Nuclear Engineering and Design 245 (2012) 49–54

Contents lists available at SciVerse ScienceDirect

Nuclear Engineering and Design

journal homepage: www.elsevier.com/locate/nucengdes

Adjoint-based sensitivity analysis for multi-component models

Hany S. Abdel-Khalik*

Department of Nuclear Engineering, North Carolina State University, Raleigh, NC 27695-7909, United States

ARTICLE INFO

Article history:
Received 23 May 2011
Received in revised form 13 January 2012
Accepted 18 January 2012

ABSTRACT

In typical design calculations, a multi-component model (i.e. a chain of codes) is often employed to calculate the quantity of interest. For design optimization, sensitivity analysis studies are often required to find optimum operating conditions or to propagate uncertainties required to set design margins. This manuscript presents a hybrid approach to enable the transfer of sensitivity information between the various components in an efficient manner that precludes the need for a global sensitivity analysis procedure, often envisaged to be computationally intractable. The presented method has two advantages over existing methods which may be classified into two broad categories: brute force-type methods and amalgamated-type methods. First, the presented method determines the minimum number of adjoint evaluations for each component as opposed to the brute force-type methods which require full evaluation of all sensitivities for all responses calculated by each component in the overall model, which proves computationally prohibitive for realistic problems. Second, the new method treats each component as a black-box as opposed to amalgamated-type methods which requires explicit knowledge of the system of equations associated with each component in order to reach the minimum number of adjoint evaluations. The discussion in this manuscript will be limited to the evaluation of first-order derivatives only. Current work focuses on the extension of this methodology to capture higher order derivatives.

Published by Elsevier B.V.

1. Introduction

The challenges facing sensitivity analysis algorithms continue to grow as engineering design codes become more complex. In particular, a modeling strategy that has found wide application in many engineering disciplines is the so-called multi-scale multi-physics phenomena modeling. In this modeling strategy, several models are employed to describe system behavior starting with detailed first principles fine scale models and ending with coarse scale models to predict the system's macroscopic performance metrics. From a high level, this modeling strategy may be viewed as an assembly of numerous models coupled together in various manners to account for the different scales and physics that affect system behavior. The interconnectivity of the models complicates the manner in which sensitivity information is transferred between the models. In particular, we focus in this manuscript on adjoint sensitivity analysis. Although, powerful adjoint sensitivity analysis tools may exist for the individual scales and/or physics models (often referred to as single-physics or single-scale adjoints, and hereinafter denoted by single-component adjoints), there is often no generally accepted way to formulating a global adjoint for the multi-scale multi-physics model (denoted hereinafter by multi-component model). A global adjoint is often much more complicated to implement and must be planned in advance for the particular set of components' models. Given the dynamic nature of code development and the need to utilize and exchange models frequently, it is paramount to design sensitivity analysis algorithms that can generate sensitivity information for multi-component models from the single-component adjoint. This is a challenging task since a global adjoint for a multi-component model depends on the manner in which the single-components models are connected.

This manuscript proposes a new method to elucidate the coupling of adjoint sensitivity information between different components' models in a multi-component model (Abdel-Khalik et al., 2011). The proposed method combines the advantages of two existing methods for evaluating sensitivity information for a multi-component model: the brute force methods[1] (Jessee et al., 2009a) and the amalgamated methods (Dan Cacuci, 2003). The brute force method is simple to implement but as will be shown in the next section requires significant computational overhead. The amalgamated method minimizes the number of adjoint evaluations but requires great insight into the components' models. Revealed in

[1] In this context, *brute force* does not refer to finite difference methods for evaluation of first-order derivatives. However, it refers to brute forcing the application of adjoint methods to all the components of a multi-component model. To our knowledge, this terminology has not been used before in the literature on global adjoint methods.

* Tel.: +1 919 515 4600; fax: +1 919 515 5115.
E-mail address: abdelkhalik@ncsu.edu

0029-5493/$ – see front matter Published by Elsevier B.V.
doi:10.1016/j.nucengdes.2012.01.017

the next section are the differences between these two methods and the proposed method to combine their advantages.

2. Mathematical description of the problem

Consider a general multi-component model composed of K components (each component's model is implemented in the form of a computer code), where each component's model may represent a different scale and/or physics. This general representation supports the generic manner proposed to transfer information between the various components' models. Let $\vec{\Theta}$ denote the corresponding mathematical operator for the overall model which may be comprised of linear and/or nonlinear equations:

$$\vec{y} = \vec{\Theta}(\vec{x}), \tag{1}$$

where the input data (independent variables) are denoted by an n-tuples vector $\vec{x} \in \mathbb{R}^n$ ($\mathbb{R}^n$ is referred to as the input space); and the output responses (dependent variables) are denoted by a vector $\vec{y} \in \mathbb{R}^m$ ($\mathbb{R}^m$ is the output space). The various components' models comprising the overall model can be described by:

$$\begin{aligned}\vec{y} &= \vec{\Theta}_K(\vec{z}_{K-1}), \vec{z}_{K-1} = \vec{\Theta}_{K-1}(\vec{z}_{K-2}), \ldots, \vec{z}_k = \vec{\Theta}_k(\vec{z}_{k-1}), \ldots, \vec{z}_2 \\ &= \vec{\Theta}_2(\vec{z}_1), \vec{z}_1 = \vec{\Theta}_1(\vec{x}),\end{aligned} \tag{2}$$

where $\vec{z}_k$ is the output (assumed to consist of m_k elements, $\vec{z}_k \in \mathbb{R}^{m_k}$) of the kth component, $\vec{\Theta}_k$, which is passed as input to the next component, $\vec{\Theta}_{k+1}$. Using the chain rule of differentiation, the first-order derivatives of the output responses $\vec{y}$ with respect to the input data $\vec{x}$ are given by:

$$\begin{aligned}&\frac{d\vec{y}}{d\vec{x}} = \frac{d\vec{y}}{d\vec{z}_{K-1}} \times \frac{d\vec{z}_{K-1}}{d\vec{z}_{K-2}} \times \ldots \times \frac{d\vec{z}_k}{d\vec{z}_{k-1}} \times \ldots \times \frac{d\vec{z}_2}{d\vec{z}_1} \times \frac{d\vec{z}_1}{d\vec{x}} \\ &\boldsymbol{\Theta} = \boldsymbol{\Theta}_K \times \boldsymbol{\Theta}_{K-1} \times \ldots \times \boldsymbol{\Theta}_k \times \ldots \times \boldsymbol{\Theta}_2 \times \boldsymbol{\Theta}_1\end{aligned} \tag{3}$$

This is a product of K matrices, each representing the sensitivity matrix associated with one component. For example, the sensitivity matrix $\boldsymbol{\Theta}_k = d\vec{z}_k/d\vec{z}_{k-1}$ contains the sensitivities of the m_k responses calculated by the kth component with respect to its m_{k-1} inputs and therefore has dimensions $m_k \times m_{k-1}$. The matrix $\boldsymbol{\Theta}$, representing the product of all sensitivities matrices, will be denoted hereinafter by the global sensitivity matrix. Evaluating the global sensitivity matrix with minimized executions of each component's model represents the goal of this paper.

Two types of methods may be utilized to determine the global sensitivity matrix. The first method, denoted hereinafter by the brute force method, determines the sensitivity matrix associated with each component's model, and then performs the K-matrix products in Eq. (3). In this method, one could use either the forward or the adjoint sensitivity analysis approach for each component, depending on which approach would be computationally more favorable.[2] To simplify the initial discussion, we assume that the adjoint approach is considered more favorable for all the components. This assumption will be relaxed later in the discussion. With only the adjoint SA employed, each component model has to be executed in an adjoint mode a number of times equal to the number of its output responses, thereby requiring a total number of adjoint evaluations for all components that is given by:

$$\begin{aligned}\text{total adjoint evals}_{\text{brute force}} &= m + m_{K-1} + \ldots + m_k \\ &\quad + \ldots + m_2 + m_1\end{aligned} \tag{4}$$

The second method, referred to hereinafter as the amalgamated method, recognizes that if the equations describing each component's model are available, one could combine all equations to form one overall model with m output responses. Given the equations describing the overall model, one could formulate the associated global adjoint problem which would require only m adjoint executions of the overall model to evaluate the global sensitivity matrix. If implemented carefully, each adjoint execution of the overall model implies one adjoint execution for each component. With K components comprising the overall model, the total number of adjoint evaluations (as compared to the brute force approach) would be:

$$\text{total adjoint evals}_{\text{amalgamated}} = K \times m \tag{5}$$

which is independent of the number of output responses of the first $K-1$ components. To achieve that, one needs to write down all the equations corresponding to the overall model, set up the associated adjoint problem, and break it down into K adjoint problems, each associated with a component's model as done by amalgamated-type methods (Dan Cacuci, 2003).

We present below an alternative approach to reach the same goal but without access to the equations comprising each of the K components. This approach will be developed in the following two subsections for two different scenarios; with the first designed for well-conditioned sensitivity matrices, and the other when the global sensitivity matrix is expected to be poorly conditioned, i.e. with a very high condition number. Earlier work has shown that sensitivity matrices associated with reactor calculations are poorly conditioned, and has demonstrated that one may exploit this poor conditioning to reduce the amount of computational burden required to propagate sensitivity information (Jessee et al., 2009b).

2.1. Well-conditioned global sensitivity matrix

To reveal the mechanics of the proposed method, we need to recall some properties of the adjoint sensitivity analysis approach. To avoid clouding the discussion with too many subscripts, we adopt simplified notations in which the kth component's model and its associated inputs and outputs are described by:

$$\vec{u} = \vec{\Theta}(\vec{v}) \tag{6}$$

where $\vec{z}_k \in \mathbb{R}^{m_k}$, $\vec{z}_{k-1} \in \mathbb{R}^{m_{k-1}}$, and $\vec{\Theta}_k$ are replaced, respectively, by $\vec{u} \in \mathbb{R}^m$, $\vec{v} \in \mathbb{R}^n$, and $\vec{\Theta}$. The sensitivity matrix $\boldsymbol{\Theta}$ associated with this component's model is given by:

$$[\boldsymbol{\Theta}]_{ij} = \frac{du_i}{dv_j}, \quad \text{and} \quad \boldsymbol{\Theta} \in \mathbb{R}^{m \times n} \tag{7}$$

In the adjoint sensitivity analysis approach, the above system of equations is broken up into two systems of equations, one for the state dependent variable $\vec{\phi}$ and one for the response of interest, written below for the ith response u_i as:

$$\vec{\Pi}(\vec{v}, \vec{\phi}) = \vec{Q}(\vec{v}); \quad \text{and} \quad u_i = u_i(\vec{v}, \vec{\phi}) \tag{8}$$

A change in the input data $\vec{v}$ induces changes in the operator $\vec{\Pi}$ and the right hand side $\vec{Q}$, which perturbs the state dependent variable $\vec{\phi}$ and consequently the response u_i. In the adjoint sensitivity analysis approach one avoids solving for the state dependent variable perturbations by setting up an adjoint problem of the form:

$$\left(\frac{\partial \vec{\Pi}(\vec{v}, \vec{\phi})}{\partial \vec{\phi}}\right)^* \vec{\phi}_i^* = \frac{\partial u_i(\vec{v}, \vec{\phi})}{\partial \vec{\phi}} \tag{9}$$

[2] The forward sensitivity analysis suits problems with few input data and many output responses. The derivatives are calculated using a finite differencing approach by running the code as many times as the number of input data. In each run, one input data is perturbed and the corresponding derivatives of all responses with respect to the perturbed input data are calculated. The adjoint sensitivity analysis is employed for problems with few responses and many input data. In this approach, the derivatives of one response with respect to all input data could be calculated in one execution of the adjoint code. The adjoint code is executed once for each response of interest.