Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 17-18, 2012

Operator Decomposition Framework for Perturbation Theory

Hany S. Abdel-Khalik, Congjian Wang, and Youngsuk Bang
Department of Nuclear Engineering, North Carolina State University
{abdelkhalik, cwang21, ybang}@ncsu.edu

1. Introduction

This summary describes a new framework for perturbation theory intended to improve its performance, in terms of the associated computational cost and the complexity of implementation, for routine reactor calculations in support of design, analysis, and regulation. Since its first introduction in reactor analysis by Winger [1], perturbation theory has assumed an aura of sophistication with regard to its implementation and its capabilities. Only few reactor physicists, typically mathematically proficient, have contributed to its development, with the general body of the nuclear engineering community remaining unaware of its current status, capabilities, and challenges. Given its perceived sophistication and the small body of community users, the application of perturbation theory has been limited to investigatory analyses only. It is safe to say that the nuclear community is split into two groups, a small one which understands the theory and, and a much bigger group with the perceived notion that perturbation theory is nothing but a fancy mathematical approach that has very little use in practice.

Over the past three years, research has demonstrated two goals [2]. First, reduce the computational cost of perturbation theory in order to enable its use for routine reactor calculations. Second, expose some of the myth about perturbation theory and present it in a form that is simple and relatable in order to stimulate the interest of nuclear practitioners, especially those who are currently working on the development of next generation reactor design and analysis tools.

The operator decomposition approach has its roots in linear algebra and can be easily understood by code developers, especially those involved in the design of iterative numerical solution strategies.

2. Perturbation Theory Background

Perturbation theory addresses the following problem. Consider a model with *n* input parameters which are represented by an *n* component vector, $p \in \mathbb{R}^n$. The model calculates a solution, referred to as the state of the engineering system. In neutronics, the flux represents the state; in thermal analysis, the temperature distribution plays the role of the state. Mathematically, this is described as follows:

$$\Theta(p,\phi) = 0 \qquad (1)$$

This represents a system of algebraic equations, solved numerically for the state represented by a vector of *k* components, $\phi \in \mathbb{R}^k$. The state rarely represents the goal of the simulation. However, one calculates the engineering quantities of interest (QoI) as functions of the state and parameters, described in general by:

$$\eta = \eta(p,\phi) \qquad (2)$$

where $\eta \in \mathbb{R}^m$ is a vector representing the *m* QoIs. The multiple execution of the model is a typical requirement for any real-world engineering analysis. For example, consider the neutronics model describing the radiation transport in a given fuel lattice. One needs to execute this model in the order of 10^4 times to account for different lattice designs, fuel composition change due to depletion, spectrum changes due to fuel temperature and thermal-hydraulics conditions. Another application is the propagation of uncertainties, where one tries to quantify the impact of parameters uncertainties on the QoIs. Mathematically, perturbation theory describes this situation as follows: for a given input parameter perturbation, Δp, find the corresponding variation in the QoI, $\Delta\eta$. This formulation implies that one does not need to know the corresponding state variation, only the QoI variation is sought. There are two general ways to solve this problem. The first one is denoted by the forward approach. As the name implies, it re-executes the model with the parameter perturbation and calculates the state variation, then the variation in QoI. In the adjoint approach however, one can show that by calculating another function called the adjoint state, one can estimate directly the QoI variation without having to calculate the state variation. If this assertion is valid in general, one can treat all required executions of the model for design calculations as simple perturbations from some reference state, and save the huge computational cost that would be otherwise required by the forward approach. Unfortunately, this assertion is not generally true. It is only true first-order variations implying that nonlinear variations are ignored, i.e. assumed negligible. This limits the use of perturbation theory to problems where only small variations are of interest, such as in the case of uncertainty analysis for neutronics model. Our objective is to design a framework that enables perturbation theory to calculate both the linear and nonlinear variations in the QoI with inexpensive computational cost.

Mathematically, assume Eq. 1 and 2 take the form:[1]

$$\boldsymbol{\Theta}(p)\phi = \pi(p) \quad \text{and} \quad \eta = \varpi^T\phi \qquad (3)$$

The symbol $\boldsymbol{\Theta} \in \mathbb{R}^{k\times k}$ is a matrix operator representing the numerically descritized continuous

[1] For simplicity, assume a source driven model, the operator acts linearly on the state, and no source perturbations.

operator in Eq. 1 with the appropriate boundary conditions. A single QoI is described by an inner product of some vector ϖ with the state. Solving for the state, one obtains:

$$\phi = \mathbf{\Theta}^{-1}\pi \quad \text{and} \quad \eta = \varpi^T \mathbf{\Theta}^{-1}\pi \tag{4}$$

Consider a perturbation in the parameters:

$$\Delta\eta = \varpi^T \left\{ \left(\mathbf{\Theta} + \Delta\mathbf{\Theta}\right)^{-1}\pi - \mathbf{\Theta}^{-1}\pi \right\} \tag{5}$$

One can expand the $\left(\mathbf{\Theta} + \Delta\mathbf{\Theta}\right)^{-1}$ term yielding:

$$\Delta\eta = -\varpi^T \mathbf{\Theta}^{-1} \left\{ \mathbf{I} - \left[\Delta\mathbf{\Theta}\mathbf{\Theta}^{-1}\right] + \left[\Delta\mathbf{\Theta}\mathbf{\Theta}^{-1}\right]^2 + ... \right\} \Delta\mathbf{\Theta}\phi \tag{6}$$

To estimate the first order variations of the QoI, ignore all terms that are higher than one to obtain:

$$\Delta^1\eta = -\varpi^T \mathbf{\Theta}^{-1}\Delta\mathbf{\Theta} \times \mathbf{\Theta}^{-1}\pi$$

In practice, one needs to calculate this variation for many parameters. To accomplish this, note that this equation could be interpreted in two different ways, mathematically described as follows:

$$\Delta^1\eta = -\varpi^T \left\{ \mathbf{\Theta}^{-1} \left(\Delta\mathbf{\Theta} \times \mathbf{\Theta}^{-1}\pi \right) \right\}$$

or

$$\Delta^1\eta = -\left(\mathbf{\Theta}^{-T}\varpi\right)^T \left(\Delta\mathbf{\Theta} \times \mathbf{\Theta}^{-1}\pi \right)$$

The first expression implies that for a given parameter variation, one needs to calculate what is between the round brackets first, then apply the inverse of the operator, implying a forward model execution, then take the inner product with the QoI's vector. This forward approach requires *n* forward model executions for a model with *n* input parameters.

The second expression however requires that one to calculate the inverse of the operator transpose times the QoI's vector. This is equivalent to executing the adjoint model once. The resulting vector is denoted the adjoint state. The QoI variation reduces to a simple inner product between the adjoint state and the variations of the terms in the round bracket. In doing so, one executes the forward and adjoint models one time each at the reference conditions.

Let's consider second order variations, Eq. 6 gives:

$$\Delta^2\eta = \left\{ \mathbf{\Theta}^{-T} \left[\Delta\mathbf{\Theta}^T \left(\mathbf{\Theta}^{-T}\varpi \right) \right] \right\}^T \Delta\mathbf{\Theta}\mathbf{\Theta}^{-1}\pi$$

First, evaluate $\mathbf{\Theta}^{-T}\varpi$ which is a single adjoint model execution at reference conditions. For a given perturbation, form the matrix $\Delta\mathbf{\Theta}$ and calculate its product with the reference adjoint. Evaluate the matrix-inverse-transpose product with the resulting vector (what is indie the curly brackets). This is equivalent to an adjoint model execution per perturbation. The result is then multiplied by $\Delta\mathbf{\Theta}\mathbf{\Theta}^{-1}\pi$ which requires a single forward model execution at reference conditions. In this approach, a total of *n*+1 adjoints and 1 forward model executions are required to get the second order variations. The cost for higher order variations using the adjoint approach increases as n^{s-1}, where *s* is the order of variation.

3. Proposed Framework

The proposed research is based on few observations. First, the state while living in a *k* dimensional space, the state variations from all possible parameters variations have been reported to live in a subspace of much smaller dimension $r \ll k$ [2-3,5]. Second, the rank *r* is much smaller than the number of responses and input parameters often associated with reactor models. One could rigorously calculate the error ε_{user} resulting from reducing the state to live in a subspace [3]. Mathematically, this could be described as follows:

$$\left\| \phi^{(r)} - \phi \right\| = \left\| \left(\mathbf{I} - \mathbf{Q}_r \mathbf{Q}_r^T \right) \phi \right\| \le \varepsilon_{user}$$

The ε_{user} can be made as small as the precision of the forward model calculations. This expression implies that the reduced flux may be expressed as linear combination of the *r* columns of the matrix $\mathbf{Q}_r$. This matrix is calculated using a statistical snapshots approach, which records all possible state variations.

Now, recalling a simple relationship from linear algebra, the so-called Sherman Morrison inversion [4]:

$$\left(\mathbf{\Theta} + uv^T\right)^{-1} = \mathbf{\Theta}^{-1} - \frac{\mathbf{\Theta}^{-1}uv^T\mathbf{\Theta}^{-1}}{1 + v^T\mathbf{\Theta}^{-1}u}$$

The $\mathbf{\Theta}^{-1}$ operator denotes the forward model with no perturbation. Consider that the rank-one update matrix uv^T represents a perturbation due to a given parameter perturbation. This relationship assures that the state will be modified from the reference value by a vector that is pointing in the direction $\mathbf{\Theta}^{-1}u$. This observation represents the core of the proposed framework. Since the flux varies only along a subspace of rank *r*, this implies all the operator perturbations in the form of matrices must have a combined range of dimension *r* since the matrix $\mathbf{\Theta}$ is nonsingular. This simple observation could be rigorously employed to recast the Eq. 6 in a much simpler format. Moreover, one can show that all higher order variations can be calculated, i.e. evaluating the entire infinite series expansion, rather than one term at a time, with only *r* adjoint model evaluations and *r* forward model executions. Numerical results of this proposed framework applied to representative nuclear reactor models will be presented at the meeting.

REFERENCES

[1] E.P. Wigner, "Effect of Small Perturbations on Pile Period", Manhattan Project Report CP-G-3048, 1945.

[2] C.J. Wang and H.S. Abdel-Khalik, Exact-to-Precision Generalized Perturbation Theory for Source Driven Problems, Nuclear Engineering and Design, September 2011.

[3] N. Halko, P.G. Martinsson, and J. Tropp, Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions, SIAM Review, 53(2) , 2011.

[4] J. Sherman and W.J. Morrison, Adjustment of an Inverse Matrix Corresponding to Changes in the Elements of a Given Column or a Given Row of the Original Matrix, Annals of Mathematical Statistics, 1949.