SECOND-ORDER ADJOINT SENSITIVITY ANALYSIS METHODOLOGY (2nd-ASAM) FOR COMPUTING EXACTLY AND EFFICIENTLY FIRST- AND SECOND-ORDER SENSITIVITIES IN LARGE-SCALE LINEAR SYSTEMS: I. COMPUTATIONAL METHODOLOGY

Dan G. Cacuci

Department of Mechanical Engineering, University of South Carolina

E-mail: cacuci@cec.sc.edu

Corresponding author:

Department of Mechanical Engineering, University of South Carolina

300 Main Street, Columbia, SC 29208, USA

Email: cacuci@cec.sc.edu; Phone: (919) 909 9624; Submitted to JCP: August 14, 2014

ABSTRACT

This work presents the *second-order forward and adjoint sensitivity analysis methodologies* (2nd-FSAM and 2nd-ASAM) for computing exactly and efficiently the second-order functional derivatives of physical (engineering, biological, etc.) system responses (i.e., “system performance parameters”) to the system’s model parameters. The definition of “system parameters” used in this work includes all computational input data, correlations, initial and/or boundary conditions, etc. For a physical system comprising N_α parameters and N_r responses, we note that the 2nd-FSAM requires a total of $(N_\alpha^2/2 + 3N_\alpha/2)$ large-scale computations for obtaining all of the first- and second-order sensitivities, for all N_r system responses. On the other hand, for one functional-type system response, the 2nd-ASAM requires one large-scale computation using the first-level adjoint sensitivity system for obtaining all of the first-order sensitivities, followed by at most N_α large-scale computations using the second-level adjoint sensitivity systems for obtaining exactly all of the second-order sensitivities. Therefore, the 2nd-FSAM should be used when $N_r \gg N_\alpha$, while the 2nd-ASAM should be used when $N_\alpha \gg N_r$. The original 2nd-ASAM presented in this work should enable the hitherto very difficult, if not intractable, *exact* computation of *all* of the second-order response sensitivities (i.e., functional Gateaux-derivatives) for large-systems involving many parameters, as usually encountered in practice. Very importantly, the implementation of the 2nd-ASAM requires very little additional effort beyond the construction of the adjoint sensitivity system needed for computing the first-order sensitivities.

KEYWORDS: second-order adjoint sensitivity analysis methodology (2nd-ASAM); exact and efficient computation of first- and second-order functional derivatives; large-scale systems.

1. INTRODUCTION

This work presents a new method for computing, exactly and efficiently, *second-order functional derivatives of system responses* (i.e., “system performance parameters” in physical, engineering, biological systems) *to the system’s parameters*, which are considered in this work in the most comprehensive sense, i.e., including all input data, correlations, initial and/or boundary conditions, etc. This new method builds on the *first-order adjoint sensitivity analysis methodology* (denoted in this work as 1^{st} -ASAM) for linear and nonlinear systems introduced by Cacuci in [1] and [2], and further developed in [3-5]; see also [6]. The aims and means of sensitivity theory/analysis are occasionally confused with the aims and means of optimization theory; such confusions usually arise because of some shared terminology, and because some of the mathematical derivations underlying sensitivity theory appear (superficially) to be similar to those underlying optimization theory. To delineate clearly the aims this work, it is therefore useful to recall briefly the terminology and aims of sensitivity theory/analysis, while highlighting the distinctions between sensitivity analysis and optimization theory.

In general, a physical system and/or the result of an indirect experimental measurement are modeled mathematically in terms of:

- (a) A system of linear and/or nonlinear equations that relate the system's independent variables and parameters to the system's state (i.e., dependent) variables;
- (b) Probability distributions, moments thereof, inequality and/or equality constraints that define the range of variations of the system's parameters;
- (c) One or several quantities, customarily referred to as system responses (or objective functions, or indices of performance), which are computed using the mathematical model.

Sensitivity theory/analysis aims at computing the functional derivatives of system responses to the system's parameters, while optimization theory aims at optimizing (minimizing or maximizing in some user-defined sense) the system responses of interest. Therefore, both sensitivity theory and optimization compute first- and occasionally second-order response derivatives with respect to the state functions and/or parameters. The 1^{st} -ASAM [1] uses the solution of certain adjoint equations for computing efficiently all of the first-order response sensitivities. Adjoint equations also appear in constrained optimization algorithms, particularly in those used in control theory and for data assimilation in the atmospheric and earth sciences (see, e.g., [7] and [8] for recent expositions), where the Lagrange multipliers used to append the constraints to the “objective” response functional are determined as the solutions of certain adjoint equations. However, in contrast to optimization theory, in which some or all of the first-order response derivatives are driven by the respective computational algorithm to zero, sensitivity theory evaluates first- (and occasionally higher-) order response derivatives with respect to system parameters at the parameters' and state-functions' nominal values. Thus, the adjoint systems/equations found in constrained optimization and, respectively, sensitivity analysis serve different purposes and differ from each other conceptually. These important distinctions have been thoroughly analyzed [9] in the setting of both local and global optimization and, respectively, local and global sensitivity analysis. Since this paper deals exclusively with sensitivity theory and analysis, the vast literature dealing with the computation of “cost-functional response” derivatives for the purposes of “cost functional response” optimization will not be discussed further.

Response sensitivities to model parameters are needed in many activities, including:

- (i) understanding the system by identifying and ranking the importance of model parameters in influencing the response under consideration;
- (ii) determining the effects of parameter variations on the system's behavior;

- (iii) improving the system design, possibly reducing conservatism and redundancy;
- (iv) prioritizing possible improvements for the system under consideration;
- (v) quantifying uncertainties in responses due to quantified parameter uncertainties (e.g., by using the method of “propagation of uncertainties” [4]);
- (vi) performing “predictive modeling” [10], including data assimilation and model calibration, for the purpose of obtaining best-estimate predicted results with reduced predicted uncertainties.

The work presented in this article is organized as follows: Section 2 briefly recalls the first-order adjoint sensitivity analysis methodology (1^{st} -ASAM), which was introduced in [1-2] and which provides the foundation for the new *second-order adjoint sensitivity analysis methodology* (2^{nd} -ASAM) presented in Section 3. The 2^{nd} -ASAM computes exactly and most efficiently all of the second-order response sensitivities, requiring only at most $O(N_\alpha)$, rather than in $O(N_\alpha^2)$, large-scale computations. Section 4 will review briefly the main characteristics of other methods (both deterministic and statistical) currently used for computing second-order response sensitivities, highlighting the fundamental distinctions between these methods and the 2^{nd} -ASAM. This comparative discussion will highlight the reasons why the 2^{nd} -ASAM is the most efficient general method to date for computing *exactly all* of the second-order sensitivities. Section 5 concludes this work by highlighting the significance of the 2^{nd} -ASAM and the main directions I am currently pursuing for further generalizing the 2^{nd} -ASAM. In Part II [11], the sequel to this work will present an illustrative paradigm application of the 2^{nd} -ASAM to a particle diffusion problem.

2. BACKGROUND: THE FIRST-ORDER ADJOINT SENSITIVITY ANALYSIS METHODOLOGY (I^{st} -ASAM) FOR LARGE-SCALE LINEAR SYSTEMS

Consider that the physical system is represented mathematically by means of K_u coupled linear operator equations of the form

$$\sum_{j=1}^{K_u} L_{ij}(\mathbf{a}) u_j(\mathbf{x}) = Q_i(\mathbf{a}), \quad i = 1, \dots, K_u, \quad \mathbf{x} \in \Omega_x, \quad (1.a)$$

which can be written in matrix form as

$$\mathbf{L}[\mathbf{a}(\mathbf{x})] \mathbf{u}(\mathbf{x}) = \mathbf{Q}[\mathbf{a}(\mathbf{x})], \quad \mathbf{x} \in \Omega_x, \quad (1.b)$$

where:

1. $\mathbf{x} = (x_1, \dots, x_{J_x})$ denotes the J_x -dimensional phase-space position vector for the primary system; $\mathbf{x} \in \Omega_x \subset \mathbb{R}^{J_x}$, where Ω_x is a subset of the J_x -dimensional real vector space $\mathbb{R}^{J_x}$;

2. $\mathbf{u}(\mathbf{x}) = [u_1(\mathbf{x}), \dots, u_{K_u}(\mathbf{x})]$ denotes a K_u -dimensional column vector whose components are the system's dependent (i.e., state) variables; $\mathbf{u}(\mathbf{x}) \in \mathcal{E}_u$, where $\mathcal{E}_u$ is a normed linear space over the scalar field $\mathcal{F}$ of real numbers;

3. $\mathbf{a}(\mathbf{x}) = [\alpha_1(\mathbf{x}), \dots, \alpha_{N_\alpha}(\mathbf{x})]$ denotes an N_α -dimensional column vector whose components are the system's parameters; $\mathbf{a} \in \mathcal{E}_\alpha$, where $\mathcal{E}_\alpha$ is also a normed linear space;

4. $\mathbf{Q}[\mathbf{a}(\mathbf{x})] = [Q_1(\mathbf{a}), \dots, Q_{K_u}(\mathbf{a})]$ denotes a K_u -dimensional column vector whose elements represent inhomogeneous source terms that depend either linearly or nonlinearly on $\mathbf{a}$; $\mathbf{Q} \in \mathcal{E}_Q$, where $\mathcal{E}_Q$ is also a normed linear space; the components of $\mathbf{Q}$ may be operators, rather than just functions, acting on $\mathbf{a}(\mathbf{x})$ and $\mathbf{x}$;

5. $\mathbf{L} \equiv [L_1(\mathbf{a}), \dots, L_{K_u}(\mathbf{a})]$ denotes a K_u -component column vector whose components are *operators* (including differential, difference, integral, distributions, and/or infinite matrices) acting *linearly* on $\mathbf{u}$ and *nonlinearly* on $\mathbf{a}$. For notational convenience, all

vectors in this work are considered to be column vectors; transposition will be indicated by a dagger ($\dagger$).

In view of the definitions given above, $\mathbf{L}$ represents the mapping $\mathbf{L}:\mathcal{D} \subset \mathcal{E} \rightarrow \mathcal{E}_Q$, where $\mathcal{D} = \mathcal{D}_u \times \mathcal{D}_\alpha$, $\mathcal{D}_u \subset \mathcal{E}_u$, $\mathcal{D}_\alpha \subset \mathcal{E}_\alpha$, and $\mathcal{E} = \mathcal{E}_u \times \mathcal{E}_\alpha$. Note that an arbitrary element $\mathbf{e} \in \mathcal{E}$ is of the form $\mathbf{e} = (\mathbf{u}, \boldsymbol{\alpha})$. If differential operators appear in Eq. (1), then a corresponding set of boundary and/or initial conditions (which are essential to define $\mathcal{D}$) must also be given. Since we consider here only systems that are linear in the state function $\mathbf{u}(\mathbf{x}) \in \mathcal{E}_u$, the accompanying boundary and/or initial conditions must also be linear in $\mathbf{u}(\mathbf{x})$, so they can therefore be represented in operator form as

$$\left[\mathbf{B}(\boldsymbol{\alpha})\mathbf{u} - \mathbf{A}(\boldsymbol{\alpha}) \right]_{\partial\Omega_x} = \mathbf{0}, \quad \mathbf{x} \in \partial\Omega_x, \quad (2)$$

where $\partial\Omega_x$ denotes the boundary of Ω_x , the operator $\mathbf{B}(\boldsymbol{\alpha})$ acts *linearly* on $\mathbf{u}$ but *nonlinearly* on the model parameters $\boldsymbol{\alpha}$, and $\mathbf{A}(\boldsymbol{\alpha})$ denotes an operator that acts *nonlinearly* on $\boldsymbol{\alpha}$.

The vector-valued function $\mathbf{u}(\mathbf{x})$ is considered to be the unique nontrivial solution of the physical problem described by Eqs. (1) and (2). The system response (i.e., result of interest), associated with the problem modeled by Eqs. (1) and (2) will be denoted here as $R(\mathbf{u}, \boldsymbol{\alpha})$; in this work, $R(\mathbf{u}, \boldsymbol{\alpha})$ is considered to be a real-valued *nonlinear functional* of $(\mathbf{u}, \boldsymbol{\alpha})$, which can be generally represented in operator form as

$$R(\mathbf{u}, \boldsymbol{\alpha}):\mathcal{D}_R \subset \mathcal{E} \rightarrow \mathcal{F}, \quad (3)$$

where $\mathcal{F}$ denotes the field of real scalars.

The nominal parameter values $\boldsymbol{\alpha}^0(\mathbf{x})$ are used in Eqs. (1) and (2) to obtain the nominal solution $\mathbf{u}^0(\mathbf{x})$ by solving these equations; mathematically, therefore, the nominal value $\mathbf{u}^0(\mathbf{x})$ of the state-function is obtained by solving

$$\mathbf{L}(\boldsymbol{\alpha}^0)\mathbf{u}^0 = \mathbf{Q}(\boldsymbol{\alpha}^0), \quad \mathbf{x} \in \Omega_x, \quad (4)$$

$$\mathbf{B}(\boldsymbol{\alpha}^0)\mathbf{u}^0 = \mathbf{A}(\boldsymbol{\alpha}^0), \quad \mathbf{x} \in \partial\Omega_x. \quad (5)$$

Equations (4) and (5) represent the “*base-case*” or *nominal* state of the physical system; the superscript “zero” will be used in this work to denoted “nominal” values . After solving Eqs. (4) and (5), the *nominal solution*, $\mathbf{u}^0(\mathbf{x})$, thus obtained is used to obtain the *nominal value*, $R(\mathbf{e}^0)$, of the response $R(\mathbf{e})$, using the nominal values $\mathbf{e}^0 = (\mathbf{u}^0, \boldsymbol{\alpha}^0) \in \mathcal{E}$ of the model’s state function and parameters.

Next, we consider (a vector of) arbitrary variations $\mathbf{h} \equiv (\mathbf{h}_u, \mathbf{h}_\alpha) \in \mathcal{E} = \mathcal{E}_u \times \mathcal{E}_\alpha$, with $\mathbf{h}_u \equiv (\delta u_1, \dots, \delta u_{N_u}) \in \mathcal{E}_u$ and $\mathbf{h}_\alpha \equiv (\delta \alpha_1, \dots, \delta \alpha_{N_\alpha}) \in \mathcal{E}_\alpha$, around $\mathbf{e}^0 = (\mathbf{u}^0, \boldsymbol{\alpha}^0) \in \mathcal{E}$. The variation (sensitivity) of the response R to variations $\mathbf{h}$ in the system parameters is given by the Gâteaux- (G)-differential $\delta R(\mathbf{e}^0; \mathbf{h})$ of the response $R(\mathbf{e})$ at $\mathbf{e}^0 = (\mathbf{u}^0, \boldsymbol{\alpha}^0)$ with increment $\mathbf{h}$, which is defined as

$$\delta R(\mathbf{e}^0; \mathbf{h}) \equiv \left\{ \frac{d}{d\varepsilon} \left[R(\mathbf{e}^0 + \varepsilon \mathbf{h}) \right] \right\}_{\varepsilon=0} = \lim_{\varepsilon \rightarrow 0} \frac{R(\mathbf{e}^0 + \varepsilon \mathbf{h}) - R(\mathbf{e}^0)}{\varepsilon}, \quad (6)$$

for $\varepsilon \in \mathcal{I}$, and all (i.e., arbitrary) vectors $\mathbf{h} \in \mathcal{E}$. When the response $R(\mathbf{e})$ is functional of the form $R: \mathcal{D}_R \rightarrow \mathcal{I}$, the sensitivity $\delta R(\mathbf{e}^0; \mathbf{h})$ is also an operator, defined on the same domain, and with the same range as $R(\mathbf{e})$. The G-differential $\delta R(\mathbf{e}^0; \mathbf{h})$ is related to the total variation $\left[R(\mathbf{e}^0 + \varepsilon \mathbf{h}) - R(\mathbf{e}^0) \right]$ of R at $\mathbf{e}^0$ through the relation $R(\mathbf{e}^0 + \varepsilon \mathbf{h}) - R(\mathbf{e}^0) = \delta R(\mathbf{e}^0; \mathbf{h}) + \Delta(\mathbf{h})$, with $\lim_{\varepsilon \rightarrow 0} [\Delta(\varepsilon \mathbf{h})] / \varepsilon = 0$.

As discussed in [1] and [2], *the most general definition of the first-order sensitivity of a response to variations in the model parameter is the G-differential $\delta R(\mathbf{e}^0; \mathbf{h})$ defined in Eq. (6). Since the system’s state vector $\mathbf{u}$ and parameters $\boldsymbol{\alpha}$ are related to each other through Eqs.*

(1) and (2), it follows that $\mathbf{h}_u$ and $\mathbf{h}_\alpha$ are also related to each other. Therefore, the sensitivity $\delta R(\mathbf{e}^0; \mathbf{h})$ of $R(\mathbf{e})$ at $\mathbf{e}^0$ can only be evaluated after determining the vector of variations $\mathbf{h}_u$ in terms of the vector of parameter variations $\mathbf{h}_\alpha$. The first-order relationship between $\mathbf{h}_u$ and $\mathbf{h}_\alpha$ is determined by taking the G-differentials of Eqs. (1) and (2). Thus, taking the G-differential at $\mathbf{e}^0$ of Eq. (1) yields

$$\sum_{j=1}^{K_u} L_{ij}(\mathbf{a}^0) \delta u_j(\mathbf{x}) = \sum_{k=1}^{N_\alpha} \left[\frac{\partial Q_i(\mathbf{a}^0)}{\partial \alpha_k} - \sum_{j=1}^{K_u} \frac{\partial L_{ij}(\mathbf{a}^0)}{\partial \alpha_k} u_j^0(\mathbf{x}) \right] \delta \alpha_k, \quad i = 1, \dots, K_u, \quad \mathbf{x} \in \Omega_x, \quad (7.a)$$

which can be written compactly in the following block-matrix form

$$\mathbf{L}(\mathbf{a}^0) \mathbf{h}_u = \delta \mathbf{Q}(\mathbf{a}^0; \mathbf{h}_\alpha) - \delta \left[\mathbf{L}(\mathbf{a}^0) \mathbf{u}^0; \mathbf{h}_\alpha \right], \quad \mathbf{x} \in \Omega_x. \quad (7.b)$$

The boundary conditions associated with Eq. (7) are obtained by taking the G-differential at $\mathbf{e}^0$ of the boundary and initial conditions represented by Eq. (2), which yields

$$\mathbf{B}(\mathbf{a}^0) \mathbf{h}_u = \delta \mathbf{A}(\mathbf{a}^0; \mathbf{h}_\alpha) - \delta \left[\mathbf{B}(\mathbf{a}^0) \mathbf{u}^0; \mathbf{h}_\alpha \right], \quad \mathbf{x} \in \partial \Omega_x. \quad (8)$$

Equations (7) and (8) represent the “*first-level forward sensitivity system*” (I^{st} -LFSS), which is also occasionally called the “forward variational model” or the “tangent linear model.” For a given vector of parameter variations $\mathbf{h}_\alpha$ around $\mathbf{a}^0$, the I^{st} -LFSS represented by Eqs. (7) and (8) is solved to obtain $\mathbf{h}_u$. Once $\mathbf{h}_u$ is available, it is in turn used in Eq. (6) to compute the sensitivity $\delta R(\mathbf{e}^0; \mathbf{h})$ of $R(\mathbf{e})$ at $\mathbf{e}^0$, for a given vector of parameter variations $\mathbf{h}_\alpha$. *The direct computation of the response sensitivity $\delta R(\mathbf{e}^0; \mathbf{h})$ by using the ($\mathbf{h}_\alpha$ -dependent) solution $\mathbf{h}_u$ of the I^{st} -LFSS is called [1] the (first-order) forward sensitivity analysis method (I^{st} -FSAM).* As is well known, the I^{st} -FSAM requires require $O(N_\alpha)$ large-scale forward computations; therefore, the I^{st} -FSAM is advantageous to employ only if, in the problem under consideration, the number N_r of responses of interest exceeds the number of system parameters and/or parameter variations of interest.

In most practical situations, however, the number of model parameters exceeds significantly the number of functional responses of interest, i.e., $N_r \ll N_\alpha$. In such cases, *the (first-order) adjoint sensitivity analysis methodology (I^{st} -ASAM)*, generally developed by Cacuci [1-2], is *the most efficient method for computing exactly the first-order sensitivities, since it requires only $O(N_r)$ large-scale computations*. The I^{st} -ASAM requires that $\delta R(\mathbf{e}^0; \mathbf{h})$ be linear in $\mathbf{h}$, which implies that $R(\mathbf{e})$ must satisfy a weak Lipschitz condition at $\mathbf{e}^0$, and also satisfy the following condition

$$\begin{aligned} R(\mathbf{e}^0 + \varepsilon \mathbf{h}_1 + \varepsilon \mathbf{h}_2) - R(\mathbf{e}^0 + \varepsilon \mathbf{h}_1) - R(\mathbf{e}^0 + \varepsilon \mathbf{h}_2) + R(\mathbf{e}^0) = o(t); \\ \mathbf{h}_1, \mathbf{h}_2 \in \mathcal{H}_u \times \mathcal{H}_\alpha; \quad \varepsilon \in \mathcal{F}. \end{aligned} \quad (9)$$

If $R(\mathbf{e})$ satisfies the two conditions above, then the total response variation $\delta R(\mathbf{e}^0; \mathbf{h})$ is indeed linear in $\mathbf{h}$, and can therefore be denoted as $DR(\mathbf{e}^0; \mathbf{h})$. Consequently, $R(\mathbf{e})$ admits a total G-derivative at $\mathbf{e}^0 = (\mathbf{u}^0, \boldsymbol{\alpha}^0)$, such that the relationship

$$DR(\mathbf{e}^0; \mathbf{h}) = R'_u(\mathbf{e}^0) \mathbf{h}_u + R'_\alpha(\mathbf{e}^0) \mathbf{h}_\alpha \quad (10)$$

holds, where $R'_u(\mathbf{e}^0)$ and $R'_\alpha(\mathbf{e}^0)$ denote the partial G-derivatives at $\mathbf{e}^0$ of $R(\mathbf{e})$ with respect to $\mathbf{u}$ and $\boldsymbol{\alpha}$. It is convenient to refer to the quantities $R'_u(\mathbf{e}^0) \mathbf{h}_u$ and $R'_\alpha(\mathbf{e}^0) \mathbf{h}_\alpha$ appearing in Eq. (10) as the “*indirect effect term*” and the “*direct effect term*,” respectively. The operator $R'_u(\mathbf{e}^0)$ acts linearly on the vector of (arbitrary) variations $\mathbf{h}_u$, from $\mathcal{E}_u$ into $\mathcal{F}$, while the operator $R'_\alpha(\mathbf{e}^0)$ acts linearly on the vector of (arbitrary) variations $\mathbf{h}_\alpha$, from $\mathcal{E}_\alpha$ into $\mathcal{F}$.

To implement the I^{st} -ASAM for computing the first-order G-differential $DR(\mathbf{e}^0; \mathbf{h})$, the spaces $\mathcal{E}_u$, $\mathcal{E}_\alpha$, and $\mathcal{E}_Q$ will henceforth be considered to be Hilbert spaces and denoted as $\mathcal{H}_u(\Omega_x)$ and $\mathcal{H}_Q(\Omega_x)$, respectively. The elements of $\mathcal{H}_u(\Omega_x)$ and $\mathcal{H}_Q(\Omega_x)$ are, as before, vector-valued functions defined on the open set $\Omega_x \subset \mathbb{R}^{J_x}$, with smooth boundary $\partial\Omega_x$. On $\mathcal{H}_u(\Omega_x)$, the inner product of two vectors $\mathbf{u}^{(1)} \in \mathcal{H}_u$ and $\mathbf{u}^{(2)} \in \mathcal{H}_u$ will be

denoted as $\langle \mathbf{u}^{(1)}, \mathbf{u}^{(2)} \rangle_u$; similarly, the inner product on $\mathcal{H}_Q(\Omega_x)$ of two vectors $\mathbf{Q}^{(1)} \in \mathcal{H}_Q$ and $\mathbf{Q}^{(2)} \in \mathcal{H}_Q$ will be denoted as $\langle \mathbf{Q}^{(1)}, \mathbf{Q}^{(2)} \rangle_Q$. The inner product of two vectors $\boldsymbol{\alpha}^{(1)} \in \mathcal{H}_\alpha$ and $\boldsymbol{\alpha}^{(2)} \in \mathcal{H}_\alpha$ on the Hilbert space $\mathcal{H}_\alpha$ will be denoted as $\langle \boldsymbol{\alpha}^{(1)}, \boldsymbol{\alpha}^{(2)} \rangle_\alpha$.

Note that $R'_u(\mathbf{e}^0)\mathbf{h}_u$ is a continuous linear (in $\mathbf{h}_u$) functional on $\mathcal{H}_u(\Omega_x)$, and thus lies in the dual space $\mathcal{H}_u^*(\Omega_x)$. As is well known, real Hilbert spaces are self-dual; for complex Hilbert spaces, it is conventional to identify $\mathcal{H}_u^*(\Omega_x)$ with $\mathcal{H}_u(\Omega_x)$ in the sense of correspondence under the usual isometric anti-isomorphism introduced by complex conjugation. Furthermore, the well known Riesz representation theorem ensures that there exists a unique (column) vector $\mathbf{D}_u R(\mathbf{e}^0) \in \mathcal{H}_u$, defined as $\mathbf{D}_u R(\mathbf{e}) \equiv [\partial R(\mathbf{e})/\partial u_1, \dots, \partial R(\mathbf{e})/\partial u_{K_u}]$, which is customarily called the partial gradient of $R(\mathbf{e})$ with respect to $\mathbf{u}$, evaluated at $\mathbf{e}^0$, such that

$$R'_u(\mathbf{e}^0)\mathbf{h}_u = \langle \mathbf{D}_u R(\mathbf{e}^0), \mathbf{h}_u \rangle_u, \quad \mathbf{h}_u \in \mathcal{H}_u. \quad (11)$$

Similarly, the functional $R'_\alpha(\mathbf{e}^0)\mathbf{h}_\alpha$ is linear in $\mathbf{h}_\alpha$; consequently, the Riesz representation theorem ensures that there exists a unique vector $\mathbf{D}_\alpha R(\mathbf{e}^0) \in \mathcal{H}_\alpha$, where $\mathbf{D}_\alpha R(\mathbf{e}) \equiv [\partial R(\mathbf{e})/\partial \alpha_1, \dots, \partial R(\mathbf{e})/\partial \alpha_{N_\alpha}]$, which is customarily called the partial gradient of $R(\mathbf{e})$ with respect to $\boldsymbol{\alpha}$, evaluated at $\mathbf{e}^0$, such that

$$R'_\alpha(\mathbf{e}^0)\mathbf{h}_\alpha = \langle \mathbf{D}_\alpha R(\mathbf{e}^0), \mathbf{h}_\alpha \rangle_\alpha, \quad \mathbf{h}_\alpha \in \mathcal{H}_\alpha. \quad (12)$$

Furthermore, Eqs. (7) and (8) can also be written, respectively, in the following inner-product forms

$$\mathbf{L}(\boldsymbol{\alpha}^0)\mathbf{h}_u = \langle \mathbf{D}_\alpha \mathbf{Q}(\boldsymbol{\alpha}^0) - \mathbf{D}_\alpha [\mathbf{L}(\boldsymbol{\alpha}^0)\mathbf{u}^0], \mathbf{h}_u \rangle_\alpha, \quad \mathbf{x} \in \Omega_x, \quad (13)$$

and

$$\mathbf{B}(\boldsymbol{\alpha}^0)\mathbf{h}_u = \langle \mathbf{D}_\alpha \mathbf{A}(\boldsymbol{\alpha}^0) - \mathbf{D}_\alpha \mathbf{B}(\boldsymbol{\alpha}^0)\mathbf{u}^0, \mathbf{h}_u \rangle_\alpha, \quad \mathbf{x} \in \partial\Omega_x, \quad (14)$$

respectively, where

$$\mathbf{D}_\alpha \mathbf{Q}(\boldsymbol{\alpha}) \equiv \begin{pmatrix} \frac{\partial Q_1}{\partial \alpha_1} & \dots & \frac{\partial Q_1}{\partial \alpha_{N_\alpha}} \\ \vdots & \ddots & \vdots \\ \frac{\partial Q_{K_u}}{\partial \alpha_1} & \dots & \frac{\partial Q_{K_u}}{\partial \alpha_{N_\alpha}} \end{pmatrix}; \mathbf{D}_\alpha [\mathbf{L}(\boldsymbol{\alpha}) \mathbf{u}] \equiv \begin{pmatrix} \frac{\partial \left[\sum_{j=1}^{K_u} L_{1,j}(\boldsymbol{\alpha}) u_j \right]}{\partial \alpha_1} & \dots & \frac{\partial \left[\sum_{j=1}^{K_u} L_{1,j}(\boldsymbol{\alpha}) u_j \right]}{\partial \alpha_{N_\alpha}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \left[\sum_{j=1}^{K_u} L_{K_u,j}(\boldsymbol{\alpha}) u_j \right]}{\partial \alpha_1} & \dots & \frac{\partial \left[\sum_{j=1}^{K_u} L_{K_u,j}(\boldsymbol{\alpha}) u_j \right]}{\partial \alpha_{N_\alpha}} \end{pmatrix}.$$

The partial-gradient operators $\mathbf{D}_\alpha \mathbf{A}(\boldsymbol{\alpha}^0)$ and $\mathbf{D}_\alpha \mathbf{B}(\boldsymbol{\alpha}^0) \mathbf{u}^0$ are defined analogously to the ones above.

Following [1] for the construction of the I^{st} -ASAM, the formal adjoint $\mathbf{L}^*(\boldsymbol{\alpha}^0)$ of $\mathbf{L}(\boldsymbol{\alpha}^0)$ is now introduced by requiring that the following relationship hold for an arbitrary vector $\boldsymbol{\psi} \in \mathcal{H}_Q$:

$$\langle \boldsymbol{\psi}, \mathbf{L}(\boldsymbol{\alpha}^0) \mathbf{h}_u \rangle_Q = \langle \mathbf{L}^*(\boldsymbol{\alpha}^0) \boldsymbol{\psi}, \mathbf{h}_u \rangle_u + \{P(\mathbf{h}_u; \boldsymbol{\psi})\}_{\partial\Omega_x} \quad (15)$$

In the above equation, the *formal adjoint operator* $\mathbf{L}^*(\boldsymbol{\alpha}^0)$ is the $K_u \times K_u$ matrix $\mathbf{L}^*(\boldsymbol{\alpha}^0) \equiv [L_{ji}^*(\boldsymbol{\alpha}^0)]$, $(i, j = 1, \dots, K_u)$, comprising the elements $L_{ji}^*(\boldsymbol{\alpha}^0)$ obtained by transposing the formal adjoints of the operators $L_{ij}(\boldsymbol{\alpha}^0)$, while $\{P(\mathbf{h}_u; \boldsymbol{\psi})\}_{\partial\Omega_x}$ is the *associated bilinear form* evaluated on $\partial\Omega_x$. The domain of $\mathbf{L}^*(\boldsymbol{\alpha}^0)$ is determined by selecting appropriate *adjoint boundary and/or initial conditions*, represented here in operator form as

$$\mathbf{B}^*(\boldsymbol{\alpha}^0) \boldsymbol{\psi}^0 - \mathbf{A}^*(\boldsymbol{\alpha}^0) = \mathbf{0}, \quad \mathbf{x} \in \partial\Omega_x. \quad (16)$$

The above boundary conditions for $\mathbf{L}^*(\boldsymbol{\alpha}^0)$ are obtained by requiring that:

- (a) They must be independent of *unknown* values of $\mathbf{h}_u$ and $\mathbf{h}_\alpha$;
- (b) The substitution of the boundary and/or initial conditions represented by Eqs. (14) and (16) into the expression of $\{P(\mathbf{h}_u; \boldsymbol{\psi})\}_{\partial\Omega_x}$ must cause all terms containing unknown values of $\mathbf{h}_u$ to vanish.

This selection of the boundary conditions for $\mathbf{L}^*(\mathbf{e}^0)$ reduces the bilinear concomitant $\{P(\mathbf{h}_u; \boldsymbol{\psi})\}_{\partial\Omega_x}$ to a quantity that contains boundary terms involving only known values of $\mathbf{h}_\alpha$, $\boldsymbol{\psi}$, and, possibly, $\boldsymbol{\alpha}^0$; this quantity will be denoted by $\hat{P}(\mathbf{h}_\alpha, \boldsymbol{\psi}; \boldsymbol{\alpha}^0)$. In general, $\hat{P}$ does not automatically vanish as a result of these manipulations, although it may do so in particular instances. In principle, $\hat{P}$ could be forced to vanish by considering extensions of $\mathbf{L}(\boldsymbol{\alpha}^0)$, in the operator sense, but this is seldom needed in practice.

Introducing now Eqs. (13), (14), and (16) into Eq. (15) and re-arranging the resulting equation yields

$$\langle \mathbf{L}^*(\boldsymbol{\alpha}^0)\boldsymbol{\psi}, \mathbf{h}_u \rangle_u = \left\langle \boldsymbol{\psi}, \left\langle \mathbf{D}_\alpha \mathbf{Q}(\boldsymbol{\alpha}^0) - \mathbf{D}_\alpha \mathbf{L}(\boldsymbol{\alpha}^0) \mathbf{u}^0, \mathbf{h}_\alpha \right\rangle_\alpha \right\rangle_Q - \hat{P}(\mathbf{h}_\alpha, \boldsymbol{\psi}; \boldsymbol{\alpha}^0). \quad (17)$$

Since $\boldsymbol{\psi}$ is not completely defined yet, we now complete its definition by requiring that the left-side of Eq. (17) and the right-side of Eq. (11) represent the same functional, i.e.,

$$\langle \mathbf{L}^*(\boldsymbol{\alpha}^0)\boldsymbol{\psi}, \mathbf{h}_u \rangle_u = \langle \mathbf{D}_u R(\mathbf{e}^0), \mathbf{h}_u \rangle_u, \quad \mathbf{h}_u \in \mathcal{H}_u, \quad (18.a)$$

which implies that the adjoint function $\boldsymbol{\psi}$ is the *weak solution* of

$$\mathbf{L}^*(\boldsymbol{\alpha}^0)\boldsymbol{\psi}^0 \stackrel{w}{=} \mathbf{D}_u R(\mathbf{e}^0), \quad (18.b)$$

The superscript “zero” in Eqs. (18) emphasizes the fact that the function $\boldsymbol{\psi}^0$ satisfies Eqs. (18) and (16) for the nominal parameter values $\boldsymbol{\alpha}^0$. Note that the well-known Riesz representation theorem ensures that the above relationship, where $\boldsymbol{\psi}^0$ satisfies the adjoint boundary conditions given in Eq. (16), holds uniquely. The construction of the *first-level adjoint sensitivity system* (I^{st} -LASS), consisting of Eqs. (18) and (16), has thus been accomplished. Furthermore, Eqs. (10) through (18) can now be used to obtain the following expression for the total sensitivity $DR(\mathbf{e}^0; \mathbf{h})$ of $R(\mathbf{e})$ at:

$$\begin{aligned}
DR(\mathbf{e}^0; \mathbf{h}) &= \left\langle \mathbf{D}_\alpha R(\mathbf{e}^0), \mathbf{h}_\alpha \right\rangle_\alpha + \left\langle \boldsymbol{\psi}^0, \left\langle \mathbf{D}_\alpha \mathbf{Q}(\boldsymbol{\alpha}^0) - \mathbf{D}_\alpha \mathbf{L}(\boldsymbol{\alpha}^0) \mathbf{u}^0, \mathbf{h}_\alpha \right\rangle_\alpha \right\rangle_\psi - \hat{P}(\mathbf{h}_\alpha, \boldsymbol{\psi}^0; \boldsymbol{\alpha}^0) \\
&\equiv \mathbf{S}(\mathbf{u}^0, \boldsymbol{\alpha}^0, \boldsymbol{\psi}^0) \mathbf{h}_\alpha = \sum_{i=1}^{N_\alpha} S_i(\mathbf{u}^0, \boldsymbol{\alpha}^0, \boldsymbol{\psi}^0) \delta \alpha_i,
\end{aligned} \tag{19}$$

where $\mathbf{S}(\mathbf{u}^0, \boldsymbol{\alpha}^0, \boldsymbol{\psi}^0) \equiv (S_1, \dots, S_{N_\alpha})^\dagger$, and where the i^{th} -partial first-order sensitivity (G-derivative), $S_i(\mathbf{u}^0, \boldsymbol{\alpha}^0, \boldsymbol{\psi}^0)$, of $R(\mathbf{e})$ with respect to the i^{th} -model parameter α_i , $i = 1, \dots, N_\alpha$, is given by the expression

$$\begin{aligned}
S_i(\mathbf{u}^0, \boldsymbol{\alpha}^0, \boldsymbol{\psi}^0) &\equiv (\mathbf{D}_\alpha R)_i + \left\langle \boldsymbol{\psi}^0, \left[\mathbf{D}_\alpha \mathbf{Q}(\boldsymbol{\alpha}^0) - \mathbf{D}_\alpha \mathbf{L}(\boldsymbol{\alpha}^0) \mathbf{u}^0 \right]_i \right\rangle_\psi - \left[\mathbf{D}_\alpha \hat{P}(\mathbf{u}^0, \boldsymbol{\psi}^0, \boldsymbol{\alpha}^0) \right]_i \\
&= \frac{\partial R(\mathbf{u}^0, \boldsymbol{\alpha}^0)}{\partial \alpha_i} + \left\langle \boldsymbol{\psi}^0, \frac{\partial \mathbf{Q}(\mathbf{u}^0, \boldsymbol{\alpha}^0)}{\partial \alpha_i} - \frac{\partial [\mathbf{L}(\boldsymbol{\alpha}^0) \mathbf{u}^0]}{\partial \alpha_i} \right\rangle_\psi - \frac{\partial \hat{P}(\mathbf{u}^0, \boldsymbol{\psi}^0, \boldsymbol{\alpha}^0)}{\partial \alpha_i}, \quad i = 1, \dots, N_\alpha.
\end{aligned} \tag{20}$$

All partial derivatives in the above expressions are to be understood as partial G-derivatives, of course. As Eq. (20) indicates, the desired elimination of all unknown values of $\mathbf{h}_u$ from the expressions of the sensitivities $S_i(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\psi})$, $i = 1, \dots, N_\alpha$, of $R(\mathbf{e})$ at $\mathbf{e}^0$ has been accomplished. Note that we have designated the space $\mathcal{D}_{\mathcal{Q}}(\Omega_x)$ as $\mathcal{D}_\psi(\Omega_x)$, in order to emphasize that we are dealing with the Hilbert space on which the adjoint function $\boldsymbol{\psi}$ is defined. *The sensitivities $S_i(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\psi})$ can therefore be computed by means of Eq. (20), after solving **only once** the adjoint sensitivity system, consisting of Eqs. (16) and (18), to obtain the adjoint function $\boldsymbol{\psi}$.* It is very important to note that this adjoint system is independent not only of the functions $\mathbf{h}_u$ but also of the nominal values $\mathbf{u}^0$ of $\mathbf{u}$. This means that the adjoint system, namely Eqs. (16) and (18), can be solved independently of the solution $\mathbf{u}^0$ of the original equations. In turn, this fact simplifies considerably the choice of numerical methods for solving the adjoint system. *It is also important to note that this advantageous situation arises if and only if the original equations that model the physical problem, i.e., Eqs. (1) and (2), are linear in the state-variable $\mathbf{u}$.*

3. THE SECOND-ORDER FORWARD AND ADJOINT SENSITIVITY ANALYSIS METHODOLOGIES FOR LINEAR SYSTEMS

We now note the very important fact that, since Eq. (20) holds for any nominal parameter values, α^0 , it follows that the first-order sensitivities $S_i(\mathbf{u}, \alpha, \psi)$, $i = 1, \dots, N_\alpha$, can be generally considered as functionals of the original state-function $\mathbf{u}$, the parameters α , and the adjoint function ψ , namely

$$S_i(\mathbf{u}, \alpha, \psi) = \frac{\partial R(\mathbf{u}, \alpha)}{\partial \alpha_i} + \left\langle \psi, \frac{\partial \mathbf{Q}(\mathbf{u}, \alpha)}{\partial \alpha_i} - \frac{\partial [\mathbf{L}(\alpha) \mathbf{u}]}{\partial \alpha_i} \right\rangle_\psi - \frac{\partial \hat{P}(\mathbf{u}, \psi, \alpha)}{\partial \alpha_i}, \quad i = 1, \dots, N_\alpha, \quad (21)$$

where the original state-function $\mathbf{u}$ satisfies Eqs. (1) and (2), while the adjoint function ψ satisfies Eqs. (16) and (18) for any physically-defined values of the parameters α , i.e.,

$$\mathbf{L}^*(\alpha) \psi = \mathbf{D}_u^w R(\mathbf{u}, \alpha), \quad \mathbf{x} \in \Omega_x, \quad (22)$$

together with the corresponding adjoint boundary and/or initial conditions

$$\mathbf{B}^*(\alpha) \psi - \mathbf{A}^*(\alpha) = \mathbf{0}, \quad \mathbf{x} \in \partial\Omega_x. \quad (23)$$

3.1. The Second-Order Forward Sensitivity Analysis Methodology (2nd-FSAM) for Linear Systems

As Eq. (21) indicates, the first-order sensitivities are functionals of the form $S_i(\mathbf{u}, \alpha, \psi) : \mathcal{D}_{R_i} \subset \mathcal{H}_u(\Omega_x) \times \mathcal{H}_\alpha \times \mathcal{H}_\psi(\Omega_x) \rightarrow \mathcal{F}$. Hence, it is possible to define the first-order G-differential, $\delta S_i(\mathbf{e}^0, \psi^0, \mathbf{g})$, of any of the functionals $S_i(\mathbf{u}, \alpha, \psi)$, at the point $(\mathbf{e}^0, \psi^0)$ in the usual manner, namely

$$\delta S_i(\mathbf{e}^0, \boldsymbol{\psi}^0; \mathbf{g}) \equiv \left\{ \frac{d}{d\boldsymbol{\varepsilon}} \left[S_i(\mathbf{u}^0 + \boldsymbol{\varepsilon} \mathbf{h}_u, \boldsymbol{\psi}^0 + \boldsymbol{\varepsilon} \mathbf{h}_\psi, \boldsymbol{\alpha}^0 + \boldsymbol{\varepsilon} \mathbf{h}_\alpha) \right] \right\}_{\boldsymbol{\varepsilon}=0}, \quad \boldsymbol{\varepsilon} \in \mathcal{T}, \quad (24)$$

for all (i.e., arbitrary) vectors $\mathbf{g} \equiv (\mathbf{h}_u, \mathbf{h}_\alpha, \mathbf{h}_\psi) \in \mathcal{H}_u(\Omega_x) \times \mathcal{H}_\alpha \times \mathcal{H}_\psi(\Omega_x)$. Applying the above definition to the expression of $S_i(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\psi})$ given by Eq. (21) yields

$$\delta S_i(\mathbf{e}^0, \boldsymbol{\psi}^0; \mathbf{g}) = (\delta S_i)_{\text{direct}} + (\delta S_i)_{\text{indirect}}, \quad (25)$$

where $(\delta S_i)_{\text{direct}}$ denotes the “*direct-effect term*” and $(\delta S_i)_{\text{indirect}}$ denotes the “*indirect-effect term*.” The “*direct-effect term*,” $(\delta S_i)_{\text{direct}}$, is defined as

$$\begin{aligned} (\delta S_i)_{\text{direct}} &\equiv \langle \mathbf{D}_\alpha S_i, \mathbf{h}_\alpha \rangle_\alpha \\ &= \frac{\partial}{\partial \alpha_i} \left\{ \left\langle \mathbf{D}_\alpha R(\mathbf{u}, \boldsymbol{\alpha}) - \mathbf{D}_\alpha \hat{P}(\mathbf{u}, \boldsymbol{\psi}, \boldsymbol{\alpha}) + \langle \boldsymbol{\psi}, \mathbf{D}_\alpha \mathbf{Q}(\boldsymbol{\alpha}) - \mathbf{D}_\alpha [\mathbf{L}(\boldsymbol{\alpha}) \mathbf{u}] \rangle_\psi, \mathbf{h}_\alpha \right\rangle_\alpha \right\}_{(\mathbf{u}^0, \boldsymbol{\psi}^0; \boldsymbol{\alpha}^0)}, \end{aligned} \quad (26)$$

and can be computed immediately at this stage, without needing any additional large-scale computations. The “*indirect-effect term*,” $(\delta S_i)_{\text{indirect}}$, is defined as

$$(\delta S_i)_{\text{indirect}} \equiv \left\langle \left(\mathbf{h}_u, \mathbf{h}_\psi \right)^\dagger \begin{pmatrix} \mathbf{D}_u S_i(\mathbf{u}^0, \boldsymbol{\psi}^0, \boldsymbol{\alpha}^0) \\ \mathbf{D}_\psi S_i(\mathbf{u}^0, \boldsymbol{\psi}^0, \boldsymbol{\alpha}^0) \end{pmatrix} \right\rangle_{u\psi}, \quad (27)$$

where we used the notation $(\mathbf{h}_u, \mathbf{h}_\psi)^\dagger \equiv (\mathbf{h}_u^\dagger, \mathbf{h}_\psi^\dagger)$ and where

$$\mathbf{D}_\psi S_i(\mathbf{u}, \boldsymbol{\psi}, \boldsymbol{\alpha}) \equiv \frac{\partial}{\partial \alpha_i} \left[\mathbf{Q}(\boldsymbol{\alpha}) - \mathbf{L}(\boldsymbol{\alpha}) \mathbf{u} - \mathbf{D}_\psi \hat{P}(\mathbf{u}, \boldsymbol{\psi}, \boldsymbol{\alpha}) \right], \quad (28)$$

$$\mathbf{D}_u S_i(\mathbf{u}, \boldsymbol{\psi}, \boldsymbol{\alpha}) \equiv \frac{\partial}{\partial \alpha_i} \left[\mathbf{D}_u R(\mathbf{u}, \boldsymbol{\alpha}) - \mathbf{D}_u \hat{P}(\mathbf{u}, \boldsymbol{\psi}, \boldsymbol{\alpha}) - \boldsymbol{\Lambda}(\boldsymbol{\psi}, \boldsymbol{\alpha}) \right], \quad (29)$$

with

$$\boldsymbol{\Lambda}(\boldsymbol{\psi}, \boldsymbol{\alpha}) \equiv (\Lambda_1, \dots, \Lambda_{K_u}); \quad \Lambda_k(\boldsymbol{\psi}, \boldsymbol{\alpha}) \equiv \sum_{j=1}^{K_u} \psi_j(\mathbf{x}) \frac{\partial L_{jk}(\boldsymbol{\alpha})}{\partial \alpha_k}, \quad k = 1, \dots, K_u. \quad (30)$$

Note that the “indirect-effect term” *cannot be computed at this stage*, since the vectors of variations $\mathbf{h}_u$ and $\mathbf{h}_\psi$ are unknown. Recall that the vector $\mathbf{h}_u$ is the $\mathbf{h}_\alpha$ -dependent solution of the I^{st} -LFSS, comprising Eqs. (7) and (8), which are computationally impractical to solve for large-scale systems with many parameters. Furthermore, the vector of variations $\mathbf{h}_\psi$ (around the nominal value $\boldsymbol{\psi}^0$) is the solution of the system of operator equations that results from applying the definition of the G-differential to the I^{st} -LASS, i.e., to Eqs. (22) and (23), to obtain

$$\left[\frac{d}{d\varepsilon} \left\{ \mathbf{L}^* \left[(\boldsymbol{\alpha}^0 + \varepsilon \mathbf{h}_\alpha) \right] (\boldsymbol{\psi}^0 + \varepsilon \mathbf{h}_\psi) - \mathbf{D}_u R(\mathbf{u}^0 + \varepsilon \mathbf{h}_u, \boldsymbol{\alpha}^0 + \varepsilon \mathbf{h}_\alpha) \right\} \right]_{\varepsilon=0} = \mathbf{0}, \quad \mathbf{x} \in \Omega_x, \quad (31)$$

$$\left[\frac{d}{d\varepsilon} \left\{ \mathbf{B}^* \left[(\boldsymbol{\alpha}^0 + \varepsilon \mathbf{h}_\alpha) \right] (\boldsymbol{\psi}^0 + \varepsilon \mathbf{h}_\psi) - \mathbf{A}^* (\boldsymbol{\alpha}^0 + \varepsilon \mathbf{h}_\alpha) \right\} \right]_{\varepsilon=0} = \mathbf{0}, \quad \mathbf{x} \in \partial\Omega_x. \quad (32)$$

As the above equations clearly show, the vector of variations $\mathbf{h}_\psi$ (around the nominal value $\boldsymbol{\psi}^0$) is related to the parameters variations $\mathbf{h}_\alpha$. Together, the I^{st} -LFSS, namely Eqs. (7) and (8), and the G-differentiated I^{st} -LASS, namely Eqs. (31) and (32), can be written in the following block-matrix-operator form:

$$\begin{pmatrix} \mathbf{L}(\boldsymbol{\alpha}^0) & \mathbf{0} \\ -\mathbf{D}_{uu}^2 R(\mathbf{e}^0) & \mathbf{L}^*(\boldsymbol{\alpha}^0) \end{pmatrix} \begin{pmatrix} \mathbf{h}_u \\ \mathbf{h}_\psi \end{pmatrix} = \begin{pmatrix} \left\langle \mathbf{D}_\alpha \mathbf{Q}(\boldsymbol{\alpha}^0) - \mathbf{D}_\alpha [\mathbf{L}(\boldsymbol{\alpha}^0) \mathbf{u}^0], \mathbf{h}_\alpha \right\rangle_\alpha \\ \left\langle \mathbf{D}_{\alpha u}^2 R(\mathbf{e}^0) - \mathbf{L}^*[(\boldsymbol{\alpha}^0) \boldsymbol{\psi}^0], \mathbf{h}_\alpha \right\rangle_\alpha \end{pmatrix}, \quad \mathbf{x} \in \Omega_x, \quad (33)$$

together with the corresponding G-differentiated boundary and/or initial conditions

$$\begin{pmatrix} \mathbf{B}(\boldsymbol{\alpha}^0) & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^*(\boldsymbol{\alpha}^0) \end{pmatrix} \begin{pmatrix} \mathbf{h}_u \\ \mathbf{h}_\psi \end{pmatrix} = \begin{pmatrix} \left\langle \mathbf{D}_\alpha \mathbf{A}(\boldsymbol{\alpha}^0) - \mathbf{D}_\alpha [\mathbf{B}(\boldsymbol{\alpha}^0) \mathbf{u}^0], \mathbf{h}_\alpha \right\rangle_\alpha \\ \left\langle \mathbf{D}_\alpha \mathbf{A}^*(\boldsymbol{\alpha}^0) - \mathbf{D}_\alpha [\mathbf{B}^*(\boldsymbol{\alpha}^0) \boldsymbol{\psi}^0], \mathbf{h}_\alpha \right\rangle_\alpha \end{pmatrix}, \quad \mathbf{x} \in \partial\Omega_x. \quad (34)$$

In Eq. (33), the operators $\mathbf{D}_{uu}^2 R(\mathbf{e}^0)$ and, respectively, $\mathbf{D}_{\alpha u}^2 R(\mathbf{e}^0)$ are matrices of partial G-derivatives, of the form

$$\mathbf{D}_{\alpha u}^2 R(\mathbf{e}) \equiv \begin{pmatrix} \frac{\partial^2 R(\mathbf{e})}{\partial u_1 \partial \alpha_1} & \cdots & \frac{\partial^2 R(\mathbf{e})}{\partial u_1 \partial \alpha_{N_\alpha}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 R(\mathbf{e})}{\partial u_{K_u} \partial \alpha_1} & \cdots & \frac{\partial^2 R(\mathbf{e})}{\partial u_{K_u} \partial \alpha_{N_\alpha}} \end{pmatrix}; \quad \mathbf{D}_{uu}^2 R(\mathbf{e}) \equiv \begin{pmatrix} \frac{\partial^2 R(\mathbf{e})}{\partial u_1^2} & \cdots & \frac{\partial^2 R(\mathbf{e})}{\partial u_1 \partial u_{K_u}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 R(\mathbf{e})}{\partial u_{K_u} \partial u_1} & \cdots & \frac{\partial^2 R(\mathbf{e})}{\partial u_{K_u}^2} \end{pmatrix}. \quad (35)$$

The block matrix Eq. (33) together with the boundary and/or initial conditions represented by Eq. (34) constitute the *second-level forward sensitivity system* (2^{nd} -LFSS). These equations can be solved to obtain the vectors $\mathbf{h}_u$ and $\mathbf{h}_\psi$, which can, in turn, be used in Eq. (27) to compute the indirect-effect term, $(\delta S_i)_{indirect}$. Using Eq. (25), this indirect effect term would then be added together with the already computed direct-effect term, $(\delta S_i)_{direct}$, to obtain the vector, $S_i(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\psi})$, of partial mixed second-order sensitivities of the form $\{\partial^2 R(\mathbf{u}, \boldsymbol{\alpha}) / \partial \alpha_j \partial \alpha_i\}_{(\mathbf{u}^0, \boldsymbol{\alpha}^0)}$, $j = 1, \dots, N_\alpha$. Thus, Eqs. (33) and (34) would be solved to obtain $S_1(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\psi})$, for $i = 1$; the process of solving these operator equations would continue for $S_2(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\psi})$, for $i = 2$, and would end by computing $S_{N_\alpha}(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\psi})$, for $i = N_\alpha$.

Clearly, the computational process just described would yield the complete set of all second-order responses sensitivities, $\{\partial^2 R(\mathbf{u}, \boldsymbol{\alpha}) / \partial \alpha_j \partial \alpha_i\}_{(\mathbf{u}^0, \boldsymbol{\alpha}^0)}$, to all of the system parameters, parameters, in $O(N_\alpha^2)$ large-scale forward computations. Of course, such a computational burden would be impractical for large-scale systems.

3.2. The Second-Order Adjoint Sensitivity Analysis Methodology (2^{nd} -ASAM) for Linear Systems

Just as the 1^{st} -ASAM aimed at computing efficiently (using adjoint functions) the first-order indirect-effect term $R'_u(\mathbf{e}^0)\mathbf{h}_u$, the *second-order adjoint sensitivity analysis methodology* (2^{nd} -ASAM) to be presented in this Section will aim at the efficient computation of the second-order indirect-effect term $(\delta S_i)_{indirect}$, by using the 2^{nd} -level adjoint sensitivity system (2^{nd} -

LASS) that will be developed in this Section. To construct the 2^{nd} -*LASS*, consider two vectors $\mathbf{v}^{(1)} \equiv (\mathbf{u}^{(1)}, \boldsymbol{\psi}^{(1)}) \in \mathcal{D}\ell_{u\psi}$ and $\mathbf{v}^{(2)} \equiv (\mathbf{u}^{(2)}, \boldsymbol{\psi}^{(2)}) \in \mathcal{D}\ell_{u\psi}$ in the Hilbert space $\mathcal{D}\ell_{u\psi}(\Omega_x) \equiv \mathcal{D}\ell_u(\Omega_x) \times \mathcal{D}\ell_\psi(\Omega_x)$, endowed with the inner product $\langle \mathbf{v}^{(1)}, \mathbf{v}^{(2)} \rangle_{u\psi}$ defined as

$$\langle \mathbf{v}^{(1)}, \mathbf{v}^{(2)} \rangle_{u\psi} \equiv \int_{\Omega_x} (\mathbf{u}^{(1)} \cdot \mathbf{u}^{(2)} + \boldsymbol{\psi}^{(1)} \cdot \boldsymbol{\psi}^{(2)}) d\mathbf{x}. \quad (36)$$

Using the above definition, we form the inner product of Eq. (33) with a yet undefined vector $(\boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)})^\dagger$ to obtain the sequence of equalities shown below:

$$\begin{aligned} & \left\langle (\boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)})^\dagger \begin{pmatrix} \mathbf{L}(\boldsymbol{\alpha}^0) & \mathbf{0} \\ -\mathbf{D}_{uu}^2 R(\mathbf{e}^0) & \mathbf{L}^*(\boldsymbol{\alpha}^0) \end{pmatrix} \begin{pmatrix} \mathbf{h}_u \\ \mathbf{h}_\psi \end{pmatrix} \right\rangle_{u\psi} \\ &= \left\langle (\boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)})^\dagger \begin{pmatrix} \langle \mathbf{D}_\alpha \mathbf{Q}(\boldsymbol{\alpha}^0) - \mathbf{D}_\alpha [\mathbf{L}(\boldsymbol{\alpha}^0) \mathbf{u}^0], \mathbf{h}_\alpha \rangle_\alpha \\ \langle \mathbf{D}_{uu}^2 R(\mathbf{e}^0) - \mathbf{L}^*[(\boldsymbol{\alpha}^0) \boldsymbol{\psi}^0], \mathbf{h}_\alpha \rangle_\alpha \end{pmatrix} \right\rangle_{u\psi} \\ &= \left\langle (\mathbf{h}_u, \mathbf{h}_\psi)^\dagger \begin{pmatrix} \mathbf{L}^*(\boldsymbol{\alpha}^0) & -[\mathbf{D}_{uu}^2 R(\mathbf{e}^0)]^* \\ \mathbf{0} & \mathbf{L}(\boldsymbol{\alpha}^0) \end{pmatrix} \begin{pmatrix} \boldsymbol{\psi}^{(1)} \\ \boldsymbol{\psi}^{(2)} \end{pmatrix} \right\rangle_{u\psi} + \left\{ P_2(\mathbf{u}^0, \boldsymbol{\alpha}^0; \mathbf{h}_u, \mathbf{h}_\psi, \mathbf{h}_\alpha; \boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)}) \right\}_{\partial\Omega_x}, \quad \mathbf{x} \in \Omega_x, \end{aligned} \quad (37)$$

where $[\mathbf{D}_{uu}^2 R(\mathbf{e}^0)]^*$ denotes the formal adjoint of $\mathbf{D}_{uu}^2 R(\mathbf{e}^0)$ and $\left\{ P_2(\mathbf{u}^0, \boldsymbol{\alpha}^0; \mathbf{h}_u, \mathbf{h}_\psi, \mathbf{h}_\alpha; \boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)}) \right\}_{\partial\Omega_x}$ denotes the corresponding bilinear concomitant on $\mathbf{x} \in \partial\Omega_x$. Following the same principles as in Section 2, we require the first term on the right-side of the last equality in Eq. (37) to represent the same functional as the right-side of Eq. (27), i.e.,

$$\left\langle (\mathbf{h}_u, \mathbf{h}_\psi)^\dagger \begin{pmatrix} \mathbf{L}^*(\boldsymbol{\alpha}^0) & -[\mathbf{D}_{uu}^2 R(\mathbf{e}^0)]^* \\ \mathbf{0} & \mathbf{L}(\boldsymbol{\alpha}^0) \end{pmatrix} \begin{pmatrix} \boldsymbol{\psi}^{(1)} \\ \boldsymbol{\psi}^{(2)} \end{pmatrix} \right\rangle_{u\psi} = \left\langle (\mathbf{h}_u, \mathbf{h}_\psi)^\dagger \begin{pmatrix} \mathbf{D}_u S_i(\mathbf{u}^0, \boldsymbol{\psi}^0, \boldsymbol{\alpha}^0) \\ \mathbf{D}_\psi S_i(\mathbf{u}^0, \boldsymbol{\psi}^0, \boldsymbol{\alpha}^0) \end{pmatrix} \right\rangle_{u\psi} \quad (38.a)$$

which indicates that the adjoint functions $\boldsymbol{\psi}^{(1)}$ and $\boldsymbol{\psi}^{(2)}$ are *weak solutions* of the system below:

$$\begin{aligned}\mathbf{L}(\boldsymbol{\alpha}^0)\boldsymbol{\psi}^{(2)} &= \mathbf{D}_{\psi} S_i(\mathbf{u}^0, \boldsymbol{\psi}^0, \boldsymbol{\alpha}^0), \\ \mathbf{L}^*(\boldsymbol{\alpha}^0)\boldsymbol{\psi}^{(1)} &= \mathbf{D}_u S_i(\mathbf{u}^0, \boldsymbol{\psi}^0, \boldsymbol{\alpha}^0) + \left[\mathbf{D}_{uu}^2 R(\mathbf{e}^0) \right]^* \boldsymbol{\psi}^{(2)}.\end{aligned}\tag{38.b}$$

The definition of the functions $\boldsymbol{\psi}^{(1)}$ and $\boldsymbol{\psi}^{(2)}$ can now be completed by requiring them to satisfy adjoint boundary conditions denoted as

$$\begin{aligned}\mathbf{B}_1^*(\mathbf{u}^0, \boldsymbol{\psi}^0; \boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)}, \boldsymbol{\alpha}^0) &= \mathbf{0}, \quad \mathbf{x} \in \partial\Omega_x, \\ \mathbf{B}_2^*(\mathbf{u}^0, \boldsymbol{\psi}^0; \boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)}, \boldsymbol{\alpha}^0) &= \mathbf{0}, \quad \mathbf{x} \in \partial\Omega_x.\end{aligned}\tag{39}$$

The above adjoint boundary and/or initial are determined following the same principles as described in Section 2, namely that:

- (a) They must be independent of *unknown* values of $\mathbf{h}_u, \mathbf{h}_{\psi}$, and $\mathbf{h}_{\alpha}$;
- (b) The substitution of the boundary and/or initial conditions represented by Eqs. (34) and (39) into the expression of $\left\{ P_2(\mathbf{u}^0, \boldsymbol{\alpha}^0; \mathbf{h}_u, \mathbf{h}_{\psi}, \mathbf{h}_{\alpha}; \boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)}) \right\}_{\partial\Omega_x}$ must cause all terms containing unknown values of $\mathbf{h}_u$ and $\mathbf{h}_{\psi}$ to vanish, which will reduce the bilinear concomitant to an expression that will contain only known values of $\mathbf{h}_{\alpha}$; we denote this expression as $\hat{P}_2(\mathbf{u}^0, \boldsymbol{\alpha}^0; \boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)}; \mathbf{h}_{\alpha})$.

Once the adjoint functions $\boldsymbol{\psi}^{(1)}$ and $\boldsymbol{\psi}^{(2)}$ have been determined by solving Eqs. (38) and (39), they can be used in conjunction with Eq. (37) to represent the “*indirect-effect term*”, $(\delta S_i)_{indirect}$, defined in Eq. (27) in the form

$$(\delta S_i)_{indirect} = \left\langle (\boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)}) \left(\left\langle \mathbf{D}_{\alpha} \mathbf{Q}(\boldsymbol{\alpha}^0) - \mathbf{D}_{\alpha} [\mathbf{L}(\boldsymbol{\alpha}^0) \mathbf{u}^0], \mathbf{h}_{\alpha} \right\rangle_{\alpha} \right) \right\rangle_{u\psi} - \hat{P}_2(\mathbf{u}^0, \boldsymbol{\alpha}^0; \boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)}; \mathbf{h}_{\alpha}).\tag{40}$$

In terms of the adjoint functions $\boldsymbol{\psi}^{(1)}$ and $\boldsymbol{\psi}^{(2)}$, the complete expression of the second-order mixed sensitivities $\delta S_i = (\delta S_i)_{\text{direct}} + (\delta S_i)_{\text{indirect}}$ is obtained by adding the above expression to the previously computed “direct effect term” from Eq. (26), to obtain

$$\begin{aligned} \delta S_i(\mathbf{u}^0, \boldsymbol{\alpha}^0, \boldsymbol{\psi}^0, \boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)}; \mathbf{h}_\alpha) &= (\delta S_i)_{\text{direct}} + (\delta S_i)_{\text{indirect}} \\ &= \frac{\partial}{\partial \alpha_i} \left\{ \left\langle \mathbf{D}_\alpha R(\mathbf{u}, \boldsymbol{\alpha}) - \mathbf{D}_\alpha \hat{P}(\mathbf{u}, \boldsymbol{\psi}, \boldsymbol{\alpha}) + \left\langle \boldsymbol{\psi}, \mathbf{D}_\alpha \mathbf{Q}(\boldsymbol{\alpha}) - \mathbf{D}_\alpha [\mathbf{L}(\boldsymbol{\alpha}) \mathbf{u}] \right\rangle_\psi, \mathbf{h}_\alpha \right\rangle_\alpha \right\}_{(\mathbf{u}^0, \boldsymbol{\psi}^0; \boldsymbol{\alpha}^0)} \\ &\quad + \left\langle \left(\boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)} \right) \left(\left\langle \mathbf{D}_\alpha \mathbf{Q}(\boldsymbol{\alpha}^0) - \mathbf{D}_\alpha [\mathbf{L}(\boldsymbol{\alpha}^0) \mathbf{u}^0], \mathbf{h}_\alpha \right\rangle_\alpha \right) \right\rangle_{u\psi} - \hat{P}_2(\mathbf{u}^0, \boldsymbol{\alpha}^0; \boldsymbol{\psi}^{(1)}, \boldsymbol{\psi}^{(2)}; \mathbf{h}_\alpha). \end{aligned} \quad (41)$$

It is convenient to call Eqs. (38) and (39) the 2^{nd} -level adjoint sensitivity system (2^{nd} -LASS) for the second-level adjoint functions $\boldsymbol{\psi}^{(1)}$ and $\boldsymbol{\psi}^{(2)}$, which are used for computing the i^{th} row/column $\left\{ \partial^2 R(\mathbf{u}, \boldsymbol{\alpha}) / \partial \alpha_j \partial \alpha_i \right\}_{(\mathbf{u}^0, \boldsymbol{\alpha}^0)}$, $j = 1, \dots, N_\alpha$, of the matrix of second-order response sensitivities. For each index $i = 1, \dots, N_\alpha$, there corresponds a distinct right-hand side (source) for the corresponding 2^{nd} -LASS; strictly speaking, therefore, the second-level adjoint functions $\boldsymbol{\psi}^{(1)}$ and $\boldsymbol{\psi}^{(2)}$ should also be indexed for $i = 1, \dots, N_\alpha$, but such an additional notational complexity is not warranted for the scope of this work.

Note that the equations constituting the 2^{nd} -LASS are independent of parameter variations $\mathbf{h}_\alpha$. Therefore, these equations need to be solved only once to compute the adjoint functions $\boldsymbol{\psi}^{(1)}$ and $\boldsymbol{\psi}^{(2)}$. Thus, the exact computation of all second-order sensitivities, $\left\{ \partial^2 R(\mathbf{u}, \boldsymbol{\alpha}) / \partial \alpha_j \partial \alpha_i \right\}_{(\mathbf{u}^0, \boldsymbol{\alpha}^0)}$, for $i, j = 1, \dots, N_\alpha$, using the 2^{nd} -ASAM requires at most N_α large-scale computations using the 2^{nd} -LASS, rather than $O(N_\alpha^2)$ large-scale computations as would be required by forward methods. It is also important to note that the construction and solution of the 2^{nd} -LASS requires very little effort beyond that already invested in solving the original forward Eq.(1) for the state variable $\mathbf{u}(\mathbf{x})$ and the 1^{st} -LASS, cf. Eq. (22) and (23), for the first-level adjoint function $\boldsymbol{\psi}$. This is because, as Eqs. (38.b) indicate, the equation for determining the second-level adjoint function $\boldsymbol{\psi}^{(2)}$ is the same as that for determining $\mathbf{u}(\mathbf{x})$, except for a different source term, while the equations to be solved for determining the

second-level adjoint function $\psi^{(1)}$ are the same as those for determining the first-level adjoint function ψ , save again for a different source term.

4. COMPARISON OF 2nd-ASAM WITH PREVIOUS METHODS FOR COMPUTING SECOND-ORDER RESPONSE SENSITIVITIES (DERIVATIVES) WITH RESPECT TO MODEL PARAMETERS

Following a brief review of the earliest attempts at computing second-order response sensitivities to model (“input”) parameters, this Section will present a comparative discussion of the basic properties of the currently popular methods used for computing second-order sensitivities, ultimately highlighting the fundamentally new and distinctive features of the 2nd-ASAM first presented in the previous Sections of this work.

The earliest attempts at computing second-order response derivatives to model parameters for large-scale systems appear were made in the field of nuclear reactor physics, within the context of the neutron and photon diffusion and/or transport equations, in the 1970s, under the name of “(generalized) perturbation theory,” as typified in [12-15]. Although these early works addressed the formalities of computing second-order sensitivities of functional-type response to neutron cross sections (“system parameters”), the actual computations were limited to a few a priori chosen parameters, leaving the impression that second-order response derivatives were rather insignificant for reactor physics problems. These early works in the field of reactor physics were extended in the 1980s to more complicated problems [16-19], but still without enabling the exact and comprehensive computation of second-order response sensitivities. In parallel to this work (in nuclear reactor physics) during the 1980s, several authors [20-23] developed direct and adjoint approaches to derive second-order sensitivity derivatives of structural response with respect to stiffness parameters. However, these approaches were very narrowly focused, and the illustrative examples were extremely simple algebraic models, not generally applicable to any other field.

As was mentioned in the “Introduction,” [see, e.g., refs. 6-8], optimization methods using adjoint functions have been developed to compute efficiently *the vector resulting from the multiplication of an arbitrary vector* (usually a vector of steepest descent) with *the matrix of second-order response derivatives with respect to the state-variables*. Note that these

Hessians are *not* matrices of second-order response sensitivities with respect to the model “input” parameters (but are, as mentioned, response derivatives with respect to the state-variables, and are therefore significantly smaller in size than the matrix of second-order response sensitivities to parameters!) Inspired by these methods, Cacuci [24] has developed a procedure which computes efficiently, albeit still *approximately*, the second-order response derivatives with respect to the model “input” parameters, together with second-order response moments, using the 1st-level adjoint sensitivity system (*Ist-LASS*), while also computing the quantity

$$q_k(\rho_{ij}) \equiv \sum_{i=1}^{N_\alpha} \sum_{j=1}^{N_\alpha} \frac{\partial^2 r_k(\boldsymbol{\alpha}^0)}{\partial \alpha_i \partial \alpha_j} \text{cov}(\alpha_i, \alpha_j) = \sum_{i=1}^{N_\alpha} \sum_{j=1}^{N_\alpha} \frac{\partial^2 r_k(\boldsymbol{\alpha}^0)}{\partial \alpha_i \partial \alpha_j} \rho_{ij} \sigma_i \sigma_j. \quad (42)$$

In the above expression, ρ_{ij} denotes the correlation between parameters α_i and α_j , while σ_i and σ_j denote the standard deviations of the respective parameters. All of the derivatives appearing in the above expressions are evaluated at the nominal parameter values $\boldsymbol{\alpha}^0$. The above quantity appears the expression of the expected value, $E(r_k)$, of the response r_k

$$E(r_k) = r_k^c(\boldsymbol{\alpha}^0) + \frac{1}{2} \sum_{i=1}^{N_\alpha} \sum_{j=1}^{N_\alpha} \frac{\partial^2 r_k(\boldsymbol{\alpha}^0)}{\partial \alpha_i \partial \alpha_j} \text{cov}(\alpha_i, \alpha_j). \quad (43)$$

The second-order response sensitivities also appear in the expression of the higher-order response moments $\text{cov}(r_k, r_\ell)$, $\text{var}(r_k)$, $\mu_3(r_k)$ and $\mu_4(r_k)$, when these are computed using the propagation of errors method [see, e.g., ref. 4 and Part II of this work]. For fully correlated parameters, i.e., when $\rho_{ij} = 1$, the quantity $q_k(\rho_{ij} = 1)$ can be computed very efficiently considering the (column) vector of standard deviations $\boldsymbol{\sigma} \equiv (\sigma_1, \sigma_2, \dots, \sigma_{N_\alpha})$ and by noting that the Hessian-vector product $\nabla^2 r_k(\boldsymbol{\alpha}^0) \boldsymbol{\sigma}$ can be quantified most efficiently by using two computations of the gradients $\nabla_\alpha(r_k)$, as follows:

$$\nabla^2 r_k(\boldsymbol{\alpha}^0) \boldsymbol{\sigma} \cong \varepsilon^{-1} \left\{ \nabla_{\alpha} [r_k(\boldsymbol{\alpha}^0 + \varepsilon \boldsymbol{\sigma})] - \nabla_{\alpha} [r_k(\boldsymbol{\alpha}^0)] \right\} \equiv \mathbf{y}_k(\boldsymbol{\alpha}^0), \quad k = 1, \dots, N_r. \quad (44)$$

where $\varepsilon \leq O(10^{-5})$ is a small scalar quantity. The gradients $\nabla_{\alpha}(r_k)$ can be computed most efficiently using the I^{st} ASAM (one I^{st} -LASS computation per response yields all of the gradients $\nabla_{\alpha}(r_k)$ with respect to the parameters $\boldsymbol{\alpha}$). Thus, the vector $\mathbf{y}_k(\boldsymbol{\alpha}^0)$ defined above can be obtained using *two* I^{st} -LASS computations. Once $\mathbf{y}_k(\boldsymbol{\alpha}^0)$ has been obtained, the quantity $q_k(\rho_{ij} = 1)$ can be obtained by an additional vector scalar-product computation of the form

$$q_k(\rho_{ij} = 1) = \boldsymbol{\sigma}^{\dagger} \mathbf{y}_k(\boldsymbol{\alpha}^0). \quad (45)$$

In practice, however, the model parameters $\boldsymbol{\alpha}$ are extremely unlikely to be fully correlated. More likely, they are partially (positively or negatively) correlated or uncorrelated. In such cases, it is still possible to reduce the number of computations required for obtaining the quantity $q_k(\rho_{ij})$, from $O(N_{\alpha})^2$ to at most $n_k \leq (N_{\alpha} + 1)$ operations for each response $r_k(\boldsymbol{\alpha}^0)$, $k = 1, \dots, N_r$, where n_k is a number to be determined as discussed below, by using Hessian-vector products of the same form as shown in Eq. (44), as follows.

- (i) The leading contribution to the response variance [see, e.g., ref 4 or Part II of this work] stems from the expression

$$\text{var}(r_k) \cong \sum_{i=1}^{N_{\alpha}} \left(\frac{\partial r_k}{\partial \alpha_i} \right)^2 \sigma_i^2 + \sum_{i=1}^{N_{\alpha}} \sum_{j \neq i}^{N_{\alpha}} \left(\frac{\partial r_k}{\partial \alpha_i} \frac{\partial r_k}{\partial \alpha_j} \right) \rho_{ij} \sigma_i \sigma_j. \quad (46)$$

- (ii) Each of the terms in the first sum on the right side of Eq. (46) is positive, unless some response sensitivity happens to vanish, e.g., $\partial r_k / \partial \alpha_i = 0$. However, even when $\partial r_k / \partial \alpha_i = 0$ for some parameter α_i , the other sensitivities are non-zero (i.e., $\partial r_k / \partial \alpha_{j \neq i} \neq 0$), unless the response happens to be computed at a critical point $\boldsymbol{\alpha}^{crit}$. In such a highly unusual case, the number n_k --to be used in the sequel-- would be set to $n_k = N_\alpha$ since all of the second-order response derivatives would need to be computed in order to quantify the response moments.
- (iii) On the other hand, in the highly likely situation when $\boldsymbol{\alpha}^0 \neq \boldsymbol{\alpha}^{crit}$, not all of the sensitivities $\partial r_k / \partial \alpha_i$ would vanish. For the non-vanishing sensitivities, rank the relative sensitivities $t_{ki}^{r\alpha}(\boldsymbol{\alpha}^0) \equiv [\partial r_k(\boldsymbol{\alpha}^0) / \partial \alpha_i] [\alpha_i^0 / r_k(\boldsymbol{\alpha}^0)]$ in decreasing order of absolute magnitudes $|t_{ki}^{r\alpha}(\boldsymbol{\alpha}^0)|$. This ranking will also indicate the most important parameters α_i in contributing to the uncertainties in the response $r_k(\boldsymbol{\alpha}^0)$.
- (iv) Compute the quantities $c_{ii}^k \equiv \left\{ t_{ki}^{r\alpha}(\boldsymbol{\alpha}^0) [\alpha_i^0 / r_k(\boldsymbol{\alpha}^0)] \sigma_i \right\}^2$ and rank them in descending order of their magnitudes. As Eq.(46) indicates, c_{ii}^k represents the contribution to $\text{var}(r_k)$ stemming solely from uncertainties in parameter α_i .
- (v) Based on the rankings of the partial variances c_{ii}^k , rank the standard deviations σ_i in order of importance in contribution to $\text{var}(r_k)$, thus constructing the vector sequence $(\sigma_1, \sigma_2, \dots, \sigma_{n_k})$, with $n_k \leq (N_\alpha + 1)$, where σ_{n_k} denotes a user-

defined cut-off value corresponding to a negligible (from the user's point of view) contribution to $\text{var}(r_k)$.

- (vi) For each of selected variance in the sequence $(\sigma_1, \sigma_2, \dots, \sigma_{n_k})$, construct the n_k -dimensional column vectors $\sigma_1 \equiv (\sigma_1, 0, \dots, 0)$, ..., $\sigma_i \equiv (0, \dots, \sigma_i, \dots, 0)$, ..., $\sigma_{n_k} \equiv (0, \dots, 0, \dots, \sigma_{n_k})$, and use each of these, in turn, *in conjunction with one l^{st} -LASS (adjoint) computation*, to obtain the n_k -dimensional vector $t^{ki} \equiv (t_1^{ki}, \dots, t_{n_k}^{ki})^\top$ defined as

$$\begin{aligned} t^{ki} &\equiv (t_1^{ki}, \dots, t_{n_k}^{ki})^\top \equiv \left(\frac{\partial^2 r_k(\mathbf{a}^0)}{\partial \alpha_1 \partial \alpha_i} \sigma_i, \dots, \frac{\partial^2 r_k(\mathbf{a}^0)}{\partial \alpha_{n_k} \partial \alpha_i} \sigma_i \right)^\top \\ &= \nabla^2 \mathbf{r}_k(\mathbf{a}^0) \sigma_i \cong \varepsilon^{-1} \left\{ \nabla_\alpha [\mathbf{r}_k(\mathbf{a}^0 + \varepsilon \sigma_i)] - \nabla_\alpha [\mathbf{r}_k(\mathbf{a}^0)] \right\}; \quad i = 1, \dots, n_k; \end{aligned} \quad (47)$$

- (vii) For each $i = 1, \dots, n_k$, multiply the first component, t_1^{ki} , of the vector $t^{ki} \equiv (t_1^{ki}, \dots, t_{n_k}^{ki})^\top$ obtained in Eq. (47) by the correlation coefficient ρ_{i1} , the second component of this vector by the correlation coefficient ρ_{i2} , and so on, until the last component $t_{n_k}^{ki}$, which is to be multiplied by ρ_{in_k} , in order to construct the sequence of vectors

$$(t_1^{ki} \rho_{i1}, \dots, t_{n_k}^{ki} \rho_{in_k})^\top = \left(\frac{\partial^2 r_k(\mathbf{a}^0)}{\partial \alpha_1 \partial \alpha_i} \sigma_i \rho_{i1}, \dots, \frac{\partial^2 r_k(\mathbf{a}^0)}{\partial \alpha_{n_k} \partial \alpha_i} \sigma_i \rho_{in_k} \right)^\top; \quad i = 1, \dots, n_k; \quad (48)$$

- (viii) Sum up the corresponding components of the vectors in Eq. (48) to construct the column vector

$$\left(\sum_{i=1}^{n_k} t_1^{ki} \rho_{i1}, \dots, \sum_{i=1}^{n_k} t_{n_k}^{ki} \rho_{in_k} \right)^\dagger = \left(\sum_{i=1}^{n_k} \frac{\partial^2 r_k(\mathbf{a}^0)}{\partial \alpha_1 \partial \alpha_i} \sigma_i \rho_{i1}, \dots, \sum_{i=1}^{n_k} \frac{\partial^2 r_k(\mathbf{a}^0)}{\partial \alpha_{n_k} \partial \alpha_i} \sigma_i \rho_{in_k} \right)^\dagger; \quad (49)$$

(ix) Form the scalar product of the above vector with the *row-vector*

$(\sigma_1, \sigma_2, \dots, \sigma_{n_k})$ to obtain:

$$(\sigma_1, \sigma_2, \dots, \sigma_{n_k}) \left(\sum_{i=1}^{n_k} t_1^{ki} \rho_{i1}, \dots, \sum_{i=1}^{n_k} t_{n_k}^{ki} \rho_{in_k} \right)^\dagger = \sum_{i=1}^{n_k} \sum_{j=1}^{n_k} \frac{\partial^2 r_k(\mathbf{a}^0)}{\partial \alpha_i \partial \alpha_j} \rho_{ij} \sigma_i \sigma_j. \quad (50)$$

Thus, the sum $\sum_{i=1}^{n_k} \sum_{j=1}^{n_k} \frac{\partial^2 r_k(\mathbf{a}^0)}{\partial \alpha_i \partial \alpha_j} \rho_{ij} \sigma_i \sigma_j$ can be computed most efficiently by needing only

$n_k + 1 \leq N_\alpha$ *Ist-LASS (adjoint)* computations per response. Furthermore, the second-order

derivatives $\partial^2 r_k / \partial \alpha_i^2$ of the response r_k can be computed by dividing each of the components

of $\mathbf{t}^{ki} \equiv (t_1^{ki}, \dots, t_{n_k}^{ki})^\dagger$ through σ_i , as follows:

$$\left(\frac{t_1^{ki}}{\sigma_i}, \dots, \frac{t_{n_k}^{ki}}{\sigma_i} \right)^\dagger = \left(\frac{\partial^2 r_k(\mathbf{a}^0)}{\partial \alpha_1 \partial \alpha_i}, \dots, \frac{\partial^2 r_k(\mathbf{a}^0)}{\partial \alpha_{n_k} \partial \alpha_i} \right)^\dagger; \quad i = 1, \dots, n_k. \quad (51)$$

Note also that the computation of $\mathbf{t}^{ki} \equiv (t_1^{ki}, \dots, t_{n_k}^{ki})^\dagger$ via Eq (47) does not ensure that the mixed

second-order derivatives will be symmetric in the order of differentiation. Small asymmetries

of the form

$$\partial^2 r_k(\mathbf{a}^0) / \partial \alpha_i \partial \alpha_j = \partial^2 r_k(\mathbf{a}^0) / \partial \alpha_j \partial \alpha_i + O(\varepsilon), \quad (52)$$

will occur due to the one-sided differencing combined with the finiteness of ε . Using a second-order symmetrical differencing form instead of Eq. (47) error will reduce the error on the right-side of Eq. (52) to $O(\varepsilon^2)$, but at the expense of needing twice (i.e., $2n_k$) as many large-scale computations as when using Eq. (47).

The above method was used [25-27] by Cacuci and (his PhD students) Proctor and Latten to perform (approximate) second-order sensitivity analysis of reactor physics experimental and computational benchmarks. Notably, ideas similar to those underlying Eqs. (44) and (47) above, namely the use of finite differences in combination with adjoint functions/systems, were also used in [28] and [29] to compute approximate second-order response sensitivities for reliability evaluation of electromagnetic devices [28] and for estimating second-order sensitivities exploiting time-domain transmission-line modeling [29].

The approaches to computing second-order response sensitivities to “input parameters” mentioned in the forgoing were all “deterministic” in nature. In parallel to these deterministic approaches, several statistical methods were developed, the vast majority of them for computing *approximately* the *first-order response* sensitivities to parameters. As discussed in the review presented in [5], perhaps the most influential statistical method for computing approximately higher-order contributions to the uncertainty in a multivariate function of uncertain parameters was “Sobol’s Method” [30], and ANOVA-like decompositions (see, e.g., [31]). Sobol’s Method is based on a particular case of Kolmogorov’s decomposition theorem, which states that any multivariate function, $f(x_1, \dots, x_n)$, defined on the unit cube $[0, 1]^n$ can be written as a linear superposition of univariate functions $h_j(x_i)$, of the form

$$f(x_1, \dots, x_n) = \sum_{j=1}^{2n+1} g \left[a_1 h_j(x_1) + a_2 h_j(x_2) + \dots + a_n h_j(x_n) \right], \quad (53)$$

where the functions $h_j(x_i)$ are continuous, albeit highly non-smooth. Sobol’s decomposition takes on the following particular form of the above superposition:

$$f(\mathbf{x}) \equiv f(x_1, \dots, x_n) = f_0 + \sum_{i=1}^n f_i(x_i) + \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) + \dots + f_{12\dots n}(x_1, x_2, \dots, x_n), \quad (54)$$

where

- (i) $f_0 \equiv \int_{[0,1]^n} f(\mathbf{x}) d\mathbf{x} = \text{constant},$
- (ii) $\int_{[0,1]^n} f_{i_1 i_2 \dots i_n}(\mathbf{x}) f_{j_1 j_2 \dots j_m}(\mathbf{x}) d\mathbf{x} = 0, \text{ if } (i_1, i_2, \dots, i_n) \neq (j_1, j_2, \dots, j_m);$
- (iii) $\int_0^1 f_{i_1 i_2 \dots i_n}(x_{i_1}, x_{i_2}, \dots, x_{i_n}) dx_{i_m} = 0, \text{ if } 1 \leq m \leq n.$

By squaring Eq. (54) and integrating the resulting expression over the unit cube $[0,1]^n$, the following relation is obtained for the total variance, denoted as D , of $f(\mathbf{x})$:

$$D \equiv \int_{[0,1]^n} f^2(\mathbf{x}) d\mathbf{x} - f_0^2 = \sum_{i=1}^n D_i + \sum_{1 \leq i < j \leq n} D_{ij} + \dots + D_{12\dots n}, \quad (55)$$

where the partial variances of $f(\mathbf{x})$ are defined as :

$$D_{i_1 i_2 \dots i_m} \equiv \int_0^1 \dots \int_0^1 f_{i_1 i_2 \dots i_m}(x_{i_1}, x_{i_2}, \dots, x_{i_m}) dx_{i_1} \dots dx_{i_m}, \text{ for } 1 \leq i_1 < \dots < i_m \leq n, \text{ } m=1, \dots, n. \quad (56)$$

Finally, the sensitivity indices are defined as

$$S_{i_1 i_2 \dots i_m} \equiv D_{i_1 i_2 \dots i_m} / D, \text{ for } 1 \leq i_1 < \dots < i_m \leq n, \text{ } m=1, \dots, n. \quad (57)$$

In view of the above definitions, the first-order sensitivity index, S_i , for the parameter x_i indicates the fractional contribution of x_i to the variance D of $f(\mathbf{x})$. The second-order sensitivity index, S_{ij} , ($i \neq j$), measures the part of the variation in $f(\mathbf{x})$ due to x_i and x_j that cannot be explained by the sum of the individual effects of x_i and x_j ; and so on. Note also that Eqs. (56) and (57) imply that

$$\sum_{i=1}^n S_i + \sum_{1 \leq i < j \leq n} S_{ij} + \dots + S_{12\dots n} = 1. \quad (58)$$

In the absence of a priori knowledge regarding the relationship between the parameters x_i and the response $f(\mathbf{x})$, the construction of a full-space sensitivity analysis typically requires $O(s^I)$ large-scale computations, where s is the sample size and I denotes the number of model parameters. For example, if the response $f(\mathbf{x})$ is a function of $I = 8$ parameters, and if the sample size is $s = 64$, then Sobol's method requires 1088 model evaluations; if the sample size increases to 1024, then Sobol's method requires 17 408 model evaluations; and so on.

Attempting to reduce the number of model evaluations inherent to Sobol's method, Rabitz and co-workers (see, e.g., Ref [32]), have introduced the so-called *high-dimensional model representations (HDMRs)*, which aim at identifying relationships between sets of inputs (e.g., parameters) and outputs (responses). HDMRs express the model output as a finite additive sum of correlated functions with increasing numbers of input variables up to the total number of inputs, in the same spirit as Sobol's expansion shown in Eq. (54). The resultant HDMR expansion is subsequently used as a *reduced-order surrogate model* that depends on fewer parameters and, more importantly, fewer coupled parameters (interactions among parameters). The two commonly used expansions, called ANOVA-HDMR and Cut-HDMR, respectively. ANOVA-HDMR requires computationally expensive numerical integrations of the model over the entire parameter space, while Cut-HDMR requires response-function evaluations along the so-called cut lines or hyperplanes with respect to pre-selected reference points in the domain of the inputs. Although Cut-HDMR is computationally less expensive, it is also less accurate and its accuracy depends strongly on the choice of the reference point; constructing a second-order Cut-HDMR expansion can still be computationally expensive. Recently, Hu et al [33] have improved the efficiency of the second-order Cut-HDMR by first using a screening algorithm (based on the so-called "New Morris Method") to pre-screen the parameter (input) space for significant inputs and interactions while eliminating parameters that were deemed insignificant, and subsequently applied the Cut-HDMR expansion in the reduced parameter space.

Statistical methods (e.g., Sobol’s ANOVA-HDMR, Cut-HDMR) are non-intrusive, requiring no (adjoint) model developments, so they can be readily used when expediency is paramount. However, statistical methods cannot compute *exactly* any response sensitivities (of any order), and become prohibitively expensive for routine sensitivity and uncertainty analyses of the large-scale models with many (hundreds and beyond) input parameters customarily encountered in practice.

5. CONCLUSIONS

This work has presented the *second-order forward and adjoint sensitivity analysis methodologies* (2^{nd} -FSAM and 2^{nd} -ASAM) for computing exactly and efficiently the second-order functional derivatives of system responses (i.e., “system performance parameters”) to the system’s model parameters. The definition of “system parameters” used in this work include, in the most comprehensive sense, all computational input data, correlations, initial and/or boundary conditions, etc. The 2^{nd} -ASAM builds on the first-order adjoint sensitivity analysis methodology (1^{st} -ASAM) for nonlinear systems introduced ([1], [2]) and developed ([3]-[5]) by Cacuci; see also [6]. For a physical system comprising N_α parameters and N_r responses, we note the following essential computational properties of the 2^{nd} -FSAM and 2^{nd} -ASAM, respectively:

- (i) The 2^{nd} -FSAM requires N_α large-scale forward model computations to obtain all of the first-order response sensitivities $\{\partial R(\mathbf{u}, \boldsymbol{\alpha})/\partial \alpha_i\}_{(\mathbf{u}^0, \boldsymbol{\alpha}^0)}$, $i = 1, \dots, N_\alpha$, and an additional number $N_\alpha(N_\alpha + 1)/2$ of large-scale computations to determine the sensitivities $\{\partial^2 R(\mathbf{u}, \boldsymbol{\alpha})/\partial \alpha_j \partial \alpha_i\}_{(\mathbf{u}^0, \boldsymbol{\alpha}^0)}$, $i, j = 1, \dots, N_\alpha$, for a total of $(N_\alpha^2/2 + 3N_\alpha/2)$ large scale computations for obtaining all of the first- and second-order sensitivities, for all N_r system responses;
- (ii) On the other hand, for one functional-type system response, the 2^{nd} -ASAM requires one large-scale computation using the *first-level adjoint sensitivity system* (1^{st} -LASS) for obtaining all of the first-order sensitivities $\{\partial R(\mathbf{u}, \boldsymbol{\alpha})/\partial \alpha_i\}_{(\mathbf{u}^0, \boldsymbol{\alpha}^0)}$, $i = 1, \dots, N_\alpha$, followed by at most N_α large-scale computations using the *second-level adjoint sensitivity systems* (2^{nd} -LASS) for

obtaining exactly all of the second-order sensitivities

$$\left\{ \partial^2 R(\mathbf{u}, \boldsymbol{\alpha}) / \partial \alpha_j \partial \alpha_i \right\}_{(\mathbf{u}^0, \boldsymbol{\alpha}^0)}, \quad i, j = 1, \dots, N_\alpha, \dots$$

The above considerations clearly highlight the fact that the 2^{nd} -FSAM should be used when $N_r \gg N_\alpha$, while the 2^{nd} -ASAM should be used when $N_\alpha \gg N_r$, which is the most often encountered situation in practice. The 2^{nd} -ASAM presented, in premiere, in this work should enable the hitherto intractable exact computation of all of the second-order response sensitivities (i.e., functional Gateaux-derivatives) to the large-number of parameters that are characteristically encountered in the large-systems of practical interest. Additional work is currently in progress to generalize the 2^{nd} -ASAM to computing efficiently and exactly the third- and higher-order response sensitivities. In the accompanying PART II [11], we present an illustrative application of the 2^{nd} -ASAM to a paradigm particle diffusion problem that admits a unique analytical solution, thereby making transparent the mathematical derivations presented in this paper. Very importantly, this illustrative application will show that:

- (i) The construction and solutions of the *second-level adjoint sensitivity systems* (2^{nd} -LASS) requires very little additional effort beyond the construction of the first-level adjoint sensitivity system (1^{st} -LASS) needed for computing the first-order sensitivities; and
- (ii) The actual number of adjoint computations needed for computing all of the first- and second-order response sensitivities will turn out to be considerably less than N_α computations per functional-type response.

We note that the 2^{nd} -ASAM was developed in this work using *real* (as opposed to complex) Hilbert spaces, since real Hilbert spaces provide the natural mathematical setting for computational purposes. This setting does *not* restrict, in any way, the generality of the 2^{nd} -ASAM theory presented here. The 2^{nd} -ASAM theory can be readily set in complex Hilbert spaces by simply changing some terminology, without affecting its substance.

The use of the 2^{nd} -ASAM to compute exactly all of the second-order response sensitivities to model input parameters is expected to induce significant advances in related scientific disciplines, particularly the areas of uncertainty quantification, model validation, reduced-order modeling, and predictive modeling/data assimilation. Combining the 2^{nd} -ASAM with Cut-HDMR, for example, appears to be a timely endeavor in this vein.

ACKNOWLEDGMENTS

This work was partially supported by contract 15540-FC51 from Gen4Energy, Inc., and partially by contract 15540-FC59 from the US Department of Energy (NA-22), respectively, with the University of South Carolina.

REFERENCES

- [1] D.G. Cacuci, “Sensitivity Theory for Nonlinear Systems: I. Nonlinear Functional Analysis Approach”, *J. Math. Phys.*, **22**, (1981) 2794-2802.
- [2] D.G. Cacuci, “Sensitivity Theory for Nonlinear Systems: II. Extensions to Additional Classes of Responses”, *J. Math. Phys.*, **22**, (1981) 2803-2812.
- [3] D.G. Cacuci, “The Forward and the Adjoint Methods of Sensitivity Analysis”, **Chapter 3**, in “*Uncertainty Analysis*”, Y. Ronen, Ed., CRC Press, Inc., Boca Raton, Florida, 1988 pp. 71-144,.
- [4] D. G. Cacuci, *Sensitivity and Uncertainty Analysis: Theory*, **Vol.1**, Chapman & Hall/CRC, Boca Raton, 2003.
- [5] D.G. Cacuci, M. Ionescu-Bujor, and M. I. Navon, *Sensitivity and Uncertainty Analysis: Applications to Large Scale Systems*, **Vol. 2**, Chapman & Hall/CRC, Boca Raton, 2005.
- [6] Tamás Práger and Fanni Dóra Kelemen, “Adjoint Methods and their Application in Earth Sciences”, **Chapter 4A**, in *Advanced Numerical Methods for Complex Environmental Models: Needs and Availability*, I. Faragó, Á. Havasi, and Z. Zlatev, Editors, Bentham Science Publishers, 2014, pp 203-275.
- [7] D.G. Cacuci, M.I. Navon, and M. Ionescu-Bujor, *Computational Methods for Data Evaluation and Assimilation*, Chapman & Hall/CRC, Boca Raton, 2014.
- [8] Dacian N. Daescu and Ionel M. Navon, “Sensitivity Analysis in Nonlinear Variational Data Assimilation: Theoretical Aspects and Applications”, *Ch 4B, Advanced Numerical Methods for Complex Environmental Models: Needs and Availability*, I. Faragó, Á. Havasi, and Z. Zlatev, Editors, Bentham Science Publishers, 2014, in pp. 276-300.
- [9] D.G. Cacuci, “Global Optimization and Sensitivity Analysis”, *Nucl. Sci. Eng.*, **104**, 78-88, 1990.
- [10] D. G. Cacuci, “Predictive Modeling of Coupled Multi-Physics Systems: I. Theory”, *Annals of Nuclear Energy*, **70**, 266–278, 2014.
- [11] Dan G. Cacuci, “Second-Order Adjoint Sensitivity Analysis Methodology (2nd-ASAM) for Computing Exactly and Efficiently First- and Second-Order Sensitivities in Large-

- Scale Linear Systems: II. Illustrative Application to a Paradigm Particle Diffusion Problem,” *J. Comp. Physics*, *accepted*, 2014.
- [12] H. Mitani, “Higher-Order Perturbation Method in Reactor Calculation”, *Nucl. Sci. Eng.*, **51**, 180-188, 1973.
- [13] Y. Seki, “Evaluation of the Second-Order Perturbation Terms by the Generalized Perturbation Method”, *Nucl. Sci. Eng.*, **51**, 243-251, 1973.
- [14] A. Gandini, “Implicit and Explicit Higher-Order Perturbation Methods for Nuclear Reactor Analysis”, *Nucl. Sci. Eng.*, **67**, 347-355, 1978.
- [15] E. Greenspan, D. Gilai, and E. M. Obloy, “Second-Order Generalized Perturbation Theory for Source-Driven Systems”, *Nucl. Sci. Eng.*, **68**, 1-8, 1978.
- [16] J. Barhen and D.G. Cacuci, “An Efficient Procedure to Evaluate Second-Order Sensitivities for Nonlinear Problems with Nonlinear Responses”, *Trans. Am. Nucl. Soc.*, **35**, 266 (1980).
- [17] D.G. Cacuci, E. Greenspan, J.H. Marable, and M.L. Williams, “Developments in Sensitivity Theory”, in ANS Topical Conference “1980 Advances in Reactor Physics and Shielding”, Sun Valley, Idaho, September 14-17, NAS/70048, 1980, 692-704, 1980.
- [18] E. Greenspan, “New Developments in Sensitivity Theory”, *Ch. 7 in Advances in Nuclear Science and Technology*, **Vol 14**, J. Lewins and M. Becker, Eds, Plenum Press, 1982.
- [19] A. Gandini, “Generalized Perturbation Theory (GPT). A Heuristic Approach”, *Ch. 7 in Advances in Nuclear Science and Technology*, **Vol 19**, J. Lewins and M. Becker, Eds, Plenum Press, 1987.
- [20] E. J. Haug, “Second-Order Design Sensitivity Analysis of Structural Systems,” *AIAA Journal*, **Vol. 19**, 1087-1088, 1981.
- [21] R. T. Haftka, “Second-Order Sensitivity Derivatives in Structural Analysis,” *AIAA Journal*, **Vol. 20**, 1765-1766, 1982.
- [22] E. J. Haug and P. E. Ehle, “Second-Order Design Sensitivity Analysis of Mechanical System Dynamics,” *International Journal for Numerical Methods in Engineering*, **Vol. 18**, 637-661, 1982.
- [23] R. T. Haftka and Z. Mroz, “First- and Second-Order Sensitivity Analysis of Linear and Nonlinear Structures” *AIAA Journal*, **Vol. 24**, 1187-1192, 1986.
- [24] D.G. CACUCI, “High-Order Quantitative Model Validation Metrics Integrating Experimental and Computational Data for Large-Scale Time-Independent Nonlinear Systems”, CASL Milestone #489: L3: VUQ.VVDA.P4.05 (VUQ.P4.02), ORNL, December 2011.

- [25] W. C. Proctor, R. T. Evans, D. G. Cacuci, "Second-Order Sensitivity Analysis of Godiva and Jezebel Experimental Benchmarks", *Trans. Am. Nucl. Soc.*, **106.1**, 429, 2012.
- [26] C. Latten and D. G. Cacuci, "High-Order Response Moments for Model Uncertainty Quantification and Validation", *Trans. Am. Nucl. Soc.*, **106.1**, 759, 2012.
- [27] W.C. Proctor, "Elements of High-Order Predictive Model Calibration Algorithms with Applications to Large-Scale Reactor Physics Systems", PhD Dissertation, North Carolina State University, March 27, 2012.
- [28] Z. Ren and C. S. Koh, "A Second-Order Design Sensitivity-Assisted Monte Carlo Simulation Method for Reliability Evaluation of the Electromagnetic Devices," *J Electr. Eng. Technol.*, **8**, No. 4, 780-786, 2013.
- [29] M. H. Negm, M. H. Bakr, O. S. Ahmed, N. K. Nikolova, and J. W. Bandler, "Wideband Second-Order Adjoint Sensitivity Analysis Exploiting TLM," *IEEE Trans. Microwave Theory and Techniques*, **62**, No. 3, 389-398, 2014.
- [30] I.M. Sobol', "Sensitivity Analysis for Nonlinear Mathematical Models," *Math. Model. Comput. Exp.*, **1**, 407, 1993.
- [31] G. Archer, A. Saltelli, and I. M. Sobol', "Sensitivity Measures, ANOVA-like Techniques, and the use of Bootstrap," *J. Statist. Comp. Simul.*, **58**, 99, 1997.
- [32] H. Rabitz and O. F. Alis, "General Foundation of High Dimensional Model Representations," *J. Math. Chem.*, **25**, 197, 1999.
- [33] Z. Hu, R. C. Smith, J. Willert, and C. T. Kelley, "High-Dimensional Model Representations for the Neutron Transport Equation", *Nucl. Sci. Eng.*, **177**, 350–360, 2014.