Numer Anal. 2012 ; 50(1): 22–45. doi:10.1137/100785958.

A COMPUTATIONAL MEASURE THEORETIC APPROACH TO INVERSE SENSITIVITY PROBLEMS II: A POSTERIORI ERROR ANALYSIS*

T. BUTLER[†], D. ESTEP[‡], and J. SANDELIN[§]

[†]Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, TX 78712 (tbutler@ices.utexas.edu) [‡]Department of Statistics, Colorado State University, Fort Collins, CO 80523 (estep@stat.colostate.edu) [§]Department of Mathematics, Colorado State University, Fort Collins, CO 80523 (sandelin@math.colostate.edu)

Abstract

In part one of this paper [T. Butler and D. Estep, *SIAM J. Numer. Anal.*, to appear], we develop and analyze a numerical method to solve a probabilistic inverse sensitivity analysis problem for a smooth deterministic map assuming that the map can be evaluated exactly. In this paper, we treat the situation in which the output of the map is determined implicitly and is difficult and/or expensive to evaluate, e.g., requiring the solution of a differential equation, and hence the output of the map is approximated numerically. The main goal is an a posteriori error estimate that can be used to evaluate the accuracy of the computed distribution solving the inverse problem, taking into account all sources of statistical and numerical deterministic errors. We present a general analysis for the method and then apply the analysis to the case of a map determined by the solution of an initial value problem.

Keywords

a posteriori error analysis; adjoint problem; density estimation; inverse sensitivity analysis; nonparametric density estimation; sensitivity analysis; set-valued inverse

1. Introduction

In part one of this paper [4], we develop and analyze a numerical method to solve an inverse stochastic sensitivity analysis problem for a smooth deterministic map. Namely, given a (probability) measure on the output of the map, compute the (probability) measure on the input space (comprising data and/or parameters) that produce the output measure. This is the stochastic version of the deterministic inverse problem for the map, and it is also the direct inversion of the forward stochastic sensitivity analysis problem for the map. As such, it deals directly with the inverse of the map in question, rather than, say, a statistical model of the output of the map.

In [4], we formulate this inverse problem using the law of total probability and then analyze an approximate solution method assuming that the map in question can be evaluated exactly.

*Received by the editors February 16, 2010; accepted for publication (in revised form) September 29, 2011; published electronically January 19, 2012. <http://www.siam.org/journals/sinum/50-1/78595.html>

The solution method borrows heavily from techniques used in measure theory. The computed solution provides a systematic method for approximating the probability of any specified event in the input space.

However, our interest lies in situations in which the output of the map is determined implicitly and is difficult and/or expensive to evaluate, e.g., requiring the solution of a differential equation. In addition, we wish to consider the situation in which the measure on the output of the map is only described approximately using a finite number of samples. These practical discretization choices introduce additional numerical errors that affect the computed inverse distribution.

In this paper, we carry out an analysis of the effects of these two numerical discretizations on the computed inverse distribution. As a consequence of the analysis, we prove that the numerical error of the approximate parameter density computed from the algorithm for solving the inverse problem converges to zero as the discretization (both statistical and deterministic) converge to zero. But our main goal is the derivation of an a posteriori error estimate that can be used to evaluate the accuracy of the computed distribution solving the inverse problem, taking into account all sources of statistical and numerical deterministic errors.

While our particular interest is a numerical analysis of the method constructed for the inverse problem in [4], aspects of the analysis we present hold general interest. We present a general error analysis for a computed probability distribution accounting for the effects of finite sampling and errors in each sample resulting from evaluation of a numerically approximated map. Because of the application to inverse problems, we also require error estimates for gradients of quantities of interest computed from numerical solutions of ordinary differential equations, which again is of interest in other contexts, e.g., applications to optimization.

1.1. The inverse problem

As mentioned, the problem we study in [4] is a direct inversion of the forward stochastic sensitivity problem for a deterministic model. We consider an operator $q(\lambda)$ that maps values in a parameter (and data) space Λ to an output space D . We assume there is a *parameter volume* measure μ_Λ on Λ that determines the volume of sets in Λ . The volume measure depends on the units of measure used for the parameters and also reflects the structural dependency among the parameters, e.g., depending on whether μ_Λ is a product measure. The volume measure is specified as part of the model that defines the map $q(\lambda)$ since the parameters must be explicitly defined in the physical model that determines q . We assume that μ_Λ is absolutely continuous with respect to the Lebesgue measure and the volume V of Λ is finite.

The deterministic model can be expressed in terms of a likelihood function $L(q|\lambda)$ of the output q values given the input parameter values λ , where $L(q|\lambda) = \delta(q - q(\lambda))$ is the unit mass distribution at $q = q(\lambda)$. If we specify a density $\sigma_\Lambda(\lambda)$ on the parameter space Λ , then the *law of total probability* implies

$$\rho_{\mathcal{D}}(q) = \int_{\Lambda} L(q|\lambda) \sigma_{\Lambda}(\lambda) d\mu_{\Lambda}(\lambda). \quad (1.1)$$

The stochastic inverse sensitivity analysis problem that we study is the inversion of the integral equation (1.1). Namely, we assume that an observed probability density $\rho_{\mathcal{D}}(q(\lambda))$ is given on the output value $q(\lambda)$ and we seek to compute the corresponding parameter density $\sigma_{\Lambda}(\lambda)$ that yields $\rho_{\mathcal{D}}(q(\lambda))$ via (1.1).

1.2. The solution method

In [4], we present a computational measure theoretic algorithm to approximate the solution of the inverse problem (1.1) by a simple function $\sigma_{\Lambda, M'}(\lambda)$ with respect to a partition $\{b_i\}_{i=1}^{M'}$ of Λ . Paraphrasing the main result from [4], we have the following theorem.

THEOREM 1.1. Given a measurable set $A \subset \Lambda$, we can approximate $P(A)$ using the simple function

$$\sigma_{\Lambda, M'}(\lambda) = \sum_{k=1}^{M'} \frac{P(b_i)}{\mu_{\Lambda}(b_i)} 1_{b_i}(\lambda). \quad (1.2)$$

The constructive proof yields a computational algorithm that generates a probability $P(b_i)$ for each cell b_i using only calculations of volumes in Λ . The main steps of the algorithm are based on the following observations:

- The probability of an interval of output data $[q_m, q_M) \subset \mathcal{D}$ is equal to the probability of the region generalized contours defined by $A = q^{-1}([q_m, q_M))$.
- If $\rho_{\mathcal{D}}(q)$ is constant on $[q_m, q_M)$, then the probability of $b \cap A$ for any event $b \subset \Lambda$ is equal to the probability of A times the ratio of volumes $\mu_{\Lambda}(b \cap A) / \mu_{\Lambda}(A)$.

Thus, the algorithm proceeds by first approximating $\rho_{\mathcal{D}}(q)$ with a simple function, which induces regions of contours with probabilities defined by the approximate output density.

Then, the ratios of volumes for each cell in the partition $\{b_i\}_{i=1}^{M'}$ are computed with respect to all the induced regions of contours. From this, we obtain $P(b_i)$ for each cell and obtain (1.2).

The main focus of the analysis in [4] is on the convergence of the approximate representation to the true representation on a given partition $\{b_i\}_{i=1}^{M'}$ assuming that map the map is evaluated exactly.

1.3. Sources of error

In this paper, we analyze and estimate errors affecting the values $\{P(b_i)\}$ in the representation (1.2) for a fixed partition $\{b_i\}_{i=1}^{M'}$. Since we fix the partition $\{b_i\}_{i=1}^{M'}$, we simplify the notation in [4] by dropping the hat on the piecewise-linear representation $\hat{q}(\lambda)$ of $q(\lambda)$.

In particular, we consider two sources of error that affect the approximation of the representation $\sigma_{\Lambda, M'}(\lambda)$. The first is “statistical error” that arises if the observed probability density $\rho_{\mathcal{D}}(q(\lambda))$ is known only through a finite collection of random samples. This type of error affects the left-hand side of (1.1). For example, finite sampling of the distribution of random variable $q(\lambda)$ is used when the observed distribution is complicated to evaluate or when it is determined by experimental observations. Given an analytic, easy-to-evaluate distribution function for $q(\lambda)$, we need not perform any sampling.

The second source of error arises when we use numerical approximations in the evaluation of the map q , e.g., as happens if q involves solving a differential equation. This means that we use approximate values of q and its gradient to form the approximate representation $\hat{q}(\lambda) \approx \tilde{q}(\lambda)$. This source of error affects the evaluation of the likelihood function in (1.1).

In this paper, we present two kinds of error analysis. We give an a priori convergence analysis that shows that the error tends to zero as the discretization is refined. This analysis

uses error bounds that are robust in the sense of holding under general conditions but which are generally orders of magnitude too large for particular computed solutions. Our main purpose in this paper is to give an a posteriori error analysis that provides the means to compute an relatively accurate error estimate on any particular computed solution. The latter result is important for the purposes of uncertainty quantification and for distributing computational resources in order to achieve a desired accuracy with efficiency.

We let $F(t)$ denote the probability distribution on $\mathbf{\Lambda}$ that represents (1.2), where $t \in \mathbb{R}^d$, and

$$F(t) = P(\{\lambda | \lambda \leq t\}) = P(\lambda \geq t). \quad (1.3)$$

Here the inequality, $\lambda \leq t$, is considered componentwise. We use $F_q(t)$ to denote the probability distribution function of $q(\lambda)$. To simplify the presentation, we assume b_i is contained in a region of contours A_j induced by the simple function approximation to $\rho_{\mathcal{Q}}(q)$. If no sampling is used to evaluate $\rho_{\mathcal{Q}}(q)$ or $F_q(t)$, then the algorithm yields

$$P(b_i) = F_q(q(b_i)) \frac{\int_{b_i} d\mu_{\Lambda}(\lambda)}{\int_{A_i} d\mu_{\Lambda}(\lambda)}, \quad 1 \leq i \leq M', \quad (1.4)$$

where $q(b_i) = \{q(\lambda), \lambda \in b_i\}$. (If $b_i \not\subset A_j$, then we alter (1.4) to sum over the regions of induced contours A_j such that $b_i \cap A_j \neq \emptyset$.) Using (1.4) in (1.3) gives

$$F(t) = \sum_{i=1}^{M'} F_q(q(b_i)) \frac{\int_{b_i \cap \{\lambda \leq t\}} d\mu_{\Lambda}(\lambda)}{\int_{A_i} d\mu_{\Lambda}(\lambda)}. \quad (1.5)$$

For the first source of error, we let $F_{q, \mathcal{N}}(t)$ denote a sample distribution function computed from a finite collection of *error-free* sample values $\{Q_1, \dots, Q_{\mathcal{N}}\}$,

$$F_{q, \mathcal{N}}(t) = \frac{1}{\mathcal{N}} \sum_{n=1}^{\mathcal{N}} 1(Q_n \leq t).$$

This leads to an approximation of $F_{\mathcal{N}}(t) \approx F(t)$ defined

$$F_{\mathcal{N}}(t) = \sum_{i=1}^{M'} F_{q, \mathcal{N}}(q(b_i)) \frac{\int_{b_i \cap \{\lambda \leq t\}} d\mu_{\Lambda}(\lambda)}{\int_{A_i} d\mu_{\Lambda}(\lambda)}.$$

Next we consider the use of an approximation $\tilde{q}(\lambda) \approx q(\lambda)$, which leads to an error in computation of $F_{q, \mathcal{N}}(t)$. We define the approximate sample distribution function $\tilde{F}_{\mathcal{N}}(t)$ as

$$\tilde{F}_{\mathcal{N}}(t) = \sum_{i=1}^{M'} F_{\tilde{q}, \mathcal{N}}(\tilde{q}(b_i)) \frac{\int_{b_i \cap \{\lambda \leq t\}} d\mu_{\Lambda}(\lambda)}{\int_{A_i} d\mu_{\Lambda}(\lambda)}. \quad (1.6)$$

We calculate probabilities using (1.6) and seek to determine the error $F(t) - \tilde{F}_{\mathcal{N}}(t)$. We decompose the error to get

$$|F(t) - \tilde{F}_{\mathcal{N}}(t)| \leq \underbrace{|F(t) - F_{\mathcal{N}}(t)|}_{\mathcal{E}_S} + \underbrace{|F_{\mathcal{N}}(t) - \tilde{F}_{\mathcal{N}}(t)|}_{\mathcal{E}_D}.$$

(1.7)

2. General error analysis for a computed probability distribution

We begin by bounding \mathcal{E}_S . We conduct an a posteriori analysis similar to that used for nonparametric density estimation for elliptic problems with randomly perturbed diffusion coefficients in [13]. The error in the distribution is bounded by

$$\mathcal{E}_S(t) \leq \sum_{i=1}^{M'} |F_q(q(b_i)) - F_{q_{\mathcal{N}}}(q(b_i))| \frac{\int_{b_i \cap \{\lambda \leq t\}} d\mu_{\Lambda}(\lambda)}{\int_{\Lambda_i} d\mu_{\Lambda}(\lambda)}. \quad (2.1)$$

Using standard statistical arguments [13], for any $\epsilon > 0$,

$$\sup_{t \in \mathbb{R}} |F_q(t) - F_{q_{\mathcal{N}}}(t)| \leq 2 \left(\frac{\log(\epsilon^{-1})}{2_{\mathcal{N}}} \right)^{1/2} \quad (2.2)$$

with probability greater than $1 - \epsilon$. It is possible to prove other forms of this bound [13]. Using (2.2) in (2.1) yields for any $\epsilon > 0$,

$$\mathcal{E}_S(t) \leq 2 \left(\frac{\log(\epsilon^{-1})}{2_{\mathcal{N}}} \right)^{1/2} \sum_{i=1}^{M'} \frac{\int_{b_i \cap \{\lambda \leq t\}} d\mu_{\Lambda}}{\int_{\Lambda_i} d\mu_{\Lambda}(\lambda)} \leq 2 \left(\frac{\log(\epsilon^{-1})}{2_{\mathcal{N}}} \right)^{1/2} \quad (2.3)$$

with probability greater than $1 - \epsilon$.

For \mathcal{E}_D , we assume a bound or estimate E_i for the error in $\hat{q}(b_i)$ on each cell b_i . More precisely, the piecewise linear function q is defined on the partition $\{B_j\}$ of Λ , where $q(\lambda) = q(\mu_j) + \nabla q(\mu_j)(\lambda - \mu_j)$ on B_j and μ_j is a chosen value in B_j , and $\tilde{q}(\lambda) = \tilde{q}(\mu_j) + \nabla \tilde{q}(\mu_j)(\lambda - \mu_j)$ on B_j . Hence the error has the form

$$q(\lambda) - \tilde{q}(\lambda) = (q(\mu_i) - \tilde{q}(\mu_i)) + (\nabla q(\mu_i) - \nabla \tilde{q}(\mu_i))(\lambda - \mu_i) \quad \text{on } B_i. \quad (2.4)$$

Hence, we require estimates or bounds for the errors in both $\tilde{q}(\mu_i)$ and $\nabla \tilde{q}(\mu_i)$, respectively. The derivation of the a priori bound or a posteriori error estimate is specific to a particular map q . In section 3, we derive the necessary estimates for nonlinear ordinary differential equations. Similar results hold for elliptic problems [5].

For convenience, we choose the fine partition $\{b_i\}$ so that for each $1 \leq i \leq M$, $b_i \subset B_j$ for some $1 \leq j \leq M$. Thus, for all cells $b_i \subset B_j$ for a fixed j , there is the same deterministic error term associated with $\tilde{q}(b_i)$. We let E_j , $1 \leq j \leq M$, denote the deterministic error associated with each $\tilde{q}(b_i)$ for all $b_i \subset B_j$. Using an analogous argument as in [13],

$$\begin{aligned} \mathcal{E}_D &\leq \sum_{i=1}^{M'} |F_{q, \mathcal{N}}(M_i + E) - F_{q, \mathcal{N}}(m_i - E)| \frac{\int_{A_i} \frac{d\mu_\Lambda(\lambda)}{\int_{A_i} d\mu_\Lambda(\lambda)}}{\int_{A_i} d\mu_\Lambda(\lambda)} \\ &\leq \sum_{i=1}^{M'} |F_q(M_i + E) - F_q(m_i - E)| \frac{\int_{A_i} \frac{d\mu_\Lambda(\lambda)}{\int_{A_i} d\mu_\Lambda(\lambda)}}{\int_{A_i} d\mu_\Lambda(\lambda)} \\ &\quad + \sum_{i=1}^{M'} |F_q(m_i - E) - F_{q, \mathcal{N}}(m_i - E)| \frac{\int_{A_i} \frac{d\mu_\Lambda(\lambda)}{\int_{A_i} d\mu_\Lambda(\lambda)}}{\int_{A_i} d\mu_\Lambda(\lambda)} \\ &\quad + \sum_{i=1}^{M'} |F_q(M_i + E) - F_q(m_i - E)| \frac{\int_{A_i} \frac{d\mu_\Lambda(\lambda)}{\int_{A_i} d\mu_\Lambda(\lambda)}}{\int_{A_i} d\mu_\Lambda(\lambda)}, \end{aligned}$$

where $E = \max_j |E_j|$, $M_i = \max q(b_j)$, and $m_i = \min q(b_j)$.

Using (2.2) for the first two terms on the right-hand side of the inequality we have that for any $\epsilon > 0$

$$\mathcal{E}_D \leq 2 \left(\frac{\log(\epsilon^{-1})}{2\mathcal{N}} \right)^{1/2} + \sum_{i=1}^{M'} |F_q(M_i + E) - F_q(m_i - E)| \frac{\int_{A_i} \frac{d\mu_\Lambda(\lambda)}{\int_{A_i} d\mu_\Lambda(\lambda)}}{\int_{A_i} d\mu_\Lambda(\lambda)}$$

with probability greater than $1 - \epsilon$. Assuming Lipschitz continuity of the distribution F_q with constant L , for any $\epsilon > 0$,

$$\mathcal{E}_D \leq 2 \left(\frac{\log(\epsilon^{-1})}{2\mathcal{N}} \right)^{1/2} + L \max_{1 \leq i \leq M'} (M_i - m_i) \times E$$

with probability greater than $1 - \epsilon$.

Putting together the bounds yields the next theorem.

THEOREM 2.1. For any $\epsilon > 0$,

$$|F(t) - \tilde{F}_{\mathcal{N}}(t)| \leq 2 \left(\frac{\log(\epsilon^{-1})}{2\mathcal{N}} \right)^{1/2} + L \max_{1 \leq i \leq M'} (\max q(b_i) - \min q(b_i)) \times E \quad (2.5)$$

with probability greater than $1 - \epsilon$. If no sampling is used to evaluate $\rho_{\mathcal{Q}}(q)$ or $F_q(t)$, then

$$|F(t) - \tilde{F}(t)| \leq L \max_{1 \leq i \leq M'} (\max q(b_i) - \min q(b_i)) \times E. \quad (2.6)$$

Note that $\hat{F}(t)$ is the distribution calculated using exact values of $\rho_{\mathcal{Q}}(q)$ but approximate values of q .

3. Application to nonlinear ordinary differential equations

We apply the general error analysis to a finite dimensional map q defined implicitly by the solution to a differential equation that depends on a finite number of parameters in the model. We consider the initial value problem

$$\begin{cases} \dot{y} = f(y; \lambda_1), & t > 0, \\ y(0) = \lambda_0, \end{cases} \quad (3.1)$$

where $y \in \mathbb{R}^n$, $f: \mathbb{R}^{n+p} \rightarrow \mathbb{R}^n$ is smooth, and $\lambda = (\lambda_1^\top, \lambda_0^\top)^\top \in \Lambda \subset \mathbb{R}^d$ ($d = p + n$) are the parameters. We solve (3.1) to calculate a linear functional of the solution, or a quantity of interest,

$$q(y) = \int_0^T \langle y, \psi \rangle dt. \quad (3.2)$$

We assume that the solution y of (3.1) depends (implicitly) on parameters λ in a smooth way and denote solutions of (3.1) as y_λ and the quantity of interest as $q(\lambda)$ to emphasize the implicit dependence of the quantity of interest on the parameters. The smooth dependence of solutions to (3.1) on parameters λ implies the dependence of the quantity of interest on λ is also smooth.

3.1. Construction of the piecewise-linear representation

Computing the gradient information is problematic for a differential equation. We use an adjoint equation and variational analysis to do this implicitly. We solve the initial value problem at a reference parameter value $\mu = (\mu_1^\top, \mu_0^\top)^\top$,

$$\begin{cases} \dot{y}_\mu = f(y_\mu; \mu_1), & t > 0, \\ y(0) = \mu_0, \end{cases} \quad (3.3)$$

where (y_μ, μ) is a reference point. We define the exact adjoint problem,

$$\begin{cases} -\dot{\phi} - D_y f(y_\mu; \mu_1)^\top \phi = \psi, & T > t \geq 0, \\ \phi(T) = 0. \end{cases} \quad (3.4)$$

The following theorem [25] relates the value of $q(\lambda)$ to $q(\mu)$ for λ near μ .

THEOREM 3.1. If $f(y; \lambda)$ is twice continuously differentiable with respect to both y and λ and Lipschitz continuous in both y and λ , then the quantity of interest is Fréchet differentiable at (y_μ, μ) with derivative $\nabla q(\mu): \mathbb{R}^d \rightarrow \mathbb{R}$ given by

$$\nabla q(\mu)[\lambda] = \langle (\lambda_0 - \mu_0), \phi(0) \rangle + \int_0^T \langle D_{\lambda_1} f(y_\mu; \mu)(\lambda_1 - \mu_1), \phi \rangle dt. \quad (3.5)$$

Additionally,

$$q(\lambda) \approx q(\mu) + \nabla q(\mu)[\lambda]. \quad (3.6)$$

In the absence of numerical error,

$$q(\lambda) \approx \int_0^T \langle y_\mu, \psi \rangle dt + \langle (\lambda_0 - \mu_0), \phi(0) \rangle + \int_0^T \langle D_{\lambda_1} f(y_\mu; \mu_1)(\lambda_1 - \mu_1), \phi \rangle dt \quad (3.7)$$

for λ close to μ .

The global piecewise-linear approximation on the partition $\{B_i\}_{i=1}^M$ of Λ is constructed by using the local linearization on each cell B_i to obtain

$$\widehat{q}(\lambda) := \sum_{i=1}^M \left(q(\mu_i) + \langle \nabla_q q(\mu_i), (\lambda - \mu_i) \rangle \right) 1_{B_i}(\lambda), \quad (3.8)$$

where μ_i is the reference parameter value chosen in cell B_i .

3.2. Discretization

The a posteriori error estimate uses a variational analysis after introducing an adjoint problem. The variational analysis makes it natural to write the discretization method in the finite element framework. This is not restrictive as most common finite difference schemes can be written as a finite element method with a particular choice of quadrature for evaluating integrals.

A finite element method is based on the variational formulation of the differential equation. For the differential equation,

$$\begin{cases} \dot{x} = g(x, t), & 0 < t \leq T, \\ x(0) = x_0, \end{cases} \quad (3.9)$$

the problem is to find $x \in \mathcal{C}^1([0, 1])$ such that

$$\begin{cases} \int_0^t (\dot{x}, v) ds = \int_0^t (g(x, t), v) ds, & t > 0, \\ x(0) = x_0 \end{cases} \quad (3.10)$$

for all $v \in \mathcal{C}^1([0, 1])$. (We use g instead of f because there are several problems that have to be solved below.)

We compute a solution on the interval $[0, T]$, and we discretize the interval $0 = t_0 < t_1 < \dots < t_N = T$ with time intervals $I_j = (t_{j-1}, t_j)$ and time steps $k_j = t_j - t_{j-1}$. The finite element method produces a piecewise polynomial approximation. We use $\mathcal{P}^q(I_j)$ to denote the space of polynomials of degree q and less on time interval I_j and define the space of piecewise polynomials,

$$\mathcal{V}^{(q)} = \left\{ U : U|_{I_j} \in \mathcal{P}^{(q)}(I_j), j=1, 2, 3, \dots, \right\}.$$

We consider the discontinuous Galerkin (dGq) finite element method that produces an approximate solution $X \in \mathcal{V}^{(q)}$ [18]. Since X may be discontinuous at time nodes, we define $X_{j=1}^+ = \lim_{t \downarrow t_{j-1}} X(t)$, $X_{j-1}^- = \lim_{t \uparrow t_{j-1}} X(t)$, and $[X]_{j-1} = X_{j-1}^+ - X_{j-1}^-$. The approximation is computed interval by interval. We set $X_0 = x_0$. Then we compute $X \in \mathcal{P}^q(t_{j-1}, t_j)$ successively for $j = 1, 2, \dots, N$, satisfying

$$\int_{t_{j-1}}^{t_j} (\dot{X}, v) ds + ([X]_{j-1}, v_{j-1}^+) = \int_{t_{j-1}}^{t_j} (g(X, t), v) ds, \quad \text{for all } v \in \mathcal{P}^q(t_{j-1}, t_j). \quad (3.11)$$

Remark 3.1. If $g(x, t) \equiv g(x)$ and $q = 0$, the dG0 approximation matches the backward Euler approximation at the time nodes. In general, we may obtain various difference schemes, e.g., the subdiagonal Pade schemes, by employing quadrature to evaluate the integrals in (3.11) [17, 27]. There is also a continuous Galerkin (cG) approximation that produces yet other classes of approximations [9]. We carry out the analysis below for the dG scheme assuming the integrals in (3.11) are computed exactly. The extension of the a posteriori analysis to handle quadrature and the cG method is straightforward [12] and we do not discuss this further.

3.3. The effect of using an approximate solution on the piecewise-linear representation

The main interest is in treating the effects of using a numerical approximation Y_μ in the linearization of the forward problem used to construct an adjoint. We define the approximate adjoint using (3.4) with “perturbed” operator $D_y f(Y_\mu; \mu_1)$,

$$\begin{cases} -\dot{\Phi} - D_y f(Y_\mu; \mu_1)^T \Phi = \psi, & T > t \geq 0, \\ \Phi(T) = 0. \end{cases} \quad (3.12)$$

We assume $f(y, \lambda)$ is twice continuously differentiable with respect to both y and λ , so that standard convergence results for Y_μ imply that over some (short) time interval $[0, T]$,

$$\|D_y f(y_\mu; \mu_1) - D_y f(Y_\mu; \mu_1)\| \leq K \|y_\mu - Y_\mu\|_V, \quad (3.13)$$

where $\|\cdot\|_V$ and $\|\cdot\|_U$ are the $L^2([0, T])$ norm of some appropriate matrix and vector norms of the arguments, respectively.

Let $\tilde{q}(\lambda)$ denote the approximate quantity of interest calculated using (3.6) with Y_μ and Φ in place of y_μ and ϕ ,

$$\tilde{q}(\lambda) = \int_0^T \langle Y_\mu, \psi \rangle dt + \langle (\lambda_0 - \mu_0), \Phi(0) \rangle + \int_0^T \langle D_{\lambda_1} f(Y_\mu; \mu_1)(\lambda_1 - \mu_1), \Phi \rangle dt \quad (3.14)$$

with error $q(\lambda) - \tilde{q}(\lambda)$. Taking the difference of (3.7) and (3.14) gives

$$\begin{aligned} q(\lambda) - \tilde{q}(\lambda) = & \underbrace{\int_0^T \langle (y_\mu - Y_\mu), \psi \rangle dt}_I + \underbrace{\langle (\lambda_0 - \mu_0), (\phi(0) - \Phi(0)) \rangle}_{II} \\ & + \underbrace{\int_0^T \langle (D_{\lambda_1} f(y_\mu; \mu_1)(\lambda_1 - \mu_1), \phi) - (D_{\lambda_1} f(Y_\mu; \mu_1)(\lambda_1 - \mu_1), \Phi) \rangle dt}_{III}. \end{aligned} \quad (3.15)$$

Term **I** is a linear functional of the error $y_\mu - Y_\mu$ and it can be estimated using standard a posteriori analysis techniques as described below. Term **II** measures the effect of using Y_μ and Φ on the sensitivity of $q(\lambda)$ to changes in the initial conditions. Term **III** measures the effect of using Y_μ and Φ on the sensitivity of $q(\lambda)$ to changes in model parameters.

The terms **II** and **III** depend linearly on the vector $\lambda - \mu$. The analysis below produces estimates that also depend on this vector linearly so that the error estimates for these terms

are also linear functions of this vector. Thus, following the analysis described below for p linearly independent vectors $\lambda - \mu$, we obtain a set of error estimates such that the error defined by **II** and **III** for any vector $\lambda - \mu$ can be written as a linear combination from this set of error estimates.

3.4. Convergence and order of accuracy

We can use straightforward a priori error analysis on (3.15) to show that $|q(\lambda) - \tilde{q}(\lambda)|$ converges at the same order as the dGq method over a short time period under the assumption that f is twice continuously differentiable.

3.5. Estimate of the error in a quantity of interest

We compute an a posteriori error estimate using variational analysis and adjoint problems [9, 18, 7, 8, 12, 27, 6]. We begin by recalling the a posteriori estimate of error in a quantity of interest. Let X denote the dGq approximation to (3.9) and let $e = x - X$, where x solves (3.9) exactly. We linearize around X in the sense of perturbing the operator to arrive at the adjoint problem

$$\begin{cases} -\dot{v} = \overline{g'(x, X)}^T \phi + \psi_1(t), & T > t \geq 0, \\ v(T) = \psi_2, \end{cases} \quad (3.16)$$

where $\overline{g'(x, X)} = \int_0^1 g'(sx + (1-s)X, t)$. For simplicity, we use g' for $\overline{g'(x, X)}$ below. If $\psi_1(t) \equiv 0$, then the quantity of interest is $(e(T), \psi_2)$. If $\psi_2 = 0$, then the quantity of interest is $\int_0^T (e(t), \psi_i(t)) dt$.

Assume $\psi_1(t) \equiv 0$ in (3.16). Take the inner product of the adjoint problem with e and integrate from 0 to T to obtain

$$-\int_0^T (\dot{v}, e) dt - \int_0^T ((g')^T v, e) dt = 0. \quad (3.17)$$

We decompose (3.17) into a sum of integral equations over each time interval in the discretization and integrate by parts over each interval to get

$$-\sum_{n=1}^N (e, v)|_{t_{n-1}}^{t_n} + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (v, \dot{e}) dt - \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (v, g' e) dt = 0. \quad (3.18)$$

Since $e = x - X$ might be discontinuous at the boundaries of each interval, we expand the first term on the right-hand side of (3.18) to

$$-\sum_{i=1}^N (e, v)|_{t_{n-1}}^{t_n} = (e(0), v(0)) + \sum_{n=2}^N ([X]_{n-1}, v_{n-1}) - (e(T), v(T)) \quad (3.19)$$

with $v_{n-1} = v(t_{n-1})$. Substitution of (3.19) into (3.18) and rearranging the terms yields

$$(e(T), v(T)) = (e(0), v(0)) + \sum_{n=1}^N ([X]_{n-1}, v_{n-1}) + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (v, \dot{e}) dt - \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (v, g' e) dt.$$

Substituting $e = x - X$ and using $\dot{x} - g(x) = 0$ gives

$$(e(T), v(T)) = (e(0), v(0)) + \sum_{n=2}^N ((X)_{n-1}, v_{n-1}) - \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (\dot{X} - g(X), v) dt. \quad (3.20)$$

Similarly, if $\psi_2 = 0$ and $\psi_1(t)$ is nonzero for some $t \in (0, T)$, we obtain

$$\int_0^T (e, \psi_1) dt = (e(0), v(0)) + \sum_{n=2}^N ([X]_{n-1}, v_{n-1}) - \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (\dot{X} - g(X), v) dt. \quad (3.21)$$

We summarize as the following theorem.

THEOREM 3.2. A computable estimate of the error in a quantity of interest of (3.9) is obtained by solving (3.16) and computing either (3.20) or (3.21).

Implementation details—Using the a posteriori estimate involves several important practical considerations. We discuss two.

Often “Galerkin orthogonality” is used to introduce a projection of the adjoint solution into the approximation space for the forward solution. This makes the estimate easier to compute and has the effect of “localizing” the error contributions from each time step.

The estimate (3.20) is computable provided that we can compute the adjoint solution ψ . This raises several issues. The first is that we cannot use $\overline{g'(x, X)}$ in practice since this requires the unknown solution x . Typically, we use $\overline{g'(x, X)} \approx \overline{g'(X, X)} = g'(X)$. The effect of this approximation on the computation of ψ can be analyzed, e.g., [10]. The error depends on the accuracy of X , so typical short time error bounds can be proved. The second issue is that in practice we solve the adjoint problem using a numerical method, typically using a higher-order method than used for the forward solution.

The consequence is that in practice we use an approximate adjoint solution. We can alter the analysis below to take into the account the effect on the estimate, but this significantly complicates the presentation of the results while it is generally not significant.

3.6. Estimating term II in (3.15)

We first observe that term II is a linear functional of the error arising from solving the exact adjoint with the approximate adjoint. We adapt the a posteriori analysis to estimate the error of this approximation. We define the adjoint to the approximate adjoint as

$$\begin{cases} \dot{w} = D_y f(Y_\mu; \mu_1) w, & 0 < t \leq T, \\ w(0) = (\lambda_0 - \mu_0). \end{cases}$$

Since $\dot{w} - D_y f(Y_\mu; \mu_1) w = 0$, we have

$$\begin{aligned} 0 &= \int_0^T \langle \dot{w} - D_y f(Y_\mu; \mu_1) w, (\phi - \Phi) \rangle dt \\ &= \left[\langle w(T), (\phi(T) - \Phi(T)) \rangle - \langle w(0), (\phi(0) - \Phi(0)) \rangle \right] - \int_0^T \langle w, (\dot{\phi} - \dot{\Phi}) \rangle dt - \int_0^T \langle w, D_y f(Y_\mu; \mu_1)^T (\phi - \Phi) \rangle dt \\ &= - \langle (\lambda_0 - \mu_0), (\phi(0) - \Phi(0)) \rangle + \int_0^T \left\langle w, \left[-\dot{\phi} - D_y f(Y_\mu; \mu_1)^T \phi + \dot{\Phi} + D_y f(Y_\mu; \mu_1)^T \Phi \right] \right\rangle dt. \end{aligned}$$

This gives

$$\Pi = \int_0^T \left\langle w, \left[-\dot{\phi} - D_y f(Y_\mu; \mu_1)^T \phi + \dot{\Phi} + D_y f(Y_\mu; \mu_1)^T \Phi \right] \right\rangle dt. \quad (3.22)$$

By adding and subtracting $D_y f(Y_\mu; \mu_1)^T \phi$ to the differential equation in (3.4) for the exact adjoint, we have

$$-\dot{\phi} - D_y f(Y_\mu; \mu_1)^T \phi = \left[D_y f(y_\mu; \mu_1) - D_y f(Y_\mu; \mu_1) \right]^T \phi + \psi. \quad (3.23)$$

Substituting (3.23) into (3.22) and using (3.12), we have

$$\begin{aligned} \Pi &= \int_0^T \left\langle w, \left[D_y f(y_\mu; \mu_1) - D_y f(Y_\mu; \mu_1) \right]^T \Phi \right\rangle dt \quad (3.24) \\ &+ \int_0^T \left\langle w, \left[D_y f(y_\mu; \mu_1) - D_y f(Y_\mu; \mu_1) \right]^T (\phi - \Phi) \right\rangle dt. \quad (3.25) \end{aligned}$$

We show that the second term on the right-hand side of the last equation is higher order and estimate the first term on the right-hand side. If $f(y, \lambda)$ is three times continuously differentiable, then we use Taylor's theorem to get

$$\begin{aligned} &\int_0^T \left\langle w, \left[D_y f(y_\mu; \mu_1) - D_y f(Y_\mu; \mu_1) \right]^T \Phi \right\rangle dt \\ &\approx \int_0^T \left\langle w, \underbrace{[\Phi^T \otimes \mathbf{J}]}_{n \times n^2} \underbrace{[D_y (\text{vec}(D_y f(Y_\mu; \mu_1)^T))]}_{n^2 \times n} (y_\mu - Y_\mu) \right\rangle dt \\ &= \int_0^T \left\langle [D_y (\text{vec}(D_y f(Y_\mu; \mu_1)^T))]^T [\Phi^T \otimes \mathbf{J}]^T w, (y_\mu - Y_\mu) \right\rangle dt, \end{aligned}$$

where \mathbf{J} denotes the $n \times n$ identity matrix and the vector operator denoted vec is a map from $\mathbb{R}^{l \times m} \rightarrow \mathbb{R}^{lm}$ defined by stacking the columns (in order) of a matrix to form a column vector. We let

$$\psi_{II} = \left[D_y \left(\text{vec} \left(D_y f(Y_\mu; \mu_1)^T \right) \right) \right]^T [\Phi^T \otimes \mathbf{J}]^T w.$$

The first term on the right-hand side is a linear functional of the error $y_\mu - Y_\mu$ and can be estimated by Theorem 3.2.

We now show that the second term is higher-order. Let $\eta = \phi - \Phi$ then

$$\begin{cases} -\dot{\eta} = D_y f(y_\mu; \mu_1)^T \phi - D_y f(Y_\mu; \mu_1)^T \Phi, & T > t \geq 0, \\ \eta(T) = 0. \end{cases} \quad (3.26)$$

If Y_μ is sufficiently close to y_μ over $[0, T]$, then

$$D_y f(Y_\mu; \mu_1)^T = D_y f(y_\mu; \mu_1)^T + \epsilon(t), \quad t \in [0, T], \quad (3.27)$$

where $\epsilon(t)$ is a perturbation matrix satisfying $\|\epsilon(t)\| \leq C \|y_\mu - Y_\mu\|$ for some $C > 0$ and all $t \in [0, T]$. Substituting (3.27) into (3.26) gives

$$\begin{cases} -\dot{\eta} = D_y f(y_\mu; \mu_1)^T \eta + \epsilon(t) \Phi(t), & T > t \geq 0, \\ \eta(T) = 0. \end{cases} \quad (3.28)$$

Let $\Sigma(t)$ denote the fundamental matrix of (3.29); then

$$\eta(t) = -\Sigma(t) \int_T^t [\Sigma(s)]^{-1} \epsilon(s) \Phi(s) ds.$$

This implies that

$$\|\eta(t)\| \geq \|\Sigma(t)\| \int_0^T \|\Sigma(s)^{-1}\| \|\epsilon(s)\| \|\Phi(s)\| ds \leq C \|y_\mu - Y_\mu\|_U. \quad (3.29)$$

Here, $\|\cdot\|_U$ is interpreted as before to mean the $L^2([0, T])$ norm of a given vector norm of the argument, and $C > 0$ is some constant that bounds the product of $\sup_{t \in [0, T]} \|\Sigma(t)\|$, $\sup_{t \in [0, T]} \|\Sigma(t)^{-1}\|$, and $\sup_{t \in [0, T]} \|\Phi(t)\|$. Thus, by Lipschitz continuity of the first derivatives of $f(y; \lambda)$ and (3.29),

$$\left| \int_0^T \left\langle w, [D_y f(y_\mu; \mu_1) - D_y f(Y_\mu; \mu_1)]^T (\phi - \Phi) \right\rangle dt \right| \leq C \|y_\mu - Y_\mu\|_U^2.$$

3.7. Estimating term III in (3.15)

Add and subtract $\langle D_\lambda f(Y_\mu; \mu_1)(\lambda_1 - \mu_1), \phi \rangle$ and write **III** = **IIIa** + **IIIb**, where

$$\begin{aligned} \text{IIIa} &= \int_0^T \langle (D_\lambda f(y_\mu; \mu_1) - D_\lambda f(Y_\mu; \mu_1))(\lambda_1 - \mu_1), \phi \rangle dt, \\ \text{IIIb} &= \int_0^T \langle D_\lambda f(Y_\mu; \mu_1)(\lambda_1 - \mu_1), (\phi - \Phi) \rangle dt, \end{aligned}$$

and estimate **IIIa** and **IIIb**.

Estimating term **IIIa**. Add and subtract $\langle (D_\lambda f(y_\mu; \mu_1) - D_\mu f(Y_\mu; \mu_1))(\mu_1 - \mu_1), \Phi \rangle$ and write **IIIa** = **IIIaa** + **IIIab**, where

$$\begin{aligned} \text{IIIaa} &= \int_0^T \langle (D_\lambda f(y_\mu; \mu_1) - D_\lambda f(Y_\mu; \mu_1))(\lambda_1 - \mu_1), (\phi - \Phi) \rangle dt, \\ \text{IIIab} &= \int_0^T \langle (D_\lambda f(y_\mu; \mu_1) - D_\lambda f(Y_\mu; \mu_1))(\lambda_1 - \mu_1), \Phi \rangle dt. \end{aligned}$$

We show that **IIIaa** is higher-order. We know that $\|\phi - \Phi\| \leq C \|y_\mu - Y_\mu\|_U$ for some constant $C > 0$; therefore

$$|\text{IIIaa}| \leq C \|y_\mu - Y_\mu\|_U^2$$

for some constant $C > 0$.

Again assuming that $f(y; \lambda)$ is three times continuously differentiable,

$$(D_\lambda f(y_\mu; \mu_1) - D_\lambda f(Y_\mu; \mu_1))(\lambda_1 - \mu_1) \approx [(\lambda_1 - \mu_1)^T \otimes J] [D_y (\text{vec}(D_\lambda f(Y_\mu; \mu_1)))] (y_\mu - Y_\mu).$$

We substitute this estimate into **IIIab** so that

$$\begin{aligned} \text{IIIab} &\approx \int_0^T \left\langle [(\lambda_1 - \mu_1)^T \otimes J] \left[D_y \left(\text{vec} \left(D_{\lambda} f \left(Y_{\mu}; \mu_1 \right) \right) \right) \right] (y_{\mu} - Y_{\mu}), \Phi \right\rangle dt \\ &= \int_0^T \left\langle (y_{\mu} - Y_{\mu}), \left[D_y \left(\text{vec} \left(D_{\lambda} f \left(Y_{\mu}; \mu_1 \right) \right) \right) \right]^T [(\lambda_1 - \mu_1)^T \otimes J]^T \Phi \right\rangle dt. \end{aligned}$$

We let

$$\psi_{\text{IIIab}} = \left[D_y \left(\text{vec} \left(D_{\lambda} f \left(Y_{\mu}; \mu_1 \right) \right) \right) \right]^T [(\lambda_1 - \mu_1)^T \otimes J]^T \Phi.$$

Thus, we have represented **IIIab** as a linear functional of the error in $y_{\mu} - Y_{\mu}$, which can be estimated by Theorem 3.2.

Estimating term IIIb. We let $\psi_{\text{IIIb}} = D_{\lambda} f(Y_{\mu}, \mu_1)(\lambda_1 - \mu_1)$ so that

$$\text{IIIb} = \int_0^T \langle \psi_{\text{IIIb}}, (\phi - \Phi) \rangle dt.$$

Thus, **IIIb** is a linear functional of the error in the adjoint solutions $\phi - \Phi$. We apply Theorem 3.2. We again define an adjoint to the *approximate* adjoint as

$$\begin{cases} \dot{z} - D_y f(Y_{\mu}; \mu_1) z = \psi_{\text{IIIb}}, & 0 < t \leq T, \\ z(0) = 0. \end{cases}$$

We perform a standard variational argument to obtain

$$\begin{aligned} \text{IIIb} &= \left[\underbrace{\langle z(T), (\phi(T) - \Phi(T)) \rangle}_{\phi(T) - \Phi(T) = 0} - \underbrace{\langle z(0), (\phi(0) - \Phi(0)) \rangle}_{z(0) = 0} - \int_0^T \langle z, (\dot{\phi} - \dot{\Phi}) \rangle dt \right] \\ &\quad - \int_0^T \langle z, D_y f(Y_{\mu}; \mu_1)^T (\phi - \Phi) \rangle dt \\ &= \int_0^T \langle z, -\dot{\phi} - D_y f(Y_{\mu}; \mu_1) \phi + \dot{\Phi} + D_y f(Y_{\mu}; \mu_1) \Phi \rangle dt. \end{aligned}$$

Using (3.26)–(3.28) in the right-hand side above, we have

$$\text{IIIb} = \int_0^T \left\langle z, \left[D_y f(y_{\mu}; \mu_1) - D_y f(Y_{\mu}; \mu_1) \right]^T \Phi \right\rangle dt - \int_0^T \left\langle z, \left[D_y f(y_{\mu}; \mu_1) - D_y f(Y_{\mu}; \mu_1) \right]^T (\phi - \Phi) \right\rangle dt.$$

The two terms on the right-hand side are analogous to (3.24) and (3.25). The second term on the right-hand side has already been proved to be higher-order. Therefore, the second term is neglected in the estimate. The first term is estimated similarly to how (3.24) was estimated. We define

$$\tilde{\psi}_{\text{IIIb}} = \left[D_y \left(\text{vec} \left(D_y f \left(Y_{\mu}; \mu_1 \right)^T \right) \right) \right]^T (\Phi^T \otimes J)^T z,$$

and the first term is approximated by

$$\int_0^T \langle \tilde{\psi}_{llb}, (Y_\mu - Y_\mu) \rangle dt,$$

which is a linear functional of the error of $y_\mu - Y_\mu$ and is estimable by Theorem 3.2. This completes the proof of Theorem 3.3.

THEOREM 3.3.

Let Y_μ and Φ denote the numerical solutions to the initial value problem (3.3) and the approximate adjoint problem (3.12), respectively.

Apply Theorem 3.2 to estimate term \mathbf{I} .

Let p_m be the number of model parameters and p_i the number of initial conditions ($p_m + p_i = p$)

For $I = 1, \dots, p$ do

if $i \leq p_m$ then

Let z denote the solutions to the adjoint to the approximate adjoint problem

$$\begin{cases} \dot{x} - D_y f(Y_\mu; \mu_1) x = D_\lambda f(Y_\mu, \mu_1) \delta_i, & 0 < t \leq T, \\ x(0) = 0, \end{cases}$$

where δ_i denotes the i th standard basis vector in \mathbb{R}^{p_m}

Set

$$\begin{aligned} \psi_{llab} &= \left[D_y \left(\text{vec} \left(D_\lambda f \left(Y_\mu; \mu_1 \right) \right) \right) \right]^T \left[\delta_i^T \otimes J \right]^T \Phi, \\ \tilde{\psi}_{llb} &= \left[D_y \left(\text{vec} \left(D_y f \left(Y_\mu; \mu_1 \right)^T \right) \right) \right]^T \left[\Phi^T \otimes J \right]^T z. \end{aligned}$$

Solve (3.12) with data given by the above vectors and use Theorem 3.2 to compute the standard error representations given by

$$e_1^i := \int_0^T \langle (y_\mu - Y_\mu), \psi_{llab} \rangle dt, \quad e_2^i := \int_0^T \langle \tilde{\psi}_{llb}, (y_\mu - Y_\mu) \rangle dt.$$

else

Let w denote the solutions to the adjoint to the approximate adjoint problem

$$\begin{cases} \dot{x} - D_y f(Y_\mu; \mu_1) x = 0, & 0 < t \leq T, \\ x(0) = \delta_i, \end{cases}$$

where δ_i denotes the i th standard basis vector in \mathbb{R}^{p_i}

Set

$$\psi_{ii} = \left[D_y \left(\text{vec} \left(D_y f(Y_\mu; \mu_1) \right)^T \right) \right]^T \left[\Phi^T \otimes J \right]^T w$$

Solve (3.12) with data given by the above vectors and use Theorem 3.2 to compute the standard error representations given by

$$e_3^i := \int_0^T \langle \psi_i, (y_\mu - Y_\mu) \rangle dt$$

end if end for Fix λ and set $u := \lambda - \mu$, where

$$u = (v_1 \cdots, v_{pm} \quad \rho_1 \cdots \rho_{pi})^T.$$

The error $q(\lambda) - \hat{q}(\lambda)$ is given by

$$q(\lambda) - \hat{q}(\lambda) = e_u + h.o.t.,$$

where e_u is the computable error estimate given by

$$e_u = e_0 + \sum_{i=1}^{pm} v_i \sum_{j=1}^2 e_j^i + \sum_{i=1}^{pi} \rho_i e_3^i.$$

This theorem provides a means of computing the E_j required in Theorem 2.1. Set $E_j = \max_{\lambda \in B_j} e_u$. For convex polygonal cells B_j , computation of E_j is straightforward since e_u is a linear function of λ , so the maximum occurs on the boundary.

4. Examples

We present examples that illustrate the properties of the computable a posteriori error estimate. In the case of deterministic computations, it is standard to test the accuracy of the estimate by direct comparison to the error on problems for which the actual error is known or can be approximated using an extremely accurate reference solution. However, it is more complicated to test accuracy for the estimates we have derived for stochastic computations because of the nature of the a posteriori bound we use for the stochastic component of the error.

We have explored the accuracy of the a posteriori bound on the effects of finite sampling in [13, 14]. Likewise, the accuracy of the a posterior error estimate for deterministic problems is well recorded; e.g., see [9, 18, 7, 8, 12] and many research papers. We do not repeat tests on these aspects here. Rather, in Examples 1 and 2 we explore the accuracy of the a posteriori error estimates on errors in computed derivatives that are needed for the estimate on the solution of the inverse problem. These estimates are new so their properties have not been explored in the literature. Finally, in Example 3 we present an example in which we check the a posteriori estimate against a direct approximation of the error.

In all the examples, the numerical solution and error estimates are computed using GAASP.¹ We use a first-order discontinuous Galerkin (dG1) method for the forward solve and a second-order continuous Galerkin (cG2) method for all of the adjoint solves. We use the

adaptive time step capability in GAASP to control the numerical integration error. In the first two examples, we terminate time step refinement once the error estimate corrects the estimated gradient by less than 10%.

4.1. Example 1

The first example is a coupled linear system with four parameters:

$$\begin{cases} \dot{x}_1 = \frac{x_1}{\lambda_1(1+t)} - \lambda_2 t x_2, & 0 < t \leq T, \\ \dot{x}_2 = \frac{x_2}{\lambda_3(1+t)} + \lambda_4 t x_1, & 0 < t \leq T, \\ x_1(0) = x_1, 0, \\ x_2(0) = x_2, 0. \end{cases} \quad (4.1)$$

The adjoint problem is

$$\begin{cases} -\dot{\phi}_1 - \frac{1}{\lambda_1(1+t)}\phi_1 + \lambda_4 t \phi_2 = \psi_1, & T > t \geq 0, \\ -\dot{\phi}_2 + \lambda_2 t \phi_1 - \frac{1}{\lambda_3(1+t)}\phi_2 = \psi_2, & T > t \geq 0, \\ \phi_1(T) = \phi_1, T, \\ \phi_2(T) = \phi_2 T. \end{cases} \quad (4.2)$$

Computing the true errors requires knowledge of the exact ϕ . To this end, we choose $\psi(t) = (\psi_1, \psi_2)^T$ and $\phi(T) = (\phi_{1,T}, \phi_{2,T})^T$ so that $\phi(t) = (t, 1)^T$.

In this linear example, $\mathbf{II} = 0$. We report the error estimates for term \mathbf{III} . We take $\mu = (2, 2, 2, 2)^T$, so $y_\mu = (\sqrt{1+t} \cos(t^2), \sqrt{1+t} \sin(t^2))^T$.

We consider both $T = 3$ and $T = 10$. We plot the forward solutions y_μ and Y_μ for $T = 3$ with two different time steps in Figure 4.1. Table 4.1 shows the error estimate results for $T = 3$. Since the computed error estimates tend to be accurate, we can often compute a corrected gradient by adding the error estimate to the computed (estimated) gradient. We see improvement in the corrected gradient by comparing the fourth and last columns of Table 4.1. At $T = 3$, Y_μ is a good approximation of y_μ at the coarse time step of 0.2 as seen in Figure 4.1, so the second derivative calculations involving Y_μ used in the error estimates produce accurate error estimates beginning at this time step.

We plot the forward solutions y_μ and Y_μ for $T = 10$ with four different time steps in Figures 4.2 and 4.3. The oscillations of y_μ increase in magnitude and with higher frequency as time increases. As seen in Table 4.2, when the error estimate is of the same order of magnitude as the estimated gradient, the estimate cannot be used to correct the gradient.

4.2. Example 2

The second example is a nonlinear problem with two parameters:

$$\begin{cases} \dot{x} = -(\lambda_1 + \sin(\lambda_2 t)) x^2, & 0 < t \leq T, \\ x(0) = x_0. \end{cases} \quad (4.3)$$

We set $\mu = (-0.1, 20)^T$ so $y_\mu(t) = 20 / (20 + 1 + (-0.1)20t - \cos(20t))$. The quantity of interest is $y(T)$. The adjoint problem is

¹Write to estep@math.colostate.edu for information.

$$\begin{cases} -\dot{\phi} + 2y_{\mu}(-0.1 + \sin(20t))\phi = 0, & T > t \geq 0, \\ \phi(T) = 1. \end{cases} \quad (4.4)$$

The solution to (4.4) is $\phi(t) = C(20 + 1 + 20(-0.1)t - \cos(20t))^2$, where C is chosen so $\phi(T) = 1$. Since (4.3) is nonlinear, we report the error estimates for both terms **II** and **III**.

We show results for both $T = 3.9$ and $T = 10$. We plot the forward solutions y_{μ} and Y_{μ} and adjoint solutions ϕ and Φ for $T = 3.9$ and $T = 10$ with three different time steps in Figures 4.4 and 4.5, respectively. Tables 4.3 and 4.4 show the error estimate results for $T = 3.9$ and $T = 10$, respectively.

4.3. Example 3

We consider the nonlinear example first presented in [4]:

$$\begin{cases} \dot{x} = \lambda_1 \sin(\lambda_2 x), & 0 < t \leq T, \\ x(0) = 1. \end{cases}$$

The quantity of interest is the average value of $x(t)$ over the time interval $[0, 2]$. Thus, we set $\psi(t) = \mathbf{1}_{[0,2]}(t)/2$ in the adjoint problem. We use a time step of 0.25 to solve at each point of a 20×20 grid of uniformly spaced parameter values in $\Lambda = [.8, 1.2] \times [.1, \pi \times .1]$ and compute the simple function approximation of the σ_{Λ} shown in Figure 4.6, where we use 10,000 samples of the quantity of interest to approximate the output density. We denote the associated distribution function by $\tilde{F}^{(1)}(t)$.

We have $\max_{1 \leq i \leq M} (\max q(b_i) - \min q(b_i)) = 2.69 \times 10^{-1}$ and the corresponding estimate is $E = 7.53 \times 10^{-04}$. The normal distribution imposed on the quantity of interest has a small variance (approximately 6.72×10^{-03}) so the Lipschitz constant of the distribution is bounded by 5. Thus, using (2.5) with $\epsilon = 0.05$, we have that

$$|F(t) - \tilde{F}^{(1)}(t)| \leq 2.55 \times 10^{-22}$$

with probability 95%.

In order to compare the computed a posteriori error estimate to the true error, we directly approximate the error in a computed solution by using another more accurate solution. We use a time step of 1.0×10^{-02} to compute solutions to the forward problem at each parameter in the 20×20 grid, and we invert using 10^8 samples of the output data and use the same resolution in Λ of 60×60 small cells to obtain another approximate distribution function that we denote $\tilde{F}^{(2)}(t)$. We compare $\tilde{F}^{(1)}(t)$ to $\tilde{F}^{(2)}(t)$ and compare to the error bound above. We evaluate the difference in these distributions at the upper-right corner of each b_i and plot the absolute value of the difference in Figure 4.7. The maximum computed absolute value of error at these points is less than 6.70×10^{-03} , which is within the error bound above.

Acknowledgments

The work of this author was supported in part by the Department of Energy (DE-FG02-05ER25699) and the National Science Foundation (DGE-0221595003, MSPACSE-0434354).

The work of this author was supported in part by the Defense Threat Reduction Agency (HDTRA1-09-1-0036), the Department of Energy (DE-FG02-04ER25620, DEFG02-05ER25699, DE-FC02-07ER54909, DE-SC0001724), Lawrence Livermore National Laboratory (B573139, B584647), the National Aeronautics and Space

Administration (NNG04GH63G), the National Science Foundation (DMS-0107832, DMS-0715135, DGE-0221595003, MSPA-CSE-0434354, ECCS-0700559), Idaho National Laboratory (00069249), NSF/NIGMS (R01GM096192), and the Sandia Corporation (PO299784).

The work of this author was supported in part by the Department of Energy (DE-FG02-04ER25620, DE-FG02-05ER25699) and the National Science Foundation (DGE-0221595003, MSPA-CSE-0434354).

REFERENCES

- [1]. Bernardo JM. Reference posterior distributions for Bayesian inference. *J. Roy. Statist. Soc.* 1979; 41:113–147.
- [2]. Billingsley, P. *Probability and Measure*. John Wiley & Sons; New York: 1995.
- [3]. Butler, T.; Estep, D. A measure-theoretic computational method for inverse sensitivity problems III: Multiple output quantities of interest. in preparation
- [4]. Butler T, Estep D. A computational measure theoretic approach to inverse sensitivity problems I: Basic method and analysis. *SIAM J. Numer. Anal.* to appear.
- [5]. Butler, T.; Colorado State University. Ph.D. thesis, Department of Mathematics. Fort Collins, CO; 2009. *Computational Measure Theoretic Approach to Inverse Sensitivity Analysis: Methods and Analysis*.
- [6]. Carey V, Estep D, Johansson A, Larson M, Tavener S. Blockwise adaptivity for time dependent problems based on coarse scale adjoint solutions. *SIAM J. Sci. Comput.* 2010; 32:2121–2145.
- [7]. Eriksson, K.; Estep, D.; Hansbo, P.; Johnson, C. *Acta Numer.* Cambridge University Press; Cambridge, UK: 1995. Introduction to adaptive methods for differential equations, in *Acta Numerica* 4, 1995; p. 105-158.
- [8]. Eriksson, K.; Estep, D.; Hansbo, P.; Johnson, C. *Computational Differential Equations*. Cambridge University Press; Cambridge, UK: 1996.
- [9]. Estep D, French D. Global error control for the continuous Galerkin finite element method for ordinary differential equations. *RAIRO Modél. Math. Anal. Numér.* 1994; 28:815–852.
- [10]. Estep D, Ginting V, Shadid J, Tavener S. An a posteriori-a priori analysis of multiscale operator splitting. *SIAM J. Numer. Anal.* 2008; 46:1116–1146.
- [11]. Estep D, Holst MJ, Målqvist A. Nonparametric density estimation for randomly perturbed elliptic problems III: Convergence, complexity, and generalizations. *J. Appl. Math. Comput.* to appear.
- [12]. Estep D, Larson MG, Williams RD. Estimating the error of numerical solutions of systems of reaction-diffusion equations. *Mem. Amer. Math. Soc.* 2000; 146
- [13]. Estep D, Målqvist A, Tavener S. Nonparametric density estimation for randomly perturbed elliptic problems I: Computational methods, a posteriori analysis, and adaptive error control. *SIAM J. Sci. Comput.* 2009; 31:2935–2959.
- [14]. Estep D, Målqvist A, Tavener S. Nonparametric density estimation for randomly perturbed elliptic problems II: Applications and adaptive modeling. *J. Numer. Methods Engrg.* 2009; 80:846–867.
- [15]. Estep D, Neckels D. Fast and reliable methods for determining the evolution of uncertain parameters in differential equations. *J. Comput. Phys.* 2006; 213:530–556.
- [16]. Estep D, Neckels D. Fast methods for determining the evolution of uncertain parameters in reaction-diffusion equations. *Comput. Methods Appl. Mech. Engrg.* 2007; 196:3967–3979.
- [17]. Estep D, Stewart A. The dynamical behavior of the discontinuous Galerkin method and related difference schemes. *Math. Comput.* 2002; 71:1075–1103.
- [18]. Estep D. A posteriori error bounds and global error control for approximation of ordinary differential equations. *SIAM J. Numer. Anal.* 1995; 32:1–48.
- [19]. Huelsenbeck JP, et al. Potential applications and pitfalls of Bayesian inference of phylogeny. *Syst. Biol.* 2002; 51:673–688. [PubMed: 12396583]
- [20]. Folland, G. *Real Analysis*. John Wiley & Sons; New York: 1999.
- [21]. Gentle, JE. *Random Number Generation and Monte Carlo Methods*. Springer; New York: 2003.

- [22]. Gilks, WR.; Richardson, S.; Spiegelhalter, DJ. Markov Chain Monte Carlo in Practice. CRC Press; Boca Raton, FL: 1995.
- [23]. Kaipio, J.; Somersalo, E. Statistical and Computational Inverse Problems. Springer; New York: 2005.
- [24]. Knill, DC.; Richards, W. Perception as Bayesian Inference. Cambridge University Press; Cambridge, UK: 1996.
- [25]. Neckels, D.; Colorado State University. Ph.D. thesis, Department of Mathematics. Fort Collins, CO: 2005. Variational Methods for Uncertainty Quantification.
- [26]. Robert, CP.; Casella, G. Monte Carlo Statistical Methods. Springer; New York: 2004.
- [27]. Sandelin, J.; Colorado State University. Ph.D. thesis, Department of Mathematics. Fort Collins, CO: 2006. Global Estimate and Control of Model, Numerical, and Parameter Error.
- [28]. Serfling, RJ. Approximation Theorems of Mathematical Statistics. John Wiley & Sons; New York: 1980.
- [29]. Tarantola, A. Inverse Problem Theory and Methods for Model Parameter Estimation. SIAM; Philadelphia: 2005.

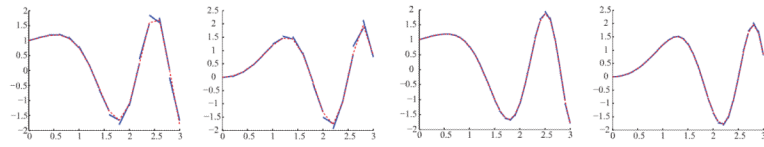


Fig. 4.1. Solutions to (4.1) for $T = 3$. Left two plots: $Y_{\mu,1}$ and $Y_{\mu,2}$ with a time step of 0.2. Right two plots: $Y_{\mu,1}$ and $Y_{\mu,2}$ with a time step of 0.1. The dotted lines indicate the corresponding exact solutions $y_{\mu,1}$ and $y_{\mu,2}$ evaluated on the same time mesh as the dashed-lined numerical approximations $\tilde{Y}_{\mu,1}$ and $\tilde{Y}_{\mu,2}$.

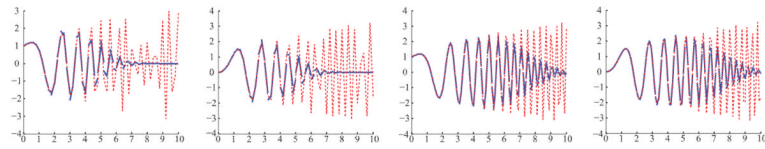


Fig. 4.2. Solutions to (4.1) for $T = 10$. Left two plots: $Y_{\mu,1}$ and $Y_{\mu,2}$ with a time step of 0.2. Right two plots: $Y_{\mu,1}$ and $Y_{\mu,2}$ with a time step of 0.1. The dotted lines indicate the corresponding exact solutions $y_{\mu,1}$ and $y_{\mu,2}$ evaluated on the same time mesh as the dashed-lined numerical approximations $Y_{\mu,1}$ and $Y_{\mu,2}$.

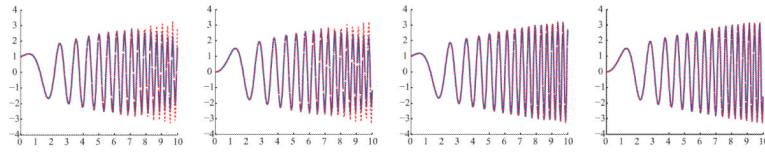


Fig. 4.3.

Solutions to (4.1) for $T = 10$. Left two plots: $Y_{\mu,1}$ and $Y_{\mu,2}$ with a time step of 0.05. Right two plots: $Y_{\mu,1}$ and $Y_{\mu,2}$ with a time step of 0.025. The dotted lines indicate the corresponding exact solutions $y_{\mu,1}$ and $y_{\mu,2}$ evaluated on the same time mesh as the dashed-lined numerical approximations $Y_{\mu,1}$ and $Y_{\mu,2}$.

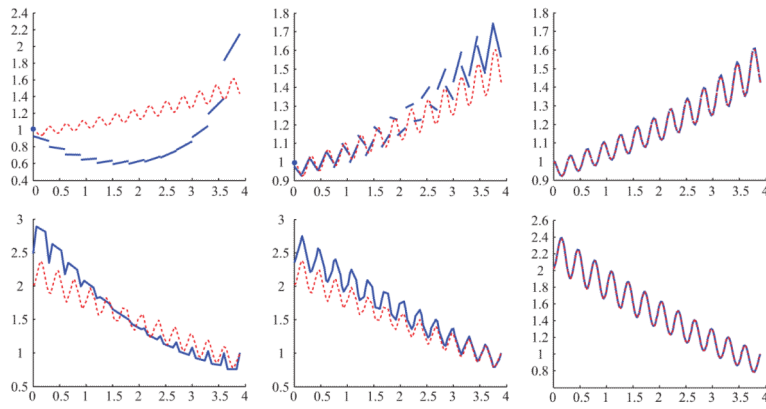


Fig. 4.4. Solutions to (4.3) for $T = 3.9$ with a time step of 0.3 (left), 0.15 (middle), and 0.075 (right). Top plots: Y_μ and y_μ . Bottom plots: Φ and ϕ . The dotted lines indicate the corresponding exact solutions y_μ and ϕ evaluated on the same time mesh as the dashed-lined numerical approximations μ and Φ .

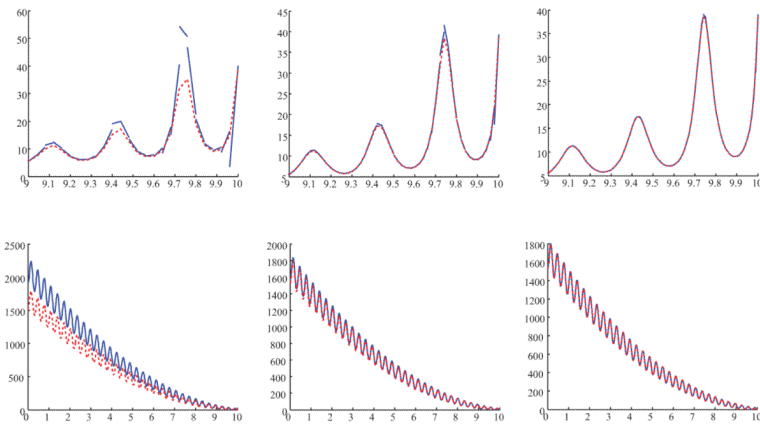


Fig. 4.5. Solutions to (4.3) for $T = 10$ with a time step of 0.04 (left), 0.02 (middle), and 0.01 (right). Top plots: Y_μ and y_μ for time interval $[9, 10]$. Bottom plots: Φ and ϕ . The dotted lines indicate the corresponding exact solutions y_μ and ϕ evaluated on the same time mesh as the dashed-lined numerical approximations Y_μ and Φ .

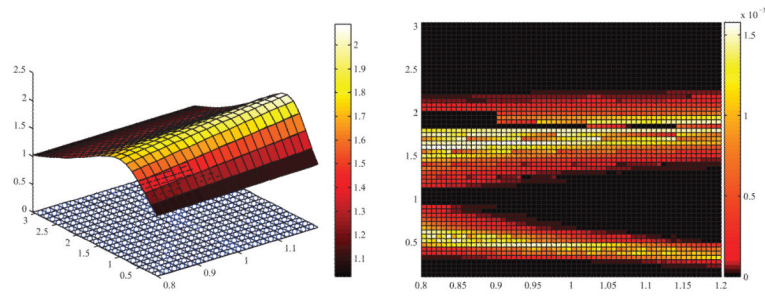


Fig. 4.6.

Left: The global piecewise-linear approximation to $q(\lambda)$ using a coarse 20×20 set of cells. The circles in each cell indicate the reference parameter used to linearize $q(\lambda)$ in that cell. Right: A contour plot of the computed probability distribution on a grid of 60×60 cells corresponding to a normal distribution on $q(\lambda)$.

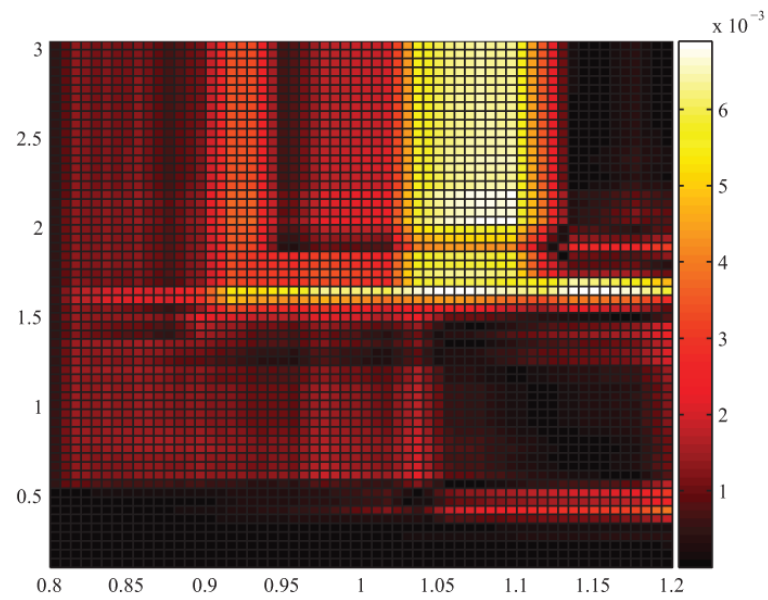


Fig. 4.7.

Plot of absolute values of approximate errors in probabilities over the 60×60 grid of cells used to approximate the solution to the inverse problem. The errors are approximated using a more accurate approximation $\tilde{F}^{(2)}(t)$ computed using a refined numerical solution with a time step of 10^{-2} , resulting in $E < 10^{-7}$, and 10^8 samples of the output density to make statistical errors small. The maximum in this plot is approximately 6.70×10^{-3} , which is less than the computed bound 2.55×10^{-2} .

Table 4.1
Example 1 results for the three partial derivatives of the solution at T = 3

Δt	true 1_q	est. 1_q	Error	IIIab	IIb	Estimate	$\frac{\text{Error}}{\text{Estimate}}$	$\frac{\text{est. } 1_q + \text{Est. true } 1_q}{\text{true } 1_q}$
0.2	1.71E-01	1.62E-01	8.97E-03	9.43E-03	-7.53E-07	9.43E-03	1.052	1.00
0.1	1.71E-01	1.69E-01	1.58E-03	1.47E-03	-7.32E-07	1.47E-03	0.931	0.999
0.05	1.71E-01	1.70E-01	2.21E-04	1.97E-04	2.58E-06	1.99E-04	0.904	1.00
Δt	true 2_q	est. 2_q	Error	IIIab	IIb	Estimate	$\frac{\text{Error}}{\text{Estimate}}$	$\frac{\text{est. } 2_q + \text{Est. true } 2_q}{\text{true } 2_q}$
0.2	-3.17E+00	-2.95E+00	-2.27E-01	-1.76E-01	1.07E-06	-1.76E-01	0.776	0.984
0.1	-3.17E+00	-3.14E+00	-2.91E-02	-2.29E-02	-7.29E-07	-2.29E-02	0.788	0.998
0.05	-3.17E+00	-3.17E+00	-3.54E-03	-2.81E-03	-9.58E-06	-2.82E-03	0.796	1.00
Δt	true 3_q	est. 3_q	Error	IIIab	IIb	Estimate	$\frac{\text{Error}}{\text{Estimate}}$	$\frac{\text{est. } 3_q + \text{Est. true } 3_q}{\text{true } 3_q}$
0.2	5.36E-01	5.30E-01	5.82E-03	4.40E-03	-4.19E-07	4.40E-03	0.757	0.997
0.1	5.36E-01	5.35E-01	7.24E-04	5.59E-04	-6.41E-07	5.58E-04	0.771	1.00
0.05	5.36E-01	5.36E-01	8.69E-05	6.77E-05	9.55E-08	6.78E-05	0.780	1.00
Δt	true 4_q	est. 4_q	Error	IIIab	IIb	Estimate	$\frac{\text{Error}}{\text{Estimate}}$	$\frac{\text{est. } 4_q + \text{Est. true } 4_q}{\text{true } 4_q}$
0.2	2.78E-01	2.45E-01	3.27E-02	3.53E-02	7.92E-07	3.53E-02	1.08	1.01
0.1	2.78E-01	2.72E-01	5.92E-03	5.58E-03	-1.09E-06	5.57E-03	0.941	0.999
0.05	2.78E-01	2.77E-01	8.36E-04	7.50E-04	-8.04E-06	7.41E-04	0.887	1.00

Table 4.2
Example 1 results for the three partial derivatives of the solution at T = 10

Δt	true 1_q	est. 1_q	Error	IIIab	IIb	Estimate	Error/Estimate	est. $1_q + \text{Est. true } 1_q$
0.2	-1.35E-02	6.30E-02	-7.66E-02	9.59E-03	2.20E-05	9.61E-03	-0.126	-5.36
0.1	-1.35E-02	5.75E-02	-7.10E-02	-1.07E-01	-5.60E-06	-1.07E-01	1.51	3.66
0.05	-1.35E-02	9.58E-03	-2.31E-02	-2.36E-02	3.40E-06	-2.36E-02	1.02	1.04
0.025	-1.35E-02	-9.20E-03	-4.39E-03	-3.91E-03	-9.80E-06	-3.92E-03	0.893	0.965
Δt	true 2_q	est. 2_q	Error	IIIab	IIb	Estimate	Error/Estimate	est. $2_q + \text{Est. true } 2_q$
0.2	1.40E+01	-3.39E-01	1.44E+01	1.26E+01	2.07E-04	1.26E+01	0.876	0.873
0.1	1.40E+01	-6.37E-01	1.47E+01	5.82E+00	-3.24E-04	5.82E+00	0.397	0.370
0.05	1.40E+01	7.68E+00	6.34E+00	4.85E+00	-2.87E-04	4.85E+00	0.765	0.894
0.0250	1.40E+01	1.30E+01	9.91E-01	7.88E-01	-2.53E-04	7.87E-01	0.794	0.985
Δt	true 3_q	est. 3_q	Error	IIIab	IIb	Estimate	Error/Estimate	est. $3_q + \text{Est. true } 3_q$
0.2	4.51E-01	4.63E-01	-1.29E-02	-1.14E-02	2.18E-05	-1.14E-02	0.882	1.02
0.1	4.51E-01	4.64E-01	-1.32E-02	-5.13E-03	-9.29E-06	-5.14E-03	0.388	1.02
0.05	4.51E-01	4.56E-01	-5.73E-03	-4.37E-03	-2.20E-06	-4.37E-03	0.763	1.00
0.025	4.51E-01	4.51E-01	-8.98E-04	-7.10E-04	-8.74E-06	-7.19E-04	0.801	1.00
Δt	true 4_q	est. 4_q	Error	IIIab	IIb	Estimate	Error/Estimate	est. $4_q + \text{Est. true } 4_q$
0.2	-9.52E-01	-1.16E-01	-8.37E-01	1.11E-01	2.18E-04	1.11E-01	-0.133	1.36
0.1	-9.52E-01	-1.78E-01	-7.75E-01	-1.17E+00	-3.28E-04	-1.18E+00	1.52	0.458
0.05	-9.52E-01	-7.01E-01	-2.51E-01	-2.58E-01	-2.83E-04	-2.58E-01	1.03	1.01
0.025	-9.52E-01	-9.04E-01	-4.81E-02	-4.27E-02	-2.52E-04	-4.29E-02	0.892	0.995

Table 4.3
Example 2 results for the two partial derivatives of the solution (upper) and the adjoint solution (lower) at $T = 3.9$

Δt	true 1_q	est. 1_q	Error	IIIab	IIb	Estimate	$\frac{\text{Error}}{\text{Estimate}}$	$\frac{\text{est. } 1_q + \text{Est.}}{\text{true } 1_q}$
0.3	-7.89E+00	-4.42E+00	-3.47E+00	-4.06E+00	-1.49E+00	-5.55E+00	1.60	1.26
0.15	-7.89E+00	-9.32E+00	1.43E+00	5.88E-01	8.18E-01	1.41E+00	0.986	1.00
0.075	-7.89E+00	-8.11E+00	2.12E-01	8.05E-02	1.01E-01	1.82E-01	0.858	1.00
Δt	true 2_q	est. 2_q	Error	IIIab	IIb	Estimate	$\frac{\text{Error}}{\text{Estimate}}$	$\frac{\text{est. } 2_q + \text{Est.}}{\text{true } 2_q}$
0.3	-1.94E-01	2.26E-01	-4.20E-01	9.44E-01	1.09E+00	2.03E+00	-4.83	-111.7
0.15	-1.93E-01	5.62E-02	-2.50E-01	-1.91E-01	-5.68E-02	-2.48E-01	0.992	0.990
0.075	-1.93E-01	-1.78E-01	-1.55E-02	-4.55E-03	-1.44E-02	-1.90E-02	1.22	1.02
Δt	$\Phi(0)$	$\Phi(0)$	Error	II	$\frac{\text{Error}}{\text{Estimate}}$	$\frac{\Phi(0) + \text{Est.}}{\Phi(0)}$		
0.3	2.02E+00	2.49E+00	-4.64E-01	1.19E+00	-2.57	1.82		
0.15	2.02E+00	2.35E+00	-3.31E-01	-3.54E-01	1.07	0.988		
0.075	2.02E+00	2.07E+00	-5.03E-02	-4.93E-02	0.981	1.00		

Table 4.4
Example 2 results for the two partial derivatives of the solution (upper) and the adjoint solution (lower) at T = 10.

Δt	1_q (true)	1_q (est)	$\frac{1_q(\text{est})}{1_q(\text{true})}$	True Error	Term IIIab	Term IIb	Error Estimate	Error Ratio	$\frac{1_q(\text{est})+\text{Error Estimate}}{1_q(\text{true})}$
0.04	$-1.52E+04$	$-1.91E+04$	1.25290	$3.85E+03$	$5.24E+02$	$7.08E+03$	$7.60E+03$	1.97601	0.75317
0.02	$-1.52E+04$	$-1.56E+04$	1.02488	$3.78E+02$	$4.93E+01$	$2.60E+02$	$3.10E+02$	0.81829	1.00452
0.01	$-1.52E+04$	$-1.53E+04$	1.00369	$5.61E+01$	$5.98E+00$	$3.61E+01$	$4.21E+01$	0.75009	1.00092
Δt	2_q (true)	2_q (est)	$\frac{2_q(\text{est})}{2_q(\text{true})}$	True Error	Term IIIab	Term IIb	Error Estimate	Error Ratio	$\frac{2_q(\text{est})+\text{Error Estimate}}{2_q(\text{true})}$
0.04	$6.66E+02$	$7.03E+02$	1.05505	$-3.67E+01$	$1.17E+03$	$-1.12E+03$	$4.73E+01$	-1.29028	1.12609
0.02	$6.66E+02$	$6.79E+02$	1.01928	$-1.28E+01$	$8.71E+01$	$-9.40E+01$	$-6.94E+00$	0.54057	1.00886
0.01	$6.66E+02$	$6.68E+02$	1.00328	$-2.19E+00$	$1.02E+01$	$-1.17E+01$	$-1.48E+00$	0.67836	1.00106
Δt	$\Phi(0)$	$\Phi(0)$	$\frac{\Phi(0)}{\Phi(0)}$	True Error	Term II	Error Ratio	$\Phi(0)+\text{Error Estimate}$		
0.04	$1.52E+03$	$1.90E+03$	1.25043	$-3.81E+02$	$-7.63E+02$	2.00188	0.74909		
0.02	$1.52E+03$	$1.56E+03$	1.02465	$-3.75E+01$	$-3.11E+01$	0.83007	1.00419		
0.01	$1.52E+03$	$1.53E+03$	1.00366	$-5.57E+00$	$-4.23E+00$	0.75982	1.00088		