Introducing the Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Volterra-Type: Mathematical Methodology and Illustrative Application to Nuclear Engineering

Dan Gabriel Cacuci [*]

Article

Introducing the Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Volterra-Type: Mathematical Methodology and Illustrative Application to Nuclear Engineering

Dan Gabriel Cacuci

University of South Carolina, Department of Mechanical Engineering, Columbia, SC 29201, USA; cacuci@cec.sc.edu

Abstract: This work presents the general mathematical frameworks of the "First and Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Volterra Type" designated as the 1st-FASAM-NIE-V and the 2nd-FASAM-NIE-V methodologies, respectively. Using a single large-scale (adjoint) computation, the 1st-FASAM-NIE-V enables the most efficient computation of the exact expressions of all first-order sensitivities of the decoder response to the feature functions and also with respect to the optimal values of the NIE-net's parameters/weights after the respective NIE-Volterra-net was optimized to represent the underlying physical system. The 2nd-FASAM-NIE-V requires as many large-scale computations as there are first-order sensitivities of the decoder response with respect to the feature functions. Subsequently, the second-order sensitivities of the decoder response with respect to the primary model parameters are obtained trivially by applying the "chain-rule of differentiation" to the second-order sensitivities with respect to the feature functions. The application of the 1st-FASAM-NIE-V and the 2nd-FASAM-NIE-V methodologies is illustrated by using a well-known model for neutron slowing down in a homogeneous hydrogenous medium, which yields tractable closed-form exact explicit expressions for all quantities of interest, including the various adjoint sensitivity functions and first- and second-order sensitivities of the decoder response with respect to all feature functions and also primary model parameters.

Keywords: first-order features adjoint sensitivity analysis; second-order features adjoint sensitivity analysis; Volterra neural integral equation; neutron slowing down

1. Introduction

It is well known that Neural Ordinary Differential Equations (NODEs) have enabled the use of deep learning for modeling discretely sampled dynamical systems. NODEs [1–3] provide a flexible trade-off between efficiency, memory costs and accuracy while bridging traditional numerical modeling with modern deep learning, as demonstrated by various applications, including time-series, dynamics and control [1–9]. However, since each time-step is determined locally in time, NODEs are limited to describing systems that are instantaneous. On the other hand, integral equations (IE) model global "long-distance" spatiotemporal relations, and IE solvers often possess stability properties that are superior to solvers for ordinary and/or partial differential equations. Therefore, differential equations are occasionally recast in integral-equation forms that can be solved more efficiently using IE solvers, as exemplified by the applications described in [10–12].

Due to their non-local behavior, IE solvers are suitable for modeling complex dynamics, learning the operator underlying the system under consideration by using data sampled from the respective system. As discussed in [13], the operator learning problem is formulated on finite grids, using finite-

difference methods that approximate the domain of the functions under investigation; the learning is performed by using an IE solver which samples the domain of integration continuously. As shown in [14], Neural Integral Equations (NIEs) and the Attentional Neural Integral Equations (ANIEs) can be used to generate dynamics and infer the spatiotemporal relations that initially generated the data, thus enabling the continuous learning of non-local dynamics with arbitrary time resolution. The ANIE interprets the self-attention mechanism as the Nystrom method for approximating integrals [15], which enables efficient integration over higher dimensions, as discussed in [10–15] and references therein.

Neural nets are trained by minimizing a "loss functional" chosen by the user to represent the discrepancy between the output produced by the neural net's decoder and some user-chosen "reference solution." However, the physical system modeled by a neural net inevitably comprises imperfectly known parameters that stem from measurements and/or computations and are therefore afflicted by uncertainties that stem from the respective experiments and/or computations. Hence, even if the neural net reproduces perfectly a given state of a physical system, the neural net's "optimized weights" are subject to the uncertainties inherent in the parameters that characterize the underlying physical system, and these uncertainties inevitably propagate to the decoder's output response. It is hence important to quantify the impact of parameters/weights uncertainties on the uncertainties induced in the decoder's output response. This impact is quantified by the sensitivities of the decoder's response with respect to the optimized weights/parameters comprised within the neural net.

Neural nets comprise not only scalar-valued weights/parameters but also functions (e.g., correlations) of such scalar model parameters, which can be conveniently called "features of primary model parameters". Cacuci [16] has developed the "n^{th}-Order Features Adjoint Sensitivity Analysis Methodology for Nonlinear Systems (n^{th}-FASAM-N)", which enables the most efficient computation of the exact expressions of arbitrarily high-order sensitivities of model responses with respect to the model's "features". In turn, the sensitivities of the responses with respect to the primary model parameters are determined, analytically and trivially, by applying the "chain rule" to the expressions obtained for the response sensitivities with respect to the model's "features." The nth-FASAM-N [16] has been applied to develop general first- and second-order sensitivity analysis methodologies for NODEs [17] and for Neural Integral Equations of Fredholm-type [18], which enable the computation, with unsurpassed efficacy, of the exact expressions of first and second-order sensitivities of decoder responses with respect to the underlying neural net's optimized weights.

This work continues the application of the nth-FASAM-N [16] methodology to develop the "First- and Second-Order Methodologies for Neural Integral Equations of Volterra Type" (acronyms "1st-FASAM-NIE-V" and, respectively, "2nd-FASAM-NIE-V"). The 1st-FASAM-NIE-V methodology, which is presented in Section 2, enables the most efficient computation of exact expressions of all of the first-order sensitivities of NIE decoder responses with respect to all of the optimal values of the NIE-net's parameters/weights, after the respective NIE-Volterra-net was optimized to represent the underlying physical system. The efficiency of the 1st-FASAM-NIE-V is illustrated in Section 3 by applying it to perform a comprehensive first-order sensitivity analysis of the well-known model [19–21] of neutron slowing down in a homogeneous medium containing fissionable material.

The general mathematical framework of the 2nd-FASAM-NIE-Volterra methodology, which is presented in Section 4, enables the most efficient computation of the exact expressions of the second-order sensitivities of NIE decoder responses with respect to all of the optimal values of the NIE-net's parameters/weights. The efficiency of the 2nd-FASAM-NIE-V is illustrated in Section 5 by applying it to perform a comprehensive second-order sensitivity analysis of the neutron slowing down model [19–21] considered in Section 3. Section 6 concludes this work by presenting a discussion that highlights the unparalleled efficiency of the 2nd-FASAM-NIE-V methodology for performing sensitivity analysis of Volterra-type neural integral equations.

2. First-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Volterra-Type (1st-FASAM-NIE-V)

Following [14], a network of nonlinear "Neural Integral Equations of Volterra-type (NIE-Volterra)" can be represented by the system of coupled equations shown below:

$$h_i(t) = g_i\left[\mathbf{F}(\boldsymbol{\theta}); t\right] + \varphi_i\left[\mathbf{F}(\boldsymbol{\theta}); t\right] \int_{t_0}^{t} \psi_i\left[\mathbf{h}(\tau); \mathbf{F}(\boldsymbol{\theta}); \tau\right] d\tau; \ t_0 \le t \le t_f; \ 1 = 1,...TH. \quad (1)$$

The quantities appearing in Eq. (1) are defined as follows:

(i) The real-valued scalar quantities t, $t_0 \le t \le t_f$, and τ, $t_0 \le \tau \le t_f$, are time-like independent variables which parameterize the dynamics of the hidden/latent neuron units. Customarily, the variable t is called the "global time" while the variable τ is called the "local time." The initial time-value is denoted as t_0 while the stopping time-value is denoted as t_f.

(ii) The components of the vector $\boldsymbol{\theta} \triangleq \left[\theta_1,...,\theta_{TW}\right]^{\dagger}$ represent scalar learnable adjustable weights, where TW denotes the total number of adjustable weights in all of the latent neural nets. The components of the column-vector $\boldsymbol{\theta} \triangleq \left[\theta_1,...,\theta_{TW}\right]^{\dagger}$ are considered to be "primary parameters" while the components of the vector-valued function $\mathbf{F}(\boldsymbol{\theta}) \triangleq \left[F_1(\boldsymbol{\theta}),...,F_{TF}(\boldsymbol{\theta})\right]^{\dagger}$ represent the "feature" functions of the respective weights. The quantity TF denotes the "total number of feature/functions of the primary model parameters" comprised in the NIE-Volterra. In general, $\mathbf{F}(\boldsymbol{\theta})$ is a nonlinear function of $\boldsymbol{\theta}$. The total number of feature functions must necessarily be smaller than the total number of primary parameters (weights), i.e., $TF < TW$. In the extreme case when there are no feature functions, it follows that $F_i(\boldsymbol{\theta}) \equiv \theta_i$, for all $i = 1,...,TW \equiv TF$. In this work, all vectors are considered to be column vectors, and the dagger "$\dagger$" symbol will be used to denote "transposition.". The symbol "$\triangleq$" will be used to denote "is defined as" or, equivalently, "is by definition equal to."

(iii) The TH-dimensional vector-valued function $\mathbf{h}(t) \triangleq \left[h_1(t),...,h_{TH}(t)\right]^{\dagger}$ represents the hidden/latent neural networks. The quantity TH denotes the total number components of $\mathbf{h}(t)$. At the initial time-value t_0, the functions $h_i(t_0)$ take on the known values $h_i(t_0) = g_i\left[\mathbf{F}(\boldsymbol{\theta}); t_0\right]$.

(iv) The functions $g_i\left[\mathbf{F}(\boldsymbol{\theta}); t\right] = h_i(t_0)$, $i = 1,...,TH$, model the initial state ("encoder") of the network. The functions $\varphi_i\left[\mathbf{F}(\boldsymbol{\theta}); t\right]$ and $\psi_i\left[\mathbf{h}(\tau); \mathbf{F}(\boldsymbol{\theta}); \tau\right]$, $i = 1,...,TH$, depend nonlinearly on $\mathbf{h}(t)$ and $\mathbf{F}(\boldsymbol{\theta})$, respectively, and model the dynamics of the latent neurons.

The "training" of the NIE-Volterra net is accomplished by using the "adjoint" or other methods to minimize the user-chosen "loss functional" intended to represent the discrepancy between the output produced by the NIE-decoder and a "reference solution" chosen by the user. After the training is completed, the primary parameters ("weights") $\boldsymbol{\theta} \triangleq \left[\theta_1,...,\theta_{TW}\right]^{\dagger}$ will have been assigned "optimal" values which are obtained as a result of having minimized the chosen loss functional. These optimal values for the primary parameters ("weights") will be denoted using a superscript "zero," as follows: $\boldsymbol{\theta}^0 \triangleq \left[\theta_1^0,...,\theta_{TW}^0\right]^{\dagger}$. Using these optimal/nominal parameter values to solve the NIE-system will yield the optimal/nominal solution $\mathbf{h}^0(t) \triangleq \left[h_1^0(t),...,h_{TH}^0(t)\right]^{\dagger}$, which will satisfy the following form of Eq. (1):

$$h_i^0(t) = g_i^0\left[\mathbf{F}(\boldsymbol{\theta}^0); t\right] + \varphi_i^0\left[\mathbf{F}(\boldsymbol{\theta}^0); t\right] \int_{t_0}^{t} \psi_i^0\left[\mathbf{h}^0(\tau); \mathbf{F}(\boldsymbol{\theta}^0); \tau\right] d\tau; t_0 \le t \le t_f; 1 = 1,...TH. \quad (2)$$

After the NIE-net is optimized to reproduce the underlying physical system as closely as possible, the subsequent responses of interest are no longer "loss functions" but become specific functionals of NIE's "decoder" response/output. Such a decoder-response, which will be denoted as $R[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})]$, can be generally represented a scalar-valued functional of $\mathbf{h}(t)$ and $\mathbf{F}(\boldsymbol{\theta})$, defined as follows:

$$R[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})] = \int_{t_0}^{t_f} D[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t]\,dt \,.$$

(3)

The function $D[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t]$ models the decoder and may contain distributions (e.g., Dirac-delta and/or Heaviside functionals, etc.), if the decoder-response is to be evaluated at some particular point in time or over a subinterval within the interval $[t_0,t_f]$.

The optimal value of the decoder-response, denoted as $R[\mathbf{h}^0;\mathbf{F}(\boldsymbol{\theta}^0)]$, is represented by evaluating Eq. (3) at the optimal/nominal parameter values $\boldsymbol{\theta}^0 \triangleq [\theta_1^0,...,\theta_{TW}^0]^\dagger$ and optimal/nominal solution $\mathbf{h}^0(t)$, as follows:

$$R[\mathbf{h}^0;\mathbf{F}(\boldsymbol{\theta}^0)] = \int_{t_0}^{t_f} D[\mathbf{h}^0(t);\mathbf{F}(\boldsymbol{\theta}^0);t]\,dt \,.$$

(4)

The true values $\boldsymbol{\theta}$ of the primary parameters ("weights") that characterize the physical system modeled by the NIE-V net are afflicted by uncertainties inherent to the experimental and/or computational methodologies employed to model the original physical system. Therefore, the true values $\boldsymbol{\theta}$ of the primary parameters ("weights") will differ from the known nominal values $\boldsymbol{\theta}^0$ (which are obtained after training the NIE-net to represent the model of the physical system) by variations denoted as $\delta\boldsymbol{\theta} \triangleq \boldsymbol{\theta}-\boldsymbol{\theta}^0$. The variations $\delta\boldsymbol{\theta} \triangleq \boldsymbol{\theta}-\boldsymbol{\theta}^0$ will induce corresponding variations $\delta\mathbf{F} \triangleq \mathbf{F}(\boldsymbol{\theta})-\mathbf{F}^0$, $\mathbf{F}^0 \triangleq \mathbf{F}(\boldsymbol{\theta}^0)$, in the feature functions, which in turn will induce variations $\mathbf{v}^{(1)}(t) \triangleq [v_1^{(1)}(t),...,v_{TH}^{(1)}(t)]^\dagger$, $v_i^{(1)}(t) \triangleq h_i(t)-h_1^0(t)$, $i=1,...,TH$, around the nominal/optimal functions $\mathbf{h}^0(t)$. Subsequently, the variations $\delta\mathbf{F}$ and $\mathbf{v}^{(1)}(t;\mathbf{x})$ will induce variations in the NIE decoder's response.

The 1st-FASAM-IDE-V methodology for computing the first-order sensitivities of the decoder's response with respect to the NIE's weights will be established by applying the same principles as those underlying the 1st-FASAM-N [16] methodology. These first-order sensitivities are embodied in the first-order G-variation $\delta R(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)};\delta\mathbf{F})$ of the response $R[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})]$, for variations $\mathbf{v}^{(1)}(t)$ and $\delta\mathbf{F}$ around the nominal values $\mathbf{h}^0(t)$ and $\mathbf{F}^0$, which is by definition obtained as follows:

$$\delta R(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)};\delta\mathbf{F}) = \left\{ \frac{d}{d\varepsilon} \int_{t_0}^{t_f} D[\mathbf{h}^0(t)+\varepsilon\mathbf{v}^{(1)}(t);\mathbf{F}^0+\varepsilon\delta\mathbf{F};t]\,dt \right\}_{\varepsilon=0}$$

$$= \left\{ \delta R(\mathbf{h}^0;\mathbf{F}^0;\delta\mathbf{F}) \right\}_{dir} + \left\{ \delta R(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)}) \right\}_{ind} \,.$$

(5)

In Eq. (5), the "direct-effect term" $\left\{ \delta R(\mathbf{h}^0;\mathbf{F}^0;\delta\mathbf{F}) \right\}_{dir}$ arises directly from variations $\delta\mathbf{F}$ (which in turn stem from parameter variations $\delta\boldsymbol{\theta}$) and is defined as follows:

$$\left\{\delta R\left(\mathbf{h}^{0};\mathbf{F}^{0};\delta\mathbf{F}\right)\right\}_{dir} \triangleq \int_{t_0}^{t_f}\left\{\frac{\partial D\left[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t\right]}{\partial\mathbf{F}}\delta\mathbf{F}\right\}_{\left(\mathbf{h}^{0};\mathbf{F}^{0}\right)} dt$$

$$\equiv \sum_{j=1}^{TF}\int_{t_0}^{t_f}\left\{\frac{\partial D\left[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t\right]}{\partial F_j}\delta F_j\right\}_{\left(\mathbf{h}^{0};\mathbf{F}^{0}\right)} dt, \tag{6}$$

while the "indirect-effect term" $\left\{\delta R\left(\mathbf{h}^{0};\mathbf{F}^{0};\mathbf{v}^{(1)}\right)\right\}_{ind}$ arises through the variations $\mathbf{v}^{(1)}(t)$ in the hidden state functions $\mathbf{h}(t)$, and is defined as follows:

$$\left\{\delta R\left(\mathbf{h}^{0};\mathbf{F}^{0};\mathbf{v}^{(1)}\right)\right\}_{ind} \triangleq \int_{t_0}^{t_f}\left\{\frac{\partial D\left[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t\right]}{\partial\mathbf{h}}\mathbf{v}^{(1)}(t)\right\}_{\left(\mathbf{h}^{0};\mathbf{F}^{0}\right)} dt$$

$$\equiv \sum_{j=1}^{TH}\int_{t_0}^{t_f}\left\{\frac{\partial D\left[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t\right]}{\partial h_j}\delta h_j\right\}_{\left(\mathbf{h}^{0};\mathbf{F}^{0}\right)} dt. \tag{7}$$

The first-order relationship between the variations $\mathbf{v}^{(1)}(t)$ and $\delta\mathbf{F}$ is obtained from the first-order G-variation of Eq. (1) for $i = 1,...TH$, as follows:

$$\left\{\frac{d}{d\varepsilon}\left[h_i^0(t) + \varepsilon v_i^{(1)}(t)\right]\right\}_{\varepsilon=0} = \left\{\frac{d}{d\varepsilon}g_i\left(\mathbf{F}^0 + \varepsilon\delta\mathbf{F};t\right)\right\}_{\varepsilon=0} + \left\{\frac{d}{d\varepsilon}\varphi_i\left(\mathbf{F}^0 + \varepsilon\delta\mathbf{F};t\right)\right\}_{\varepsilon=0}$$

$$\times\left\{\frac{d}{d\varepsilon}\int_{t_0}^{t}\psi_i\left[h_1^0(\tau) + \varepsilon v_1^{(1)}(\tau),...,h_{TH}^0(\tau) + \varepsilon v_{TH}^{(1)}(\tau);\mathbf{F}^0 + \varepsilon\delta\mathbf{F};\tau\right]d\tau\right\}_{\varepsilon=0}; t_0 \le t \le t_f. \tag{8}$$

Performing the operations indicated in Eq. (8) yields the following NIE-V net, which will be called the "1st-Level Variational Sensitivity System" (1st-LVSS), for the components $v_i^{(1)}(t)$, $i = 1,...TH$, of the "1st-level variational function" $\mathbf{v}^{(1)}(t)$:

$$v_i^{(1)}(t) = \left\{\varphi_i(\mathbf{F};t)\sum_{j=1}^{TH}\int_{t_0}^{t}\frac{\partial\psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]}{\partial h_j(\tau)}v_j^{(1)}(\tau)d\tau\right\}_{\left(\mathbf{h}^0,\mathbf{F}^0\right)} + \sum_{k=1}^{TF}\left\{q_{ik}(\mathbf{F};t)\delta F_k\right\}_{\left(\mathbf{h}^0;\mathbf{F}^0\right)}, \tag{9}$$

where:

$$q_{ik}(\mathbf{F};t) \triangleq \frac{\partial g_i(\mathbf{F};t)}{\partial F_k} + \frac{\partial\varphi_i(\mathbf{F};t)}{\partial F_k}\int_{t_0}^{t}\psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]d\tau + \varphi_i(\mathbf{F};t)\int_{t_0}^{t}\frac{\partial\psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]}{\partial F_k}d\tau. \tag{10}$$

As indicated in Eq. (9), the 1st-LVSS is to be computed at the nominal/optimal values for the respective model parameters. It is important to note that the 1st-LVSS is linear in the variational function $\mathbf{v}^{(1)}(t)$, although it generally remains nonlinear in $\mathbf{h}(t)$.

The 1st-LVSS would need to be solved anew to obtain the function $\mathbf{v}^{(1)}(t)$ that would correspond to each variation δF_j, $j = 1,...,TF$; this procedure would become prohibitively expensive computationally if TF is a large number. The need for repeatedly solving the 1st-LVSS can be avoided by recasting the indirect-effect term $\left\{\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)}\right)\right\}_{ind}$ in terms of an expression that does not involve the function $\mathbf{v}^{(1)}(t)$. This goal can be achieved by expressing $\left\{\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)}\right)\right\}_{ind}$ in terms of another function, which will be called the "1st-level adjoint function," and which will be the solution of the "1st-Level Adjoint Sensitivity System (1st-LASS)" to be constructed next.

The 1st-LASS will be constructed in a Hilbert space, denoted as $H_1(\Omega_t)$, where $\Omega_t \triangleq \left[t_0, t_f\right]$, comprising elements of the same form as $\mathbf{v}^{(1)}(t) \in H_1(\Omega_t)$. The inner product of two elements

$\boldsymbol{\chi}^{(1)}(t) \triangleq \left[\chi_1^{(1)}(t), \ldots, \chi_{TH}^{(1)}(t)\right]^\dagger \in H_1(\Omega_t)$ and $\boldsymbol{\eta}^{(1)}(t) \triangleq \left[\eta_1^{(1)}(t), \ldots, \eta_{TH}^{(1)}(t)\right]^\dagger \in H_1(\Omega_t)$ will be denoted as $\left\langle \boldsymbol{\chi}^{(1)}(t), \boldsymbol{\eta}^{(1)}(t) \right\rangle_1$ and is defined as follows:

$$\left\langle \boldsymbol{\chi}^{(1)}(t), \boldsymbol{\eta}^{(1)}(t) \right\rangle_1 \triangleq \int_{t_0}^{t_f} \left[\boldsymbol{\chi}^{(1)}(t)\right]^\dagger \boldsymbol{\eta}^{(1)}(t) dt = \sum_{j=1}^{TH} \int_{t_0}^{t_f} \chi_i^{(1)}(t)\eta_i^{(1)}(t) dt. \qquad (11)$$

The inner product $\left\langle \boldsymbol{\chi}^{(1)}(t), \boldsymbol{\eta}^{(1)}(t) \right\rangle_1$ is required to hold in a neighborhood of the nominal values $\left(\mathbf{h}^0; \mathbf{F}^0\right)$.

The next step is to form the inner product of Eq. (9) with a vector $\mathbf{a}^{(1)}(t) \triangleq \left[a_1^{(1)}(t), \ldots, a_{TH}^{(1)}(t)\right]^\dagger \in H_1(\Omega_t)$, where the superscript "(1)" indicates "1st-level", to obtain the following relationship:

$$\sum_{i=1}^{TH} \int_{t_0}^{t_f} a_i^{(1)}(t)v_i^{(1)}(t) dt - \left\{ \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_i^{(1)}(t)\varphi_i(\mathbf{F};t) dt \sum_{j=1}^{TH} \int_{t_0}^{t} \frac{\partial \psi_i[\mathbf{h}(\tau);\mathbf{F};\tau]}{\partial h_j(\tau)} v_j^{(1)}(\tau) d\tau \right\}_{(\mathbf{h}^0, \mathbf{F}^0)}$$
$$= \sum_{i=1}^{TH} \sum_{k=1}^{TF} \left[\int_{t_0}^{t_f} a_i^{(1)}(t)q_{ik}(\mathbf{F};t) dt \right] \delta F_k. \qquad (12)$$

The second term on the left-side of Eq. (12) is transformed using "integration by parts" as follows:

$$\sum_{i=1}^{TH} \int_{t_0}^{t_f} a_i^{(1)}(t)\varphi_i(\mathbf{F};t) dt \sum_{j=1}^{TH} \int_{t_0}^{t} \frac{\partial \psi_i[\mathbf{h}(\tau);\mathbf{F};\tau]}{\partial h_j(\tau)} v_j^{(1)}(\tau) d\tau$$
$$= \sum_{i=1}^{TH} \sum_{j=1}^{TH} \left\{ \left[\int_{t_0}^{t} \frac{\partial \psi_i[\mathbf{h}(\tau);\mathbf{F};\tau]}{\partial h_j(\tau)} v_j^{(1)}(\tau) d\tau \right] \left[\int_{t_0}^{t} a_i^{(1)}(\tau)\varphi_i(\mathbf{F};\tau) d\tau \right] \right\}_{t=t_f}$$
$$- \sum_{i=1}^{TH} \sum_{j=1}^{TH} \left\{ \left[\int_{t_0}^{t} \frac{\partial \psi_i[\mathbf{h}(\tau);\mathbf{F};\tau]}{\partial h_j(\tau)} v_j^{(1)}(\tau) d\tau \right] \left[\int_{t_0}^{t} a_i^{(1)}(\tau)\varphi_i(\mathbf{F};\tau) d\tau \right] \right\}_{t=t_0} \qquad (13)$$
$$- \sum_{i=1}^{TH} \sum_{j=1}^{TH} \int_{t_0}^{t_f} \left[\frac{\partial \psi_i[\mathbf{h}(t);\mathbf{F};t]}{\partial h_j(t)} v_j^{(1)}(t) dt \right] \left[\int_{t_0}^{t} a_i^{(1)}(\tau)\varphi_i(\mathbf{F};\tau) d\tau \right]$$
$$= \sum_{i=1}^{TH} \sum_{j=1}^{TH} \int_{t_0}^{t_f} v_j^{(1)}(t) \frac{\partial \psi_i[\mathbf{h}(t);\mathbf{F};t]}{\partial h_j(t)} dt \int_{t}^{t_f} a_i^{(1)}(\tau)\varphi_i(\mathbf{F};\tau) d\tau.$$

Replacing the result obtained in Eq. (13) into the left-side of Eq. (12) yields the following relation for the left-side of Eq. (12):

$$\sum_{i=1}^{TH} \int_{t_0}^{t_f} a_i^{(1)}(t)v_i^{(1)}(t) dt - \left\{ \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_i^{(1)}(t)\varphi_i(\mathbf{F};t) dt \sum_{j=1}^{TH} \int_{t_0}^{t} \frac{\partial \psi_i[\mathbf{h}(\tau);\mathbf{F};\tau]}{\partial h_j(\tau)} v_j^{(1)}(\tau) d\tau \right\}_{(\mathbf{h}^0, \mathbf{F}^0)}$$
$$= \sum_{i=1}^{TH} \int_{t_0}^{t_f} v_i^{(1)}(t) \left\{ a_i^{(1)}(t) - \sum_{k=1}^{TH} \frac{\partial \psi_k[\mathbf{h}(t);\mathbf{F};t]}{\partial h_i(t)} \int_{t}^{t_f} a_k^{(1)}(\tau)\varphi_k(\mathbf{F};\tau) d\tau \right\}_{(\mathbf{h}^0, \mathbf{F}^0)} dt. \qquad (14)$$

The term on the right-side of Eq. (14) is now required to represent the "indirect-effect" term defined in Eq. (7), which is achieved by requiring that the components of the function $\mathbf{a}^{(1)}(t)$ satisfy the following system of equations for $i = 1, \ldots, TH$:

$$a_i^{(1)}(t) - \sum_{k=1}^{TH} \frac{\partial \psi_k[\mathbf{h}(t);\mathbf{F};t]}{\partial h_i(t)} \int_{t}^{t_f} a_k^{(1)}(\tau)\varphi_k(\mathbf{F};\tau) d\tau = \frac{\partial D[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t]}{\partial h_i}. \qquad (15)$$

The Volterra-like neural system obtained in Eq. (15) will be called the "1st-Level Adjoint Sensitivity System" and its solution, $\mathbf{a}^{(1)}(t)$, will be called the "1st-level adjoint sensitivity function." The 1st-LASS is to be solved using the nominal/optimal values for the parameters and for the function

$\mathbf{h}(t)$ but this fact has not been explicitly indicated in order to simplify the notation. The 1st-LASS is linear in $\mathbf{a}^{(1)}(t)$ but is, in general, nonlinear in $\mathbf{h}(t;\mathbf{x})$. Notably, the 1st-LASS is independent of any parameter variations and needs to be solved once only to determine the 1st-level adjoint sensitivity function $\mathbf{a}^{(1)}(t)$. The 1st-LASS is a "final-value problem" since the computation of the adjoint function $\mathbf{a}^{(1)}(t)$ will commence at $t = t_f$, with the known values $a_i^{(1)}(t_f) = \left\{ \partial D\left[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t\right] \big/ \partial h_i(t) \right\}_{t=t_f}$.

It follows from Eqs. (12)–(15) that the indirect-effect term defined in Eq. (7) can be expressed in terms of the 1st-level adjoint sensitivity function $\mathbf{a}^{(1)}(t)$ as follows:

$$\left\{ \delta R\left(\mathbf{h}^0;\mathbf{F}^0;\mathbf{a}^{(1)}\right) \right\}_{ind} = \sum_{k=1}^{TF} \left\{ \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_i^{(1)}(t) q_{ik}(\mathbf{F};t)\, dt \right\}_{\left(\mathbf{h}^0;\mathbf{F}^0\right)} \delta F_k. \tag{16}$$

Using the results obtained in Eqs. (16) and (6) in Eq. (5) yields the following expression for the G-variation $\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)};\delta\mathbf{F}\right)$, which is seen to be linear in $\delta\mathbf{F}$:

$$\begin{aligned}
\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)};\delta\mathbf{F}\right) &= \sum_{j=1}^{TF} \int_{t_0}^{t_f} \left\{ \frac{\partial D\left[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t\right]}{\partial F_j} \delta F_j \right\}_{\left(\mathbf{h}^0;\mathbf{F}^0\right)} dt \\
&\quad + \left\{ \sum_{i=1}^{TH} \sum_{j=1}^{TF} \left[\int_{t_0}^{t_f} a_i^{(1)}(t) q_{ij}(\mathbf{F};t)\, dt \right] \delta F_j \right\}_{\left(\mathbf{h}^0;\mathbf{F}^0\right)} \triangleq \sum_{j=1}^{TF} \left\{ \frac{\partial R}{\partial F_j} \right\}_{\left(\mathbf{h}^0;\mathbf{F}^0\right)} \delta F_j.
\end{aligned} \tag{17}$$

Identifying in Eq. (17) the expressions that multiply the variations δF_j, yields the following expressions for the sensitivities $\partial R/\partial F_j$ of the response $R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]$ with respect to the components $F_j(\boldsymbol{\theta})$ of the feature function $\mathbf{F}(\boldsymbol{\theta})$, for $j = 1,\dots,TF$:

$$\begin{aligned}
\frac{\partial R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]}{\partial F_j} &= \int_{t_0}^{t_f} \frac{\partial D\left[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t\right]}{\partial F_j} dt + \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_i^{(1)}(t) \frac{\partial g_i(\mathbf{F};t)}{\partial F_j} dt \\
&\quad + \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_i^{(1)}(t) \left\{ \frac{\partial \varphi_i(\mathbf{F};t)}{\partial F_j} \int_{t_0}^{t} \psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right] d\tau + \varphi_i(\mathbf{F};t) \int_{t_0}^{t} \frac{\partial \psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]}{\partial F_j} d\tau \right\} dt.
\end{aligned} \tag{18}$$

The expression on the right-side of Eq. (18) is to be evaluated at the nominal/optimal values for the respective model parameters, but this fact has not been indicated explicitly in order to simplify the notation.

The sensitivities with respect to the primary model parameters can be obtained by using the result obtained in Eq. (18) together with the "chain rule" of differentiating compound functions, as follows:

$$\frac{\partial R}{\partial \theta_j} = \sum_{i=1}^{TF} \frac{\partial R}{\partial F_i} \frac{\partial F_i}{\partial \theta_j}, \quad j = 1,\dots,TW. \tag{19}$$

The sensitivities $\partial R/\partial F_j$ are obtained from Eq. (18) while the derivatives $\partial F_i/\partial \theta_j$ are obtained analytically, exactly, from the known expressions of the feature functions $F_i(\boldsymbol{\theta})$.

Particular Case: The First-Order Comprehensive Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Volterra-Type (1st-CASAM-NIE-V)

When no feature functions can be constructed from the model parameters/weights, the feature functions become identical to the parameters, i.e., $F_i(\boldsymbol{\theta}) \equiv \theta_i$ for all $i = 1,\dots,TF \triangleq TW$. In this case, the expression obtained in Eq. (18) yields directly the first-order sensitivities $\partial R/\partial \theta_j$ of the decoder

response with respect to the model weights/parameters, for all $j = 1,...,TW$, taking on the following specific form:

$$\frac{\partial R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]}{\partial \theta_j} = \int_{t_0}^{t_f} \frac{\partial D\left[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t\right]}{\partial \theta_j} dt + \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_i^{(1)}(t) \frac{\partial g_i(\mathbf{F};t)}{\partial \theta_j} dt$$
$$+ \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_i^{(1)}(t) \left\{ \frac{\partial \varphi_i(\mathbf{F};t)}{\partial \theta_j} \int_{t_0}^{t} \psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right] d\tau + \varphi_i(\mathbf{F};t) \int_{t_0}^{t} \frac{\partial \psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]}{\partial \theta_j} d\tau \right\} dt. \tag{20}$$

Since the 1st-LASS is independent of any parameter variations, the 1st-level adjoint sensitivity function $\mathbf{a}^{(1)}(t) \triangleq \left[a_1^{(1)}(t),...,a_{TH}^{(1)}(t)\right]^{\dagger}$ which appears in Eq. (20) remains the solution of the 1st-LASS defined by Eq. (15). In this case, however, all of the sensitivities $\partial R/\partial \theta_j$, for all $j = 1,...,TW$ would be obtained by computing integrals using quadrature formulas. Thus, when there are no feature functions of parameters, the 1st-FASAM-NIE-V reduces to the "First-Order Comprehensive Adjoint Sensitivity Analysis Methodology [16] applied to Neural Integral Equations of Volterra-Type" (1st-CASAM-NIE-V). On the other hand, when features of parameters can be constructed, only TF $(TF < TW)$ numerical computations of integrals using quadrature formulas are required, using Eq. (18) to obtain the sensitivities $\partial R/\partial F_j$, $j = 1,...,TF$. Subsequently, the sensitivities with respect to the model's weights/parameters are obtained analytically using the chain-rule provided in Eq. (19).

3. Illustrative Application of the 1st-CASAM-NIE-V and 1st-FASAM-NIE-V Methodologies to Neutron Slowing Down in an Infinite Homogeneous Hydrogenous Medium

The illustrative model considered in this Section is a Volterra-type integral equation that describes the energy distribution of neutrons in a homogeneous hydrogenous medium (such as a water-moderated/cooled reactor system) containing ^{238}U (among other materials), which is a heavy element that strongly absorbs neutrons. The distribution of collided neutrons in such a medium is described [19–21] by the following linear integral equation of Volterra-type, customarily called the "neutron slowing down equation" for the neutron collision density denoted as $C(E)$:

$$C(E) = \frac{\Sigma_s(E_s)S}{\Sigma_t(E_s)E_s} + \int_E^{E_s} C(e) \frac{\Sigma_s(e)}{\Sigma_t(e)} \frac{de}{e}.$$

$$(21)$$

The various quantities that appear in Eq. (21) are defined as follows:

(i) The quantity S denotes the rate at which the source neutrons, considered to be monoenergetic, are emitted at the "source energy" E_s. Neutron upscattering is considered to be negligeable; therefore, E_s is the highest energy in the medium.

(ii) The quantity E, $0 < E_\ell \le E \le E_s$, denotes the instantaneous energy of the collided neutrons; E_ℓ denotes the lowest neutron energy in the model.

(iii) The quantity $\Sigma_s(E)$ denotes the medium's macroscopic scattering cross section, which is defined as follows:

$$\Sigma_s(E) \triangleq \sum_{i=1}^{M} w_i N_i \sigma_s^{(i)}(E), \tag{22}$$

where M denotes the number of materials in the medium, w_i denotes the relative weighting of the i-th-material in the medium, N_i denotes the number density of the i-th-material, while $\sigma_s^{(i)}(E)$ denotes the energy-dependent scattering microscopic cross section of the i-th-material.

(iv) The quantity $\Sigma_t(E)$ denotes the medium's macroscopic scattering cross section, which is defined as follows:

$$\Sigma_t(E) \triangleq \sum_{i=1}^{M} w_i N_i \sigma_t^{(i)}(E), \tag{23}$$

where $\sigma_t^{(i)}(E) \ge \sigma_s^{(i)}(E)$ denotes the energy-dependent total microscopic cross section of the ith-material. The quantities w_i, N_i, $\sigma_s^{(i)}(E)$, $\sigma_t^{(i)}(E)$ are subject to uncertainties since they are determined from experimentally obtained data.

Notably, the Volterra-type Eq. (21) is a "final-value problem" since the computation is started at the highest-energy value, $C(E_s) = \Sigma_s(E_s)S / \Sigma_t(E_s)E_s$, and progresses towards the lowest energy value E_ℓ. Customarily, the solution of Eq. (21) is written in the following form:

$$C(E) = \frac{\Sigma_s(E_s)S}{\Sigma_t(E_s)E} exp\left\{ -\int_E^{E_s} \frac{\Sigma_a(e)}{\Sigma_t(e)} \frac{de}{e} \right\}, \tag{24}$$

where $\Sigma_a(E) \triangleq \Sigma_t(E) - \Sigma_s(E)$ denotes the medium's macroscopic absorption cross section. The expression provided in Eq. (24) is amenable to computations of the loss of neutrons due to absorbing materials, particularly in the so-called "resonance" energy region.

A typical "decoder response" for the NIE-Volterra network modeled by Eq. (21) is the energy-averaged collision density, denoted below as $R[C(E)]$, which would be measured by a detector having an interaction cross-section Σ_d. Mathematically, this detector-response can be expressed as follows:

$$R[C(E)] \triangleq \int_{E_\ell}^{E_s} \Sigma_d(E)C(E)dE; \;\; \Sigma_d(E) \triangleq N_d\sigma_d(E), \tag{25}$$

where N_d and $\sigma_d(E)$ denote, respectively, the detector material's atomic number density and the microscopic cross section describing the interaction (e.g., absorption) of neutrons with the detector's material; N_d and $\sigma_d(E)$ can be considered as the "weights" that characterizes the neural net's "decoder."

Since the energy-dependence of the cross sections does not play a significant role in the sensitivity analysis of the NIE-Volterra modeled by Eq. (21), the respective microscopic cross-sections will henceforth be considered to be energy-independent for the purpose of illustrating the application of the 1st-FASAM-NIE-V, in order to simplify the ensuing derivations. For energy-independent cross sections, Eqs. (21) and (25) take on the following forms, respectively:

$$C(E) = F(\boldsymbol{\theta})\left[\frac{S}{E_s} + \int_E^{E_s} C(e) \frac{de}{e} \right], \tag{26}$$

$$R[C(E)] = \Sigma_d(\boldsymbol{\theta}) \int_{E_\ell}^{E_s} C(E)dE. \tag{27}$$

In Eqs. (26) and (27), the source strength S is an imprecisely-known "weight" that characterizes the neural net's "encoder." Furthermore, the (column) vector of parameters denoted as $\boldsymbol{\theta} \triangleq (\theta_1,...,\theta_{TW})^\dagger$ comprises as components the "imprecisely known primary model parameters" (or "weights", as customarily called when referring to neural nets) and is defined as follows:

$$\boldsymbol{\theta} \triangleq (\theta_1,...,\theta_{TW})^\dagger \triangleq \left(w_1,...,w_M; N_1,...,N_M; \sigma_t^{(1)},...,\sigma_t^{(M)}; \sigma_s^{(1)},...,\sigma_s^{(M)}; N_d,\sigma_d \right)^\dagger; \tag{28}$$

where $TW \triangleq 4 \times M + 2$ denotes the "total number of imprecisely-known weights/parameters." These primary model parameters/weights are not known exactly but are affected by uncertainties since they stem from experimental procedures, which determine the nominal/mean/optimal values and the second-order moments of their unknown joint distributions; their third- and higher-order moments

are rarely known. It is convenient to denote the nominal values of these primary model parameters/weights by using the superscript "zero" as follows:

$$\boldsymbol{\theta}^0 \triangleq \left(\theta_1^0,...,\theta_{TW}^0\right)^\dagger \triangleq \left(w_1^0,...,w_M^0;N_1^0,...,N_M^0;\sigma_t^{(1,0)},...,\sigma_t^{(M,0)};\sigma_s^{(1,0)},...,\sigma_s^{(M,0)};N_d^0,\sigma_d^0\right)^\dagger; \quad (29)$$

The "feature function of primary parameters," $F(\boldsymbol{\theta})$, is defined as follows:

$$F(\boldsymbol{\theta}) \triangleq \Sigma_s(\boldsymbol{\theta})\big/\Sigma_t(\boldsymbol{\theta}) \le 1. \tag{30}$$

The closed-form solution of Eq. (26) has the following expression in terms of the feature function $F(\boldsymbol{\theta})$:

$$C(E) = \frac{S}{E_s} F(\boldsymbol{\theta}) \left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}. \tag{31}$$

The closed-form expression of the decoder response can be readily obtained by replacing the result obtained in Eq. (31) into Eq. (27) and performing the integration over the energy-variable to obtain:

$$R\big[C(E)\big] = S\Sigma_d(\boldsymbol{\theta}) \frac{F(\boldsymbol{\theta})}{1-F(\boldsymbol{\theta})} \left[1 - \left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\right]. \tag{32}$$

The expression obtained in Eq. (32) reveals that the imprecisely known quantities that affect the decoder-response $R\big[C(E)\big]$ are as follows:

(i)   the source strength S;
(ii) the detector interaction macroscopic cross section $\Sigma_d(\boldsymbol{\theta})$, which can be considered to be a "feature function" of the model parameters $\boldsymbol{\theta}$;
(iii) the feature function $F(\boldsymbol{\theta}) \triangleq \Sigma_s(\boldsymbol{\theta})\big/\Sigma_t(\boldsymbol{\theta})$.

3.1. Application of 1ˢᵗ-CASAM-NIE-V to Directly Compute the First-Order Sensitivities of the Decoder Response with Respect to the Primary Model Parameters

The first-order sensitivities of the decoder response $R\big[C(E)\big]$ with respect to the model parameters is obtained by applying the definition of the G-differential to Eq. (26), for arbitrary parameter variations $\delta\boldsymbol{\theta} \triangleq \left(\delta\theta_1,...,\delta\theta_{TW}\right)^\dagger \triangleq \boldsymbol{\theta}-\boldsymbol{\theta}^0 \triangleq \left(\theta_1-\theta_1^0,...,\theta_{TW}-\theta_{TW}^0\right)^\dagger$ around the parameters' nominal values. These parameter variations will induce variations $\delta C(E) \triangleq C(E)-C^0(E)$ in the neutron collision density, around the nominal value $C^0(E)$ of the neutron collision density. The variations $\delta\boldsymbol{\theta}$ and $\delta C(E)$ will induce variations $\delta R\big[C^0(E);\boldsymbol{\theta}^0;\delta C(E);\delta\boldsymbol{\theta}\big]$ in the decoder's response.

The first-order Gateaux (G-)variation $\delta R\big[C^0(E);\boldsymbol{\theta}^0;\delta C(E);\delta\boldsymbol{\theta}\big]$ is obtained, by definition, from Eq. (27) as follows:

$$\delta R\big(C^0;\boldsymbol{\theta}^0;\delta C;\delta\boldsymbol{\theta}\big) \triangleq \frac{d}{d\varepsilon}\left\{\big[\Sigma_d\big(\boldsymbol{\theta}^0+\varepsilon\delta\boldsymbol{\theta}\big)\big]\int_{E_\ell}^{E_s}\big[C^0(E)+\varepsilon\delta C(E)\big]dE\right\}_{\varepsilon=0} \tag{33}$$

$$= \left\{\delta R\big(C^0;\boldsymbol{\theta}^0;\delta\boldsymbol{\theta}\big)\right\}_{dir} + \left\{\delta R\big(C^0;\boldsymbol{\theta}^0;\delta C\big)\right\}_{ind},$$

where the "direct effect" term $\left\{\delta R\big(C^0;\boldsymbol{\theta}^0;\delta\boldsymbol{\theta}\big)\right\}_{dir}$ arises directly from parameter variations $\delta\boldsymbol{\theta}$ and is defined as follows:

$$\left\{\delta R\left(C^0;\boldsymbol{\theta}^0;\delta\boldsymbol{\theta}\right)\right\}_{dir} \triangleq \left[\int_{E_\ell}^{E_s} C^0\left(E\right)dE\right]\sum_{i=1}^{TW}\left\{\frac{\partial\Sigma_d\left(\boldsymbol{\theta}\right)}{\partial\theta_i}\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}\delta\theta_i$$

$$= \left\{\left(\delta N_d\right)\sigma_d + N_d\left(\delta\sigma_d\right)\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}\int_{E_\ell}^{E_s} C^0\left(E\right)dE, \tag{34}$$

while the indirect effect term arises from the variations $\delta C\left(E\right)$ and is defined as follows:

$$\left\{\delta R\left(C^0;\boldsymbol{\theta}^0;\delta C\right)\right\}_{ind} \triangleq \Sigma_d\left(\boldsymbol{\theta}^0\right)\int_{E_\ell}^{E_s}\delta C\left(E\right)dE. \tag{35}$$

As indicated in Eqs. (34) and (35), both the direct-effect and the indirect-effect term are to be evaluated at the nominal parameter values.

The first-order relation between the variation $\delta C\left(E\right)$ and the parameter variations $\delta\theta_i$ is obtained by evaluating the G-variation of Eq. (26) for variations $\delta\boldsymbol{\theta}$ around the nominal parameter values $\boldsymbol{\theta}^0$ which yields, by definition, the following NIE-Volterra equation for $\delta C\left(E\right)$:

$$\delta C\left(E\right) \triangleq \frac{d}{d\varepsilon}\left\{F\left(\boldsymbol{\theta}^0 + \varepsilon\delta\boldsymbol{\theta}\right)\frac{S^0 + \varepsilon\delta S}{E_s} + F\left(\boldsymbol{\theta}^0 + \varepsilon\delta\boldsymbol{\theta}\right)\int_E^{E_s}\left[C^0\left(e\right) + \varepsilon\delta C\left(e\right)\right]\frac{de}{e}\right\}_{\varepsilon=0}$$

$$= F\left(\boldsymbol{\theta}^0\right)\int_E^{E_s}\delta C\left(e\right)\frac{de}{e} + Q\left(E\right), \tag{36}$$

where:

$$Q\left(E\right) \triangleq \frac{F\left(\boldsymbol{\theta}^0\right)\delta S}{E_s} + \left\{\left[\frac{S^0}{E_s} + \int_E^{E_s} C^0\left(e\right)\frac{de}{e}\right]\left[\sum_{i=1}^{TW}\frac{\partial F\left(\boldsymbol{\theta}\right)}{\partial\theta_i}\delta\theta_i\right]\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}$$

$$= \frac{F\left(\boldsymbol{\theta}^0\right)\delta S}{E_s} + \left\{\frac{S^0}{E_s}\left(\frac{E_s}{E}\right)^{F\left(\boldsymbol{\theta}\right)}\left[\sum_{i=1}^{TW}\frac{\partial F\left(\boldsymbol{\theta}\right)}{\partial\theta_i}\delta\theta_i\right]\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}. \tag{37}$$

The second equality in Eq. (37) has been obtained by using Eqs. (26) and (31) to eliminate the integral term involving $C^0\left(e\right)$.

The particular form of the first-order derivative $\partial F\left(\boldsymbol{\theta}\right)/\partial\theta_i$, which appears in Eq. (37), is obtained by using the definition of $F\left(\boldsymbol{\theta}\right)$ provided in Eq. (30), which yields the following expression:

$$\frac{\partial F\left(\boldsymbol{\theta}\right)}{\partial\theta_i} = \frac{1}{\Sigma_t\left(\boldsymbol{\theta}\right)}\frac{\partial\Sigma_s\left(\boldsymbol{\theta}\right)}{\partial\theta_i} - \frac{\Sigma_s\left(\boldsymbol{\theta}\right)}{\Sigma_t^2\left(\boldsymbol{\theta}\right)}\frac{\partial\Sigma_t\left(\boldsymbol{\theta}\right)}{\partial\theta_i}, \tag{38}$$

In view of the definition provided in Eq. (22), the derivatives $\partial\Sigma_s\left(\boldsymbol{\theta}\right)/\partial\theta_i$ have the following particular expressions:

$$\textit{for } i=1,\dots,M: \ \theta_i \triangleq w_i; \ \Rightarrow \frac{\partial\Sigma_s\left(\boldsymbol{\theta}\right)}{\partial\theta_i} = N_i\sigma_s^{(i)}; \tag{39}$$

$$\textit{for } i=M+1,\dots,2M: \ \theta_i \triangleq N_i; \ \Rightarrow \frac{\partial\Sigma_s\left(\boldsymbol{\theta}\right)}{\partial\theta_i} = w_i\sigma_s^{(i)}; \tag{40}$$

$$\textit{for } i=3M+1,\dots,4M: \ \theta_i \triangleq \sigma_s^{(i)}; \ \Rightarrow \frac{\partial\Sigma_s\left(\boldsymbol{\theta}\right)}{\partial\theta_i} = w_i N_i. \tag{41}$$

In view of the definition provided in Eq. (23), the derivatives $\partial\Sigma_t\left(\boldsymbol{\theta}\right)/\partial\theta_i$ have the following particular expressions:

$$\text{for } i = 1,...,M : \ \theta_i \triangleq w_i; \ \Rightarrow \frac{\partial \Sigma_t(\mathbf{\theta})}{\partial \theta_i} = N_i \sigma_t^{(i)}; \tag{42}$$

$$\text{for } i = M+1,...,2M : \ \theta_i \triangleq N_i; \ \Rightarrow \frac{\partial \Sigma_t(\mathbf{\theta})}{\partial \theta_i} = w_i \sigma_t^{(i)}; \tag{43}$$

$$\text{for } i = 2M+1,...,3M : \ \theta_i \triangleq \sigma_t^{(i)}; \ \Rightarrow \frac{\partial \Sigma_t(\mathbf{\theta})}{\partial \theta_i} = w_i N_i. \tag{44}$$

The NIE-Volterra net represented by Eq. (36) will be called the "1$^{\text{st}}$-Level Variational Sensitivity System (1$^{\text{st}}$-LVSS)" and its solution, $\delta C(E)$, will be called the "1$^{\text{st}}$-level variational sensitivity function." It is evident that Eq. (36) would need to be solved $(TW+1)$-times in order to obtain the variation $\delta C(E)$ for the source variation δS and for every parameter variation $\delta \theta_i$ This need for repeatedly solving Eq. (36) can be circumvented by applying the principles of the 1$^{\text{st}}$-CASAM-NIE-V, generally outlined in Section 2, to eliminate the appearance of the variation $\delta C(E)$ in the indirect-effect term defined in Eq. (35) while expressing this indirect-effect term as a functional of a first-level adjoint function that does not depend on any parameter variation, as follows:

1. Consider that the function $\delta C(E) \in \mathbb{H}_1(\Omega_E)$ belongs to a Hilbert space denoted as $\mathbb{H}_1(\Omega_E)$, which is defined on the domain $\Omega_E \triangleq [E_\ell, E_s]$. The inner product in $\mathbb{H}_1(\Omega_E)$ of two functions $u(E) \in \mathbb{H}_1(\Omega_E)$ and $v(E) \in \mathbb{H}_1(\Omega_E)$ will be denoted as $\langle u(E), v(E) \rangle_1$ and is defined as follows:

$$\langle u(E), v(E) \rangle_1 \triangleq \int_{E_\ell}^{E_0} u(E) v(E) dE. \tag{45}$$

2. Form the inner product of Eq. (36) with a vector $a^{(1)}(E) \in \mathbb{H}_1(E)$, where the superscript "(1)" indicates "1$^{\text{st}}$-Level", to obtain the following relationship:

$$\left\langle a^{(1)}(E), \delta C(E) - F(\mathbf{\theta}^0) \int_E^{E_s} \delta C(e) \frac{de}{e} \right\rangle_1 = \left\langle a^{(1)}(E), Q(E) \right\rangle_1. \tag{46}$$

3. Transform the left-side of Eq. (46) as follows:

$$\left\langle a^{(1)}(E), \delta C(E) - F(\mathbf{\theta}^0) \int_E^{E_s} \delta C(e) \frac{de}{e} \right\rangle_1 = \int_{E_\ell}^{E_s} a^{(1)}(E) \delta C(E) dE$$

$$-F(\mathbf{\theta}^0) \int_{E_\ell}^{E_s} a^{(1)}(E) dE \int_E^{E_s} \delta C(e) \frac{de}{e} = \int_{E_\ell}^{E_s} \delta C(E) a^{(1)}(E) dE \tag{47}$$

$$-F(\mathbf{\theta}^0) \int_{E_\ell}^{E_s} \delta C(E) \frac{dE}{E} \int_{E_\ell}^E a^{(1)}(e) de = \int_{E_\ell}^{E_s} \delta C(E) \left[a^{(1)}(E) - \frac{F(\mathbf{\theta}^0)}{E} \int_{E_\ell}^E a^{(1)}(e) de \right] dE.$$

In obtaining the expression on the right-side of the last equality in Eq. (47), the well-known "integration by parts" formula $\int g'(x) f(x) dx = g(x) f(x) - \int g(x) f'(x) dx$ has been used to reverse the integration order in the double integral $\int_{E_\ell}^{E_s} a^{(1)}(E) dE \int_E^{E_s} \delta C(e) \frac{de}{e}$, as follows:

$$g'(E) \triangleq v(E) \triangleq a^{(1)}(E); \ g(E) = \int_{E_\ell}^E v(x) dx = \int_{E_\ell}^E a^{(1)}(E) dE;$$

$$f(E) \triangleq \int_E^{E_s} u(x) dx; \ u(E) \triangleq \delta C(E)/E; \ f'(E) \triangleq -u(E); \tag{48}$$

$$\int_{E_\ell}^{E_s} v(E)\,dE \int_E^{E_s} u(x)\,dx = \left\{\int_{E_\ell}^E v(E)\,dE \int_E^{E_s} u(x)\,dx\right\}_{E=E_s} - \left\{\int_{E_\ell}^E v(E)\,dE \int_E^{E_s} u(x)\,dx\right\}_{E=E_\ell}$$

$$-\int_{E_\ell}^{E_s}\left[-u(E)\,dE\right]\int_{E_\ell}^E v(x)\,dx = \int_{E_\ell}^{E_s} u(E)\,dE \int_{E_\ell}^E v(x)\,dx. \tag{49}$$

4. Require the last term in Eq. (47) to represent the indirect-effect term defined in Eq. (35) which yields the following "1st-Level Adjoint Sensitivity System (1st-LASS)" for the first-level adjoint sensitivity function $a^{(1)}(E)$:

$$a^{(1)}(E) - \frac{F(\boldsymbol{\theta}^0)}{E}\int_{E_\ell}^E a^{(1)}(e)\,de = \Sigma_d(\boldsymbol{\theta}^0). \tag{50}$$

The 1st-LASS represented by Eq. (50) is a linear NIE-Volterra net, which is independent of any parameter variation and needs to be solved just once to obtain the first-level adjoint sensitivity function $a^{(1)}(E)$. Notably, the 1st-LASS is an "initial-value problem," in that the computation of $a^{(1)}(E)$ commences at the lowest-energy value, where $a^{(1)}(E_\ell) = \Sigma_d(\boldsymbol{\theta}^0)$, and progresses towards the highest-energy value, E_s. For further reference, the closed-form solution of Eq. (50) can be obtained by differentiating this equation with respect to E and subsequently integrating the resulting first-order linear differential equation, to obtain the following exact expression:

$$a^{(1)}(E) = \frac{\Sigma_d(\boldsymbol{\theta})}{1-F(\boldsymbol{\theta})}\left[1 - F(\boldsymbol{\theta})\left(\frac{E}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\right]. \tag{51}$$

The expression on the right-side of Eq. (51) is to be evaluated at the nominal parameter values $\boldsymbol{\theta}^0$, but the superscript "zero" has been omitted for notational simplicity.

5. Using Eqs. (46), (47) and (50) yields the following expression for the indirect-effect term defined in Eq. (35):

$$\{\delta R(C;\boldsymbol{\theta};\delta C)\}_{ind} = \int_{E_\ell}^{E_s} a^{(1)}(E)Q(E)\,dE = \frac{F(\boldsymbol{\theta})\delta S}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)\,dE$$

$$+ \frac{S}{E_s}\left[\sum_{i=1}^{TW}\frac{\partial F(\boldsymbol{\theta})}{\partial\theta_i}\delta\theta_i\right]\int_{E_\ell}^{E_s} a^{(1)}(E)\,dE\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}. \tag{52}$$

The expression on the right-side of Eq. (52) is to be evaluated at the nominal parameter values $\boldsymbol{\theta}^0$, but the superscript "zero" has been omitted for notational simplicity.

6. Adding the expression obtained in Eq. (52) to the expression of the direct-effect term represented by Eq. (34) yields the following expression for the first-order G-variation $\delta R(C^0;\boldsymbol{\theta}^0;\delta C;\delta\boldsymbol{\theta})$:

$$\delta R(C^0;\boldsymbol{\theta}^0;\delta C;\delta\boldsymbol{\theta}) = \left[(\delta N_d)\sigma_d + N_d(\delta\sigma_d)\right]\int_{E_\ell}^{E_s} C(E)\,dE$$

$$+ \frac{F(\boldsymbol{\theta})\delta S}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)\,dE + \frac{S}{E_s}\left[\sum_{i=1}^{TW}\frac{\partial F(\boldsymbol{\theta})}{\partial\theta_i}\delta\theta_i\right]\int_{E_\ell}^{E_s} a^{(1)}(E)\,dE\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}. \tag{53}$$

7. It follows from Eq. (53) that the first-order sensitivities of the decoder response with respect to the (encoder's) source strength and the optimal weights/parameters have the following expressions:

$$\frac{\partial R}{\partial S} = \frac{F(\boldsymbol{\theta})}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)\,dE\,; \tag{54}$$

$$\frac{\partial R}{\partial N_d} = \sigma_d \int_{E_\ell}^{E_s} C(E) dE \; ; \tag{55}$$

$$\frac{\partial R}{\partial \sigma_d} = N_d \int_{E_\ell}^{E_s} C(E) dE \; ; \tag{56}$$

$$\frac{\partial R}{\partial \theta_i} = \frac{S}{E_s} \frac{\partial F(\boldsymbol{\theta})}{\partial \theta_i} \int_{E_\ell}^{E_s} a^{(1)}(E) \left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})} dE; \; i=1,...,TW-2. \tag{57}$$

Inserting into Eqs. (54)–(57) the closed-form expression for the neutron collision density obtained in Eq. (31) yields the following closed-form explicit expressions for the first-order sensitivities of the decoder response with respect to the (encoder's) source strength and the optimal weights/parameters:

$$\frac{\partial R}{\partial S} = \Sigma_d(\boldsymbol{\theta}) \frac{F(\boldsymbol{\theta})}{1-F(\boldsymbol{\theta})} \left[1 - \left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\right]; \tag{58}$$

$$\frac{\partial R}{\partial N_d} = \sigma_d S \frac{F(\boldsymbol{\theta})}{1-F(\boldsymbol{\theta})} \left[1 - \left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\right]; \tag{59}$$

$$\frac{\partial R}{\partial \sigma_d} = N_d S \frac{F(\boldsymbol{\theta})}{1-F(\boldsymbol{\theta})} \left[1 - \left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\right]; \tag{60}$$

$$\frac{\partial R}{\partial \theta_i} = \frac{\partial F(\boldsymbol{\theta})}{\partial \theta_i} \frac{S\Sigma_d(\boldsymbol{\theta})}{1-F(\boldsymbol{\theta})} \left\{ \frac{1}{1-F(\boldsymbol{\theta})} \left[1 - \left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\right] - F(\boldsymbol{\theta}) \left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1} \ln\left(\frac{E_s}{E_\ell}\right) \right\}; \tag{61}$$
$$i=1,...,TW.$$

The correctness of the expressions obtained in Eqs. (58)–(61) can be readily verified by differentiating the expressions of the decoder's response obtained in Eq. (32).

In practice, only the exact mathematical expression of the 1st-LASS, namely Eq. (50), and the exact mathematical expression of the first-order sensitivities obtained in Eqs. (54)–(57) are available. The solution of the 1st-LASS, which is a linear NIE-Volterra net for the first-level adjoint sensitivity function $a^{(1)}(E)$, would need to be obtained numerically, in practice. The numerical solution for $a^{(1)}(E)$ would be used to determine the first-order sensitivities stemming from the "indirect-effect" term by using quadrature formulas to evaluate the integrals obtained in Eqs. (54) and (57). It is very important to note that a single "large-scale" computation, for determining numerically the adjoint function $a^{(1)}(E)$ by solving the 1st-LASS (a NIE-Volterra type equation), would be needed for evaluating all of the first-order sensitivities. The numerical computations using quadrature formulas for evaluating the integrals in Eqs. (54) and (57) are considered to be "small-scale" computations.

As has been already observed in the brief remarks following Eq. (37), the computation of the first-order sensitivities of the decoder response with respect to the encoder source strength S and model weights/parameters could also have been computed by numerically solving repeatedly the NIE-Volterra net (1st-LVSS) represented by Eq. (36). This procedure would be very expensive computationally, since it would require $(TW+1)$ large scale computations to solve the 1st-LVSS defined by Eq. (36) in order to obtain the variation $\delta C(E)$ for every parameter variation $\delta \theta_i$ and the source variation δS. In addition, the same amount of "quadrature" computations would need

to be performed using Eq. (35) as would be needed for evaluating the first-order sensitivities using Eqs. (54) and (57).

3.2. Efficient Indirect Computation Using the 1ˢᵗ-FASAM-NIE-V of the First-Order Sensitivities of the Decoder Response with Respect to Pimary Model Parameters

When feature functions of model parameters such as $\Sigma_d(\boldsymbol{\theta})$ and $F(\boldsymbol{\theta})$ can be identified, as is the case with the NIE-Volterra net and decoder response represented by Eqs. (26) and (27), respectively, it is considerably more efficient to determine the first-order sensitivities of the decoder response with respect to the feature functions and subsequently derive analytically the sensitivities with respect to the primary model parameters by using the "chain rule of differentiation," as will be shown in this Section. Thus, considering arbitrary variations $\delta\Sigma_d(\boldsymbol{\theta}) \triangleq \Sigma_d(\boldsymbol{\theta}) - \Sigma_d(\boldsymbol{\theta}^0)$ and $\delta F(\boldsymbol{\theta}) \triangleq F(\boldsymbol{\theta}) - F(\boldsymbol{\theta}^0)$ around the nominal values $\Sigma_d(\boldsymbol{\theta}^0)$ and, respectively, $F(\boldsymbol{\theta}^0)$, the first-order G-variation of the decoder response has the following expression:

$$
\begin{aligned}
\delta R\left(C^0; \boldsymbol{\theta}^0; \delta C; \delta\Sigma_d\right) &\triangleq \frac{d}{d\varepsilon}\left\{\left[\Sigma_d(\boldsymbol{\theta}^0) + \varepsilon\delta\Sigma_d(\boldsymbol{\theta}^0)\right]\int_{E_\ell}^{E_s}\left[C^0(E) + \varepsilon\delta C(E)\right]dE\right\}_{\varepsilon=0} \\
&= (\delta\Sigma_d)\left\{\int_{E_\ell}^{E_s} C(E)\,dE\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} + \left\{\delta R\left(C^0; \boldsymbol{\theta}^0; \delta C\right)\right\}_{ind},
\end{aligned}
\tag{62}
$$

where the expression of the indirect effect term is defined in Eq. (35). The first-order relation between the variation $\delta C(E)$ and the variations $\delta\Sigma_d(\boldsymbol{\theta})$ and $\delta F(\boldsymbol{\theta})$ is obtained, by definition, from Eq. (26) as follows:

$$
\begin{aligned}
\delta C(E) &\triangleq \frac{d}{d\varepsilon}\left\{\left[F(\boldsymbol{\theta}^0) + \varepsilon\delta F(\boldsymbol{\theta})\right]\frac{S^0 + \varepsilon\delta S}{E_s} + \left[F(\boldsymbol{\theta}^0) + \varepsilon\delta F(\boldsymbol{\theta})\right]\right. \\
&\left. \times \int_E^{E_s}\left[C^0(e) + \varepsilon\delta C(e)\right]\frac{de}{e}\right\}_{\varepsilon=0} = F(\boldsymbol{\theta}^0)\int_E^{E_s}\delta C(e)\frac{de}{e} + Q(E),
\end{aligned}
\tag{63}
$$

where:

$$
Q(E) \triangleq \frac{F(\boldsymbol{\theta}^0)\delta S}{E_s} + \left\{\frac{S^0}{E_s}\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}\left[\delta F(\boldsymbol{\theta})\right]\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}.
\tag{64}
$$

Comparing Eq. (63) to Eq. (36) indicates that the only difference between these equations is the expression of the term $Q(E)$, which is expressed in terms of $\delta F(\boldsymbol{\theta})$ in Eq. (64). Consequently, the first-level adjoint sensitivity function that corresponds to the variational function $\delta C(E)$ is determined by following the same procedure as outlined in Eqs. (46)–(50), ultimately obtaining the same 1ˢᵗ-LASS as was obtained in Eq. (50), having as solution the same expression for $a^{(1)}(E)$ as was obtained in Eq. (51). It further follows that the expression of the indirect-effect term will have the following expression:

$$
\left\{\delta R(C; \boldsymbol{\theta}; \delta C)\right\}_{ind} = (\delta S)\frac{F(\boldsymbol{\theta})}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)\,dE + \left[\delta F(\boldsymbol{\theta})\right]\frac{S}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)\,dE\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}. \tag{65}
$$

It follows from Eqs. (62) and (65) that the first-order G-variation $\delta R\left(C^0; \boldsymbol{\theta}^0; \delta C; \delta\Sigma_d\right)$ has the following expression:

$$
\begin{aligned}
\delta R\left(C^0; \boldsymbol{\theta}^0; \delta C; \delta\Sigma_d\right) &= (\delta\Sigma_d)\left\{\int_{E_\ell}^{E_s} C(E)\,dE\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} \\
&+ (\delta S)\frac{F(\boldsymbol{\theta})}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)\,dE + \left[\delta F(\boldsymbol{\theta})\right]\frac{S}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)\,dE\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}.
\end{aligned}
\tag{66}
$$

As indicated by the expression obtained in Eq. (66), the first-order sensitivities of the decoder response with respect to the feature functions and the encoder's source strength are as follows:

$$\frac{\partial R}{\partial F(\boldsymbol{\theta})} = \frac{S}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})} dE \; ; \tag{67}$$

$$\frac{\partial R}{\partial \Sigma_d(\boldsymbol{\theta})} = \int_{E_\ell}^{E_s} C(E) dE \; ; \tag{68}$$

$$\frac{\partial R}{\partial S} = \frac{F(\boldsymbol{\theta})}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E) dE \; . \tag{69}$$

The closed-form expressions of the above sensitivities are readily determined by using in Eqs. (67)–(69) the expressions obtained in Eqs. (51) and (24), and by performing the respective integrations obtain:

$$\frac{\partial R}{\partial F(\boldsymbol{\theta})} = \frac{S\Sigma_d(\boldsymbol{\theta})}{1-F(\boldsymbol{\theta})}\left\{\frac{1}{1-F(\boldsymbol{\theta})}\left[1-\left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\right] - F(\boldsymbol{\theta})\left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\ln\left(\frac{E_s}{E_\ell}\right)\right\}; \tag{70}$$

$$\frac{\partial R}{\partial \Sigma_d(\boldsymbol{\theta})} = S\frac{F(\boldsymbol{\theta})}{1-F(\boldsymbol{\theta})}\left[1-\left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\right]; \tag{71}$$

$$\frac{\partial R}{\partial S} = \Sigma_d(\boldsymbol{\theta})\frac{F(\boldsymbol{\theta})}{1-F(\boldsymbol{\theta})}\left[1-\left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\right]. \tag{72}$$

The first-order sensitivities with respect to the primary parameters are obtained analytically from Eqs. (67) and (68), respectively, by using the following "chain rule" of differentiation:

$$\frac{\partial R}{\partial N_d} = \frac{\partial R}{\partial \Sigma_d(\boldsymbol{\theta})}\frac{\partial \Sigma_d(\boldsymbol{\theta})}{\partial N_d} = \sigma_d\frac{\partial R}{\partial \Sigma_d(\boldsymbol{\theta})} \; ; \tag{73}$$

$$\frac{\partial R}{\partial \sigma_d} = \frac{\partial R}{\partial \Sigma_d(\boldsymbol{\theta})}\frac{\partial \Sigma_d(\boldsymbol{\theta})}{\partial \sigma_d} = N_d\frac{\partial R}{\partial \Sigma_d(\boldsymbol{\theta})} \; ; \tag{74}$$

$$\frac{\partial R}{\partial \theta_i} = \frac{\partial R}{\partial F(\boldsymbol{\theta})}\frac{\partial F(\boldsymbol{\theta})}{\partial \theta_i} \; ; \; i=1,...,4M. \tag{75}$$

The specific expressions of the first-order sensitivities $\partial R/\partial \theta_i$, $i=1,...,4M$, are obtained by using Eq. (75) in conjunction with Eq. (69) and Eqs. (38)–(44).

3.3. Discussion: Direct Versus Indirect Computation of the First-Order Sensitivities of Decoder Response with Respect to the Primary Model Parameters:

The principles of the 1st-CASAM-NIE-V were applied in Section 3.1 to determine the first-order sensitivities of the decoder response directly with respect to the model's primary parameters/weights. It has been shown that this procedure requires a single "large-scale" computation to solve a NIE-Volterra equation in order to determine the (single) 1st-level adjoint sensitivity function $a^{(1)}(E)$, which is subsequently used in $4M+1$ integrals that are computed using quadrature formulas. The

two additional first-order sensitivities with respect to the components of $\Sigma_d(\boldsymbol{\theta})$ require a single quadrature involving the forward function $C(E)$.

The principles of the 1st-FASAM-NIE-V were applied in Subsection 3.2 to determine the first-order sensitivities of the decoder response with respect to the feature functions. This path required just two (as opposed to $4M+1$) numerical evaluations of (two) integrals using quadrature formulas involving the 1st-level adjoint sensitivity function $a^{(1)}(E)$. The sensitivities of the decoder response with respect to the primary parameters/weights were subsequently determined analytically, using the "chain rule of differentiation" of the explicitly-known expression of the feature function $F(\boldsymbol{\theta})$. Evaluating the two additional first-order sensitivities with respect to the components of $\Sigma_d(\boldsymbol{\theta})$ require a single quadrature involving the forward function $C(E)$, as in Section 3.1. Evidently, the indirect path presented in Section 3.2 is computationally more efficient, since it requires substantially fewer numerical quadratures than the path presented in Section 3.1. The superiority of the indirect path, via "feature functions," over the direct computation of sensitivities with respect to the model parameters will be considerably more evident for the computation of second-order sensitivities, as will be shown in the forthcoming Sections 4 and 5, below.

Of course, when no feature functions can be identified, the 1st-FASAM-NIE-V methodology becomes identical to the 1st-CASAM-NIE-V methodology.

4. The Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Volterra-Type (2nd-FASAM-NIE-V)

The second-order sensitivities of the response $R\big[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\big]$ defined in Eq. (3) will be computed by conceptually using their basic definitions as being the "first-order sensitivities of the first-order sensitivities." Thus, the second-order sensitivities stemming from the first-order sensitivities $\partial R\big[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\big]/\partial F_j$ are obtained from the first-order G-differential of Eq. (18), for $j=1,...,TF$, as follows:

$$
\begin{aligned}
\delta\left(\frac{\partial R}{\partial F_j}\right) &\triangleq \left\{\frac{d}{d\varepsilon}\left[\int_{t_0}^{t_f}\frac{\partial D\big[\mathbf{h}^0(t)+\varepsilon\mathbf{v}^{(1)}(t);\mathbf{F}(\boldsymbol{\theta}^0)+\varepsilon\delta\mathbf{F};t\big]}{\partial F_j}dt\right]\right\}_{\varepsilon=0} \\
&+\left\{\sum_{i=1}^{TH}\frac{d}{d\varepsilon}\int_{t_0}^{t_f}\big[a_i^{(1,0)}(t)+\varepsilon\delta a_i^{(1)}(t)\big]\frac{\partial g_i\big(\mathbf{F}^0+\varepsilon\delta\mathbf{F};t\big)}{\partial F_j}dt\right\}_{\varepsilon=0} \\
&+\left\{\sum_{i=1}^{TH}\int_{t_0}^{t_f}\big[a_i^{(1,0)}(t)+\varepsilon\delta a_i^{(1)}(t)\big]\frac{\partial\varphi_i\big(\mathbf{F}^0+\varepsilon\delta\mathbf{F};t\big)}{\partial F_j}\right. \\
&\left.\times\int_{t_0}^{t}\psi_i\big[\mathbf{h}^0(\tau)+\varepsilon\mathbf{v}^{(1)}(\tau);\mathbf{F}^0+\varepsilon\delta\mathbf{F};\tau\big]d\tau\right\}_{\varepsilon=0} \\
&+\left\{\sum_{i=1}^{TH}\int_{t_0}^{t_f}\big[a_i^{(1,0)}(t)+\varepsilon\delta a_i^{(1)}(t)\big]\varphi_i\big(\mathbf{F}^0+\varepsilon\delta\mathbf{F};t\big)dt\right. \\
&\left.\times\int_{t_0}^{t}\frac{\partial\psi_i\big[\mathbf{h}^0(\tau)+\varepsilon\mathbf{v}^{(1)}(\tau);\mathbf{F}^0+\varepsilon\delta\mathbf{F};\tau\big]}{\partial F_j}d\tau\right\}_{\varepsilon=0} \\
&\triangleq \delta\big(\partial R/\partial F_j\big)_{dir}+\delta\big(\partial R/\partial F_j\big)_{ind};\quad j=1,...,TF.
\end{aligned}
\tag{76}
$$

In Eq. (76), the expression of the direct-effect term $\delta\big(\partial R/\partial F_j\big)_{dir}$ is obtained after performing the operations with respect to the scalar ε and comprises the variations $\delta\mathbf{F}$ (stemming from variations in the model parameters), being defined as follows:

$$
\delta\left(\frac{\partial R}{\partial F_j}\right)_{dir} \triangleq \sum_{k=1}^{TF}\int_{t_0}^{t_f}\left[\frac{\partial^2 D\left[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t\right]}{\partial F_k \partial F_j}\delta F_k\right]dt + \sum_{i=1}^{TH}\int_{t_0}^{t_f}\left[a_i^{(1)}(t)\sum_{k=1}^{TF}\frac{\partial^2 g_i(\mathbf{h};\mathbf{F})}{\partial F_k \partial F_j}\delta F_k\right]dt
$$

$$
+\sum_{i=1}^{TH}\int_{t_0}^{t_f}a_i^{(1)}(t)\left[\sum_{k=1}^{TF}\frac{\partial^2 \varphi_i(\mathbf{F};t)}{\partial F_k \partial F_j}\delta F_k\right]\int_{t_0}^{t}\psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]d\tau
$$

$$
+\sum_{i=1}^{TH}\int_{t_0}^{t_f}a_i^{(1)}(t)\frac{\partial \varphi_i(\mathbf{F};t)}{\partial F_j}\int_{t_0}^{t}\left[\sum_{k=1}^{TF}\frac{\partial \psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]}{\partial F_k}\delta F_k\right]d\tau \tag{77}
$$

$$
+\sum_{i=1}^{TH}\int_{t_0}^{t_f}a_i^{(1)}(t)\left[\sum_{k=1}^{TF}\frac{\partial \varphi_i(\mathbf{F};t)}{\partial F_k}\delta F_k\right]dt\int_{t_0}^{t}\frac{\partial \psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]}{\partial F_j}d\tau
$$

$$
+\sum_{i=1}^{TH}\int_{t_0}^{t_f}a_i^{(1)}(t)\varphi_i(\mathbf{F};t)dt\left[\sum_{k=1}^{TF}\frac{\partial^2 \psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]}{\partial F_k \partial F_j}\delta F_k\right]; \quad j=1,...,TF.
$$

The expression on the right-side of Eq. (77) is to be evaluated at the nominal/optimal values for the respective model parameters, but this fact has not been indicated explicitly in order to simplify the notation.

The expression of the indirect-effect term $\delta\left(\partial R/\partial F_j\right)_{ind}$ defined in Eq. (76) is obtained after performing the operations with respect to the scalar ε and comprises the variations $\mathbf{v}^{(1)}(t)$ and $\delta\mathbf{a}^{(1)}(t) \triangleq \left[\delta a_1^{(1)}(t),...,\delta a_{TH}^{(1)}(t)\right]^\dagger$, as follows:

$$
\delta\left(\frac{\partial R}{\partial F_j}\right)_{ind} \triangleq \sum_{k=1}^{TH}\int_{t_0}^{t_f}\frac{\partial^2 D\left[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t\right]}{\partial h_k(t)\partial F_j}v_k(t)dt + \sum_{i=1}^{TH}\int_{t_0}^{t_f}\delta a_i^{(1)}(t)\frac{\partial g_i(\mathbf{F};t)}{\partial F_j}dt
$$

$$
+\sum_{i=1}^{TH}\int_{t_0}^{t_f}\delta a_i^{(1)}(t)\frac{\partial \varphi_i(\mathbf{F};t)}{\partial F_j}\int_{t_0}^{t}\psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]d\tau
$$

$$
+\sum_{i=1}^{TH}\int_{t_0}^{t_f}a_i^{(1)}(t)\frac{\partial \varphi_i(\mathbf{F};t)}{\partial F_j}\int_{t_0}^{t}\sum_{k=1}^{TH}\frac{\partial \psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]}{\partial h_k(\tau)}v_k(\tau)d\tau \tag{78}
$$

$$
+\sum_{i=1}^{TH}\int_{t_0}^{t_f}\delta a_i^{(1)}(t)\varphi_i(\mathbf{F};t)dt\int_{t_0}^{t}\frac{\partial \psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]}{\partial F_j}d\tau
$$

$$
+\sum_{i=1}^{TH}\int_{t_0}^{t_f}a_i^{(1)}(t)\varphi_i(\mathbf{F};t)dt\int_{t_0}^{t}\sum_{k=1}^{TH}\left[\frac{\partial^2 \psi_i\left[\mathbf{h}(\tau);\mathbf{F};\tau\right]}{\partial h_k(\tau)\partial F_j}v_k(\tau)\right]d\tau; \quad j=1,...,TF.
$$

The expressions in Eq. (78) are to be evaluated at the nominal values of the respective functions and parameters, but the respective indication (i.e., the superscript "zero") has been omitted in order to simplify the notation.

The direct-effect term $\delta\left(\partial R/\partial F_j\right)_{dir}$ can be evaluated at this time for all variations $\delta\mathbf{F}$, but the indirect-effect term $\delta\left(\partial R/\partial F_j\right)_{ind}$ can be evaluated only after having determined the variations $\mathbf{v}^{(1)}(t)$ and $\delta\mathbf{a}^{(1)}(t)$. The variation $\mathbf{v}^{(1)}(t)$ is the solution of the 1st-LVSS defined by Eq. (9). On the other hand, the variational function $\delta\mathbf{a}^{(1)}(t)$ is the solution of the system of equations obtained by G-differentiating the 1st-LASS. By definition, the G-differential of Eq. (15) is obtained as follows, for $i=1,...,TH$:

$$\left\{ \frac{d}{d\varepsilon} \left[a_i^{(1,0)}(t) + \varepsilon \delta a_i^{(1)}(t) \right] \right\}_{\varepsilon=0} - \left\{ \frac{d}{d\varepsilon} \left[\sum_{k=1}^{TH} \frac{\partial \psi_k \left[\mathbf{h}^0(t) + \varepsilon \mathbf{v}^{(1)}(t); \mathbf{F}^0 + \varepsilon \delta \mathbf{F}; t \right]}{\partial h_i(t)} \right. \right.$$

$$\left. \left. \times \int_t^{t_f} \left[a_k^{(1,0)}(\tau) + \varepsilon \delta a_k^{(1)}(\tau) \right] \varphi_k \left(\mathbf{F}^0 + \varepsilon \delta \mathbf{F}; \tau \right) d\tau \right\}_{\varepsilon=0} \right. \tag{79}$$

$$= \left\{ \frac{d}{d\varepsilon} \frac{\partial D \left[\mathbf{h}^0(t) + \varepsilon \mathbf{v}^{(1)}(t); \mathbf{F}^0 + \varepsilon \delta \mathbf{F}; t \right]}{\partial h_i(t)} \right\}_{\varepsilon=0}.$$

Performing the operations indicated in Eq. (79) and rearranging the various terms yields the following relations, for $i = 1, ..., TH$:

$$\delta a_i^{(1)}(t) - \sum_{k=1}^{TH} \frac{\partial \psi_k \left[\mathbf{h}(t); \mathbf{F}; t \right]}{\partial h_i(t)} \int_t^{t_f} \delta a_k^{(1)}(\tau) \varphi_k(\mathbf{F}; \tau) d\tau - \sum_{m=1}^{TH} \frac{\partial^2 D \left[\mathbf{h}(t); \mathbf{F}; t \right]}{\partial h_m(t) \partial h_i(t)} v_m^{(1)}(t)$$

$$- \sum_{k=1}^{TH} \left[\sum_{m=1}^{TH} \frac{\partial^2 \psi_k \left[\mathbf{h}(t); \mathbf{F}; t \right]}{\partial h_m(t) \partial h_i(t)} v_m^{(1)}(t) \right] \int_t^{t_f} a_k^{(1)}(\tau) \varphi_k(\mathbf{F}; \tau) d\tau = \sum_{n=1}^{TF} S_{in}(\mathbf{F}; t) \delta F_n, \tag{80}$$

where:

$$\sum_{n=1}^{TF} S_{in}(\mathbf{F}; t) \delta F_n \triangleq \sum_{k=1}^{TH} \left[\sum_{n=1}^{TF} \frac{\partial^2 \psi_k \left[\mathbf{h}(t); \mathbf{F}; t \right]}{\partial F_n \partial h_i(t)} \delta F_n \right] \int_t^{t_f} a_k^{(1)}(\tau) \varphi_k(\mathbf{F}; \tau) d\tau$$

$$+ \sum_{k=1}^{TH} \frac{\partial \psi_k \left[\mathbf{h}(t); \mathbf{F}; t \right]}{\partial h_i(t)} \sum_{n=1}^{TF} \int_t^{t_f} \left[a_k^{(1)}(\tau) \frac{\partial \varphi_k(\mathbf{F}; \tau)}{\partial F_n} \delta F_n \right] d\tau + \sum_{n=1}^{TF} \frac{\partial^2 D \left[\mathbf{h}(t); \mathbf{F}; t \right]}{\partial F_n \partial h_i(t)} \delta F_n. \tag{81}$$

As indicated by the result obtained in Eq. (80), the variations $\delta \mathbf{a}^{(1)}(t)$ are coupled to the variations $\mathbf{v}^{(1)}(t)$. Therefore, they can be obtained by simultaneously solving Eqs. (80) and (9), which together will be called the "2nd-Level Variational Sensitivity System (2nd-LVSS)." The solution of the 2nd-LVSS, namely the vector $\mathbf{v}^{(2)}(t) \triangleq \left[\delta \mathbf{a}^{(1)}(t), \mathbf{v}^{(1)}(t) \right]^\dagger$, will be called the "2nd-level variational sensitivity function." Since the 2nd-LVSS depends on the variations $\delta \mathbf{F}$ (stemming from variations in the model parameters), it would need to be solved anew for each such variation. The repeated solving of the 2nd-LVSS can be avoided by following the general principles underlying the 2nd-FASAM [16], which considers the function $\mathbf{v}^{(2)}(t) \triangleq \left[\mathbf{v}^{(1)}(t), \delta \mathbf{a}^{(1)}(t) \right]^\dagger$ to be an element in a Hilbert space denoted as $\mathbb{H}_2(\Omega_t)$. The Hilbert space $\mathbb{H}_2(\Omega_t)$ is considered to be endowed with an inner product denoted as $\left\langle \boldsymbol{\chi}^{(2)}, \boldsymbol{\eta}^{(2)} \right\rangle_2$, between two vectors $\boldsymbol{\chi}^{(2)}(t) \triangleq \left[\boldsymbol{\chi}_1^{(2)}(t), \boldsymbol{\chi}_2^{(2)}(t) \right]^\dagger \in \mathbb{H}_2(\Omega_t)$ and $\boldsymbol{\eta}^{(2)}(t) = \left[\boldsymbol{\eta}_1^{(2)}(t), \boldsymbol{\eta}_2^{(2)}(t) \right]^\dagger \in \mathbb{H}_2(\Omega_t)$, with $\boldsymbol{\eta}_1^{(2)}(t) \triangleq \left[\eta_{1,1}^{(2)}(t), ..., \eta_{1,TH}^{(2)}(t) \right]^\dagger$, $\boldsymbol{\eta}_2^{(2)}(t) \triangleq \left[\eta_{2,1}^{(2)}(t), ..., \eta_{2,TH}^{(2)}(t) \right]^\dagger$, $\boldsymbol{\chi}_1^{(2)}(t) \triangleq \left[\chi_{1,1}^{(2)}(t), ..., \chi_{1,TH}^{(2)}(t) \right]^\dagger$, $\boldsymbol{\chi}_2^{(2)}(t) \triangleq \left[\chi_{2,1}^{(2)}(t), ..., \chi_{2,TH}^{(2)}(t) \right]^\dagger$, which is defined as follows:

$$\left\langle \boldsymbol{\chi}^{(2)}, \boldsymbol{\eta}^{(2)} \right\rangle_2 \triangleq \int_{t_0}^{t_f} \left[\boldsymbol{\chi}^{(2)}(t) \right]^\dagger \boldsymbol{\eta}^{(2)}(t) dt = \left\langle \boldsymbol{\chi}_1^{(2)}, \boldsymbol{\eta}_1^{(2)} \right\rangle_1 + \left\langle \boldsymbol{\chi}_2^{(2)}, \boldsymbol{\eta}_2^{(2)} \right\rangle_1$$

$$= \sum_{j=1}^{TH} \int_{t_0}^{t_f} \chi_{1,j}^{(2)}(t) \eta_{1,j}^{(2)}(t) dt + \sum_{j=1}^{TH} \int_{t_0}^{t_f} \chi_{2,j}^{(2)}(t) \eta_{2,j}^{(2)}(t) dt. \tag{82}$$

Following the general principles underlying the 2nd-FASAM [16], the function $\mathbf{v}^{(2)}(t) \triangleq \left[\mathbf{v}^{(1)}(t), \delta \mathbf{a}^{(1)}(t) \right]^\dagger$ will be eliminated from the expression of each indirect-effect terms $\delta \left(\partial R / \partial F_j \right)_{ind}$, $j = 1, ..., TF$, defined in Eq. (78). This elimination is achieved by considering, for each

index $j = 1,...,TF$, a vector-valued function denoted as $\mathbf{a}^{(2)}(t;j) = \left[\mathbf{a}_1^{(2)}(t;j), \mathbf{a}_2^{(2)}(t;j) \right]^{\dagger} \in \mathrm{H}_2(\Omega_t)$,

with $\mathbf{a}_1^{(2)}(t;j) \triangleq \left[a_{1,1}^{(2)}(t;j),...,a_{1,TH}^{(2)}(t;j) \right]^{\dagger}$ and $\mathbf{a}_2^{(2)}(t;j) \triangleq \left[a_{2,1}^{(2)}(t;j),...,a_{2,TH}^{(2)}(t;j) \right]^{\dagger}$. Using the

definition provided in Eq. (82), we construct the inner product of Eqs. (9) and (80) with the vector

$\mathbf{a}^{(2)}(t;j) = \left[\mathbf{a}_1^{(2)}(t;j), \mathbf{a}_2^{(2)}(t;j) \right]^{\dagger}$, to obtain the following relation:

$$
\begin{aligned}
&\sum_{i=1}^{TH} \int_{t_0}^{t_f} a_{1,i}^{(2)}(t;j) v_i^{(1)}(t) dt - \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_{1,i}^{(2)}(t;j) \varphi_i(\mathbf{F};t) dt \sum_{k=1}^{TH} \int_{t_0}^{t} \frac{\partial \psi_i \left[\mathbf{h}(\tau); \mathbf{F}; \tau \right]}{\partial h_k(\tau)} v_k^{(1)}(\tau) d\tau \\
&+ \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_{2,i}^{(2)}(t;j) \delta a_i^{(1)}(t) dt - \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_{2,i}^{(2)}(t;j) \sum_{k=1}^{TH} \frac{\partial \psi_k \left[\mathbf{h}(t); \mathbf{F}; t \right]}{\partial h_i(t)} dt \int_{t}^{t_f} \delta a_k^{(1)}(\tau) \varphi_k(\mathbf{F};\tau) d\tau \\
&- \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_{2,i}^{(2)}(t;j) \sum_{m=1}^{TH} \frac{\partial^2 D \left[\mathbf{h}(t); \mathbf{F}; t \right]}{\partial h_m(t) \partial h_i(t)} v_m^{(1)}(t) dt \\
&- \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_{2,i}^{(2)}(t;j) \sum_{k=1}^{TH} \left[\sum_{m=1}^{TH} \frac{\partial^2 \psi_k \left[\mathbf{h}(t); \mathbf{F}; t \right]}{\partial h_m(t) \partial h_i(t)} v_m^{(1)}(t) \right] dt \int_{t}^{t_f} a_k^{(1)}(\tau) \varphi_k(\mathbf{F};\tau) d\tau = Q^{(2)},
\end{aligned} \tag{83}
$$

where:

$$
Q^{(2)} \triangleq \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_{1,i}^{(2)}(t) dt \left[\sum_{k=1}^{TF} q_{ik}(\mathbf{F};t) \delta F_k \right] + \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_{2,i}^{(2)}(t) dt \left[\sum_{k=1}^{TF} S_{ik}(\mathbf{F};t) \delta F_k \right]. \tag{84}
$$

Following the principles of the 2nd-CASAM [16], the left-side of Eq. (83) will be identified with the indirect-effect term defined in Eq. (78), thereby determining the (yet undetermined) functions $\mathbf{a}^{(2)}(t;j) = \left[\mathbf{a}_1^{(2)}(t;j), \mathbf{a}_2^{(2)}(t;j) \right]^{\dagger}$. For this purpose, the right-side of Eq. (78) is cast in the form of the inner product $\left\langle \mathbf{v}^{(2)}(t), [\ ] \right\rangle_2 = \left\langle \mathbf{v}^{(1)}(t), [\ ] \right\rangle_1 + \left\langle \delta \mathbf{a}^{(1)}(t), [\ ] \right\rangle_1$. The terms on the right-side of Eq. (78) involving the components of the function $\delta \mathbf{a}^{(1)}(t)$ are already in the desired format, but the terms involving the components of the function $\mathbf{v}^{(1)}(t)$ must be re-arranged, as follows:

(i) The fourth term on the right-side of Eq. (78), is recast by using "integration by parts" as follows:

$$
\begin{aligned}
&\sum_{i=1}^{TH} \int_{t_0}^{t_f} a_i^{(1)}(t) \frac{\partial \varphi_i(\mathbf{F};t)}{\partial F_j} \int_{t_0}^{t} \sum_{k=1}^{TH} \frac{\partial \psi_i \left[\mathbf{h}(\tau); \mathbf{F}; \tau \right]}{\partial h_k(\tau)} v_k(\tau) d\tau \\
&= \sum_{i=1}^{TH} \sum_{k=1}^{TH} \left\{ \left[\int_{t_0}^{t} a_i^{(1)}(t) \frac{\partial \varphi_i(\mathbf{F};t)}{\partial F_j} dt \right] \left[\int_{t_0}^{t} \frac{\partial \psi_i \left[\mathbf{h}(\tau); \mathbf{F}; \tau \right]}{\partial h_k(\tau)} v_k(\tau) d\tau \right] \right\}_{t=t_f} \\
&- \sum_{i=1}^{TH} \sum_{k=1}^{TH} \left\{ \left[\int_{t_0}^{t} a_i^{(1)}(t) \frac{\partial \varphi_i(\mathbf{F};t)}{\partial F_j} dt \right] \left[\int_{t_0}^{t} \frac{\partial \psi_i \left[\mathbf{h}(\tau); \mathbf{F}; \tau \right]}{\partial h_k(\tau)} v_k(\tau) d\tau \right] \right\}_{t=t_0} \\
&- \sum_{i=1}^{TH} \sum_{k=1}^{TH} \int_{t_0}^{t_f} \frac{\partial \psi_i \left[\mathbf{h}(t); \mathbf{F}; t \right]}{\partial h_k(t)} v_k(t) \int_{t_0}^{t} a_i^{(1)}(\tau) \frac{\partial \varphi_i(\mathbf{F};\tau)}{\partial F_j} d\tau \\
&= \sum_{i=1}^{TH} \sum_{k=1}^{TH} \int_{t_0}^{t_f} \frac{\partial \psi_i \left[\mathbf{h}(t); \mathbf{F}; t \right]}{\partial h_k(t)} v_k(t) dt \int_{t}^{t_f} a_i^{(1)}(\tau) \frac{\partial \varphi_i(\mathbf{F};\tau)}{\partial F_j} d\tau.
\end{aligned} \tag{85}
$$

(ii) The sixth (last) term on the right-side of Eq. (78), is recast by using "integration by parts," as above, to obtain the following relation:

$$
\begin{aligned}
&\sum_{i=1}^{TH} \int_{t_0}^{t_f} a_i^{(1)}(t) \varphi_i(\mathbf{F};t) dt \int_{t_0}^{t} \sum_{k=1}^{TH} \left[\frac{\partial^2 \psi_i \left[\mathbf{h}(\tau); \mathbf{F}; \tau \right]}{\partial h_k(\tau) \partial F_j} v_k(\tau) \right] d\tau \\
&= \sum_{i=1}^{TH} \sum_{k=1}^{TH} \int_{t_0}^{t_f} \frac{\partial^2 \psi_i \left[\mathbf{h}(t); \mathbf{F}; t \right]}{\partial h_k(t) \partial F_j} v_k(t) dt \int_{t}^{t_f} a_i^{(1)}(\tau) \varphi_i(\mathbf{F};\tau) d\tau.
\end{aligned} \tag{86}
$$

Using in Eq. (78) the results obtained in Eqs. (85) and (86) yields the following expression for the indirect-effect term, for $j = 1,\dots,TF$:

$$\delta\left(\frac{\partial R}{\partial F_j}\right)_{ind} \triangleq \sum_{k=1}^{TH}\int_{t_0}^{t_f} v_k(t)\,dt\left\{\frac{\partial^2 D[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t]}{\partial h_k(t)\partial F_j}\right.$$

$$+\sum_{i=1}^{TH}\frac{\partial\psi_i[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)}\int_t^{t_f} a_i^{(1)}(\tau)\frac{\partial\varphi_i(\mathbf{F};\tau)}{\partial F_j}\,d\tau + \sum_{i=1}^{TH}\frac{\partial^2\psi_i[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)\partial F_j}\int_t^{t_f} a_i^{(1)}(\tau)\varphi_i(\mathbf{F};\tau)\,d\tau\right\} \quad (87)$$

$$+\sum_{k=1}^{TH}\int_{t_0}^{t_f}\delta a_k^{(1)}(t)\left\{\frac{\partial g_k(\mathbf{F};t)}{\partial F_j}+\frac{\partial\varphi_k(\mathbf{F};t)}{\partial F_j}\int_{t_0}^t\psi_k[\mathbf{h}(\tau);\mathbf{F};\tau]\,d\tau+\varphi_k(\mathbf{F};t)\int_{t_0}^t\frac{\partial\psi_k}{\partial F_j}\,d\tau\right\}dt.$$

The left-side of Eq. (83) is now recast in the form of the inner product $\left\langle\mathbf{v}^{(2)}(t),[\ ]\right\rangle_2$ by performing the following operations:

(i) The second term on the left-side of Eq. (83) is rearranged by using "integration by parts" as follows:

$$\sum_{i=1}^{TH}\int_{t_0}^{t_f} a_{1,i}^{(2)}(t;j)\varphi_i(\mathbf{F};t)\,dt\sum_{k=1}^{TH}\int_{t_0}^t\frac{\partial\psi_i[\mathbf{h}(\tau);\mathbf{F};\tau]}{\partial h_k(\tau)}v_k^{(1)}(\tau)\,d\tau$$

$$=\sum_{k=1}^{TH}\sum_{i=1}^{TH}\int_{t_0}^{t_f} v_k(t)\frac{\partial\psi_i[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)}\,dt\int_t^{t_f} a_{1,i}^{(2)}(\tau;j)\varphi_i(\mathbf{F};\tau)\,d\tau. \quad (88)$$

(ii) The fourth term on the left-side of Eq. (83) is rearranged by using "integration by parts" as follows:

$$\sum_{i=1}^{TH}\int_{t_0}^{t_f} a_{2,i}^{(2)}(t;j)\sum_{k=1}^{TH}\frac{\partial\psi_k[\mathbf{h}(t);\mathbf{F};t]}{\partial h_i(t)}\,dt\int_t^{t_f}\delta a_k^{(1)}(\tau)\varphi_k(\mathbf{F};\tau)\,d\tau$$

$$=\sum_{i=1}^{TH}\sum_{k=1}^{TH}\left\{\left[\int_{t_0}^t a_{2,i}^{(2)}(t;j)\sum_{k=1}^{TH}\frac{\partial\psi_k[\mathbf{h}(t);\mathbf{F};t]}{\partial h_i(t)}\,dt\right]\left[\int_t^{t_f}\delta a_k^{(1)}(\tau)\varphi_k(\mathbf{F};\tau)\,d\tau\right]\right\}_{t=t_f}$$

$$-\sum_{i=1}^{TH}\sum_{k=1}^{TH}\left\{\left[\int_{t_0}^t a_{2,i}^{(2)}(t;j)\sum_{k=1}^{TH}\frac{\partial\psi_k[\mathbf{h}(t);\mathbf{F};t]}{\partial h_i(t)}\,dt\right]\left[\int_t^{t_f}\delta a_k^{(1)}(\tau)\varphi_k(\mathbf{F};\tau)\,d\tau\right]\right\}_{t=t_0} \quad (89)$$

$$-\sum_{i=1}^{TH}\sum_{k=1}^{TH}\int_{t_0}^{t_f} dt\left[-\delta a_k^{(1)}(t)\varphi_k(\mathbf{F};t)\right]\left[\int_{t_0}^t a_{2,i}^{(2)}(\tau;j)\frac{\partial\psi_k[\mathbf{h}(\tau);\mathbf{F};\tau]}{\partial h_i(\tau)}\,d\tau\right]$$

$$=\sum_{k=1}^{TH}\sum_{i=1}^{TH}\int_{t_0}^{t_f}\delta a_k^{(1)}(t)\varphi_k(\mathbf{F};t)\,dt\int_{t_0}^t a_{2,i}^{(2)}(\tau;j)\frac{\partial\psi_k[\mathbf{h}(\tau);\mathbf{F};\tau]}{\partial h_i(\tau)}\,d\tau$$

(iii) The fifth term on the left-side of Eq. (83) is rearranged as follows:

$$\sum_{i=1}^{TH}\int_{t_0}^{t_f} a_{2,i}^{(2)}(t;j)\sum_{k=1}^{TH}\frac{\partial^2 D[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)\partial h_i(t)}v_k^{(1)}(t)\,dt$$

$$=\sum_{k=1}^{TH}\int_{t_0}^{t_f} v_k^{(1)}(t)\left[\sum_{i=1}^{TH}\frac{\partial^2 D[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)\partial h_i(t)}a_{2,i}^{(2)}(t;j)\right]dt. \quad (90)$$

(iv) The sixth term on the left-side of Eq. (83) is rearranged as follows:

$$\sum_{i=1}^{TH}\int_{t_0}^{t_f} a_{2,i}^{(2)}(t)\sum_{k=1}^{TH}\left[\sum_{m=1}^{TH}\frac{\partial^2\psi_k[\mathbf{h}(t);\mathbf{F};t]}{\partial h_m(t)\partial h_i(t)}v_m^{(1)}(t)\right]dt\int_t^{t_f} a_k^{(1)}(\tau)\varphi_k(\mathbf{F};\tau)\,d\tau$$

$$=\sum_{m=1}^{TH}\sum_{i=1}^{TH}\sum_{k=1}^{TH}\int_{t_0}^{t_f} v_m^{(1)}(t)\left[\frac{\partial^2\psi_k[\mathbf{h}(t);\mathbf{F};t]}{\partial h_m(t)\partial h_i(t)}a_{2,i}^{(2)}(t)\int_t^{t_f} a_k^{(1)}(\tau)\varphi_k(\mathbf{F};\tau)\,d\tau\right]dt \quad (91)$$

$$=\sum_{k=1}^{TH}\int_{t_0}^{t_f} v_k^{(1)}(t)\sum_{i=1}^{TH}\sum_{n=1}^{TH}\left[\frac{\partial^2\psi_n[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)\partial h_i(t)}a_{2,i}^{(2)}(t)\int_t^{t_f} a_n^{(1)}(\tau)\varphi_n(\mathbf{F};\tau)\,d\tau\right]dt$$

Inserting the results obtained in Eqs. (88)–(91) into the left-side of Eq. (83) yields the following relation:

$$
\begin{aligned}
Q^{(2)} &= \sum_{k=1}^{TH} \int_{t_0}^{t_f} dt\, v_k^{(1)}(t) \Bigg\{ a_{1,k}^{(2)}(t;j) - \sum_{i=1}^{TH} \frac{\partial \psi_i[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)} \int_t^{t_f} a_{1,i}^{(2)}(\tau;j)\varphi_i(\mathbf{F};\tau)\,d\tau \\
&- \sum_{i=1}^{TH} \frac{\partial^2 D[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)\partial h_i(t)} a_{2,i}^{(2)}(t;j) - \sum_{i=1}^{TH}\sum_{n=1}^{TH} \frac{\partial^2 \psi_n[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)\partial h_i(t)} a_{2,i}^{(2)}(t)\int_t^{t_f} a_n^{(1)}(\tau)\varphi_n(\mathbf{F};\tau)\,d\tau \Bigg\} \quad (92) \\
&+ \sum_{k=1}^{TH} \int_{t_0}^{t_f} dt\, \delta a_k^{(1)}(t) \Bigg\{ a_{2,k}^{(2)}(t;j) - \sum_{i=1}^{TH} \varphi_k(\mathbf{F};t)\, dt \int_{t_0}^{t} a_{2,i}^{(2)}(\tau;j)\frac{\partial \psi_k[\mathbf{h}(\tau);\mathbf{F};\tau]}{\partial h_i(\tau)}\,d\tau \Bigg\}.
\end{aligned}
$$

The right-side of Eq. (92) can now be required to represent the indirect-effect term defined in Eq. (87), by imposing the requirement that the hitherto arbitrary function $\mathbf{a}^{(2)}(t;j) = \left[\mathbf{a}_1^{(2)}(t;j), \mathbf{a}_2^{(2)}(t;j)\right]^\dagger$ be the solution of the following NIE-Volterra equations, for $i=1,...,TH; \; j=1,...,TF$:

$$
\begin{aligned}
& a_{1,k}^{(2)}(t;j) - \sum_{i=1}^{TH} \frac{\partial \psi_i[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)} \int_t^{t_f} a_{1,i}^{(2)}(\tau;j)\varphi_i(\mathbf{F};\tau)\,d\tau - \sum_{i=1}^{TH} \frac{\partial^2 D[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)\partial h_i(t)} a_{2,i}^{(2)}(t;j) \\
& - \sum_{i=1}^{TH}\sum_{n=1}^{TH} \frac{\partial^2 \psi_n[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)\partial h_i(t)} a_{2,i}^{(2)}(t)\int_t^{t_f} a_n^{(1)}(\tau)\varphi_n(\mathbf{F};\tau)\,d\tau = \frac{\partial^2 D[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t]}{\partial h_k(t)\partial F_j} \quad (93) \\
& + \sum_{i=1}^{TH} \frac{\partial \psi_i[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)} \int_t^{t_f} a_i^{(1)}(\tau)\frac{\partial \varphi_i(\mathbf{F};\tau)}{\partial F_j}\,d\tau + \sum_{i=1}^{TH} \frac{\partial^2 \psi_i[\mathbf{h}(t);\mathbf{F};t]}{\partial h_k(t)\partial F_j} \int_t^{t_f} a_i^{(1)}(\tau)\varphi_i(\mathbf{F};\tau)\,d\tau ;
\end{aligned}
$$

$$
\begin{aligned}
& a_{2,k}^{(2)}(t;j) - \sum_{i=1}^{TH} \varphi_k(\mathbf{F};t)\, dt \int_{t_0}^{t} a_{2,i}^{(2)}(\tau;j)\frac{\partial \psi_k[\mathbf{h}(\tau);\mathbf{F};\tau]}{\partial h_i(\tau)}\,d\tau \\
& = \frac{\partial g_k(\mathbf{F};t)}{\partial F_j} + \frac{\partial \varphi_k(\mathbf{F};t)}{\partial F_j} \int_{t_0}^{t} \psi_k[\mathbf{h}(\tau);\mathbf{F};\tau]\,d\tau + \varphi_k(\mathbf{F};t)\int_{t_0}^{t} \frac{\partial \psi_k}{\partial F_j}\,d\tau .
\end{aligned}
\quad (94)
$$

It follows from Eqs. (92)–(94) that the indirect-effect term $\delta\left(\partial R/\partial F_j\right)_{ind}$ defined by Eq. (78) or, equivalently, Eq. (87) can be expressed in terms of the function $\mathbf{a}^{(2)}(t;j) = \left[\mathbf{a}_1^{(2)}(t;j), \mathbf{a}_2^{(2)}(t;j)\right]^\dagger$ as follows, for $j=1,...,TF$:

$$
\begin{aligned}
\delta\left(\frac{\partial R}{\partial F_j}\right)_{ind} &= \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_{1,i}^{(2)}(t;j)\,dt \left[\sum_{n=1}^{TF} q_{in}(\mathbf{F};t)\delta F_n\right] + \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_{2,i}^{(2)}(t;j)\,dt \left[\sum_{n=1}^{TF} S_{in}(\mathbf{F};t)\delta F_n\right] \\
&= \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_{1,i}^{(2)}(t;j)\,dt \sum_{n=1}^{TF} \delta F_n \Bigg\{ \frac{\partial g_i(\mathbf{F};t)}{\partial F_n} + \frac{\partial \varphi_i(\mathbf{F};t)}{\partial F_n} \int_{t_0}^{t} \psi_i[\mathbf{h}(\tau);\mathbf{F};\tau]\,d\tau \\
&+ \varphi_i(\mathbf{F};t)\int_{t_0}^{t} \frac{\partial \psi_i[\mathbf{h}(\tau);\mathbf{F};\tau]}{\partial F_n}\,d\tau \Bigg\} + \sum_{i=1}^{TH} \int_{t_0}^{t_f} a_{2,i}^{(2)}(t;j)\,dt \Bigg\{ \sum_{n=1}^{TF} \frac{\partial^2 D[\mathbf{h}(t);\mathbf{F};t]}{\partial F_n\partial h_i(t)}\delta F_n \quad (95) \\
&+ \sum_{k=1}^{TH} \left[\sum_{n=1}^{TF} \frac{\partial^2 \psi_k[\mathbf{h}(t);\mathbf{F};t]}{\partial F_n\partial h_i(t)}\delta F_n\right] \int_t^{t_f} a_k^{(1)}(\tau)\varphi_k(\mathbf{F};\tau)\,d\tau \\
&+ \sum_{k=1}^{TH} \frac{\partial \psi_k[\mathbf{h}(t);\mathbf{F};t]}{\partial h_i(t)} \sum_{n=1}^{TF} \int_t^{t_f} \left[a_k^{(1)}(\tau)\frac{\partial \varphi_k(\mathbf{F};\tau)}{\partial F_n}\delta F_n\right] d\tau \Bigg\}.
\end{aligned}
$$

The second-order sensitivities $\partial^2 R[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})]/\partial F_j\partial F_n$ of the decoder-response with respect to the components of the feature function are obtained by adding the expression of the indirect-effect term obtained in Eq. (95) to the expression for the direct-effect term obtained in Eq. (77) and subsequently identifying the expressions that multiply the variations δF_n. The expressions thus obtained for $\partial^2 R[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})]/\partial F_j\partial F_n$, for $j,n=1,...,TF$, are as follows:

$$\frac{\partial^2 R\big[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\big]}{\partial F_j \partial F_n} \triangleq \frac{\partial^2 D\big[\mathbf{h}(t);\mathbf{F}(\boldsymbol{\theta});t\big]}{\partial F_n \partial F_j} + \sum_{i=1}^{TH}\int_{t_0}^{t_f} a_i^{(1)}(t)\frac{\partial^2 g_i(\mathbf{h};\mathbf{F})}{\partial F_n \partial F_j}dt$$

$$+\sum_{i=1}^{TH}\int_{t_0}^{t_f} a_i^{(1)}(t)\frac{\partial^2 \varphi_i(\mathbf{F};t)}{\partial F_n \partial F_j}\int_{t_0}^{t}\psi_i\big[\mathbf{h}(\tau);\mathbf{F};\tau\big]d\tau + \sum_{i=1}^{TH}\int_{t_0}^{t_f} a_i^{(1)}(t)\frac{\partial \varphi_i(\mathbf{F};t)}{\partial F_j}\int_{t_0}^{t}\frac{\partial \psi_i\big[\mathbf{h}(\tau);\mathbf{F};\tau\big]}{\partial F_n}d\tau$$

$$+\sum_{i=1}^{TH}\int_{t_0}^{t_f} a_i^{(1)}(t)\frac{\partial \varphi_i(\mathbf{F};t)}{\partial F_n}dt\int_{t_0}^{t}\frac{\partial \psi_i\big[\mathbf{h}(\tau);\mathbf{F};\tau\big]}{\partial F_j}d\tau + \sum_{i=1}^{TH}\int_{t_0}^{t_f} a_i^{(1)}(t)\varphi_i(\mathbf{F};t)dt\frac{\partial^2 \psi_i\big[\mathbf{h}(\tau);\mathbf{F};\tau\big]}{\partial F_n \partial F_j}$$

$$+\sum_{i=1}^{TH}\int_{t_0}^{t_f} a_{1,i}^{(2)}(t;j)dt\left\{\frac{\partial g_i(\mathbf{F};t)}{\partial F_n} + \frac{\partial \varphi_i(\mathbf{F};t)}{\partial F_n}\int_{t_0}^{t}\psi_i\big[\mathbf{h}(\tau);\mathbf{F};\tau\big]d\tau\right.$$

$$+\varphi_i(\mathbf{F};t)\int_{t_0}^{t}\frac{\partial \psi_i\big[\mathbf{h}(\tau);\mathbf{F};\tau\big]}{\partial F_n}d\tau\right\} + \sum_{i=1}^{TH}\int_{t_0}^{t_f} a_{2,i}^{(2)}(t;j)dt\left\{\frac{\partial^2 D\big[\mathbf{h}(t);\mathbf{F};t\big]}{\partial F_n \partial h_i(t)}\right.$$

$$+\sum_{k=1}^{TH}\frac{\partial^2 \psi_k\big[\mathbf{h}(t);\mathbf{F};t\big]}{\partial F_n \partial h_i(t)}\int_{t}^{t_f} a_k^{(1)}(\tau)\varphi_k(\mathbf{F};\tau)d\tau + \sum_{k=1}^{TH}\frac{\partial \psi_k\big[\mathbf{h}(t);\mathbf{F};t\big]}{\partial h_i(t)}\int_{t}^{t_f} a_k^{(1)}(\tau)\frac{\partial \varphi_k(\mathbf{F};\tau)}{\partial F_n}d\tau\right\}.$$

$$(96)$$

The NIE-Volterra system presented in Eqs. (93) and (94) is called the "2nd-Level Adjoint Sensitivity System (2nd-LASS)" and its solution, $\mathbf{a}^{(2)}(j;t) = \big[\mathbf{a}_1^{(2)}(j;t),\mathbf{a}_2^{(2)}(j;t)\big]^{\dagger}$, $j = 1,...,TF$, is called the "2nd-level adjoint sensitivity function." Since the sources on the right-sides of Eqs. (93) and (94) stem from the first-order sensitivities $\partial R\big[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\big]\big/\partial F_j$, $j = 1,...,TF$, they are dependent on the index "j", which implies that for each first-order sensitivity $\partial R\big[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\big]\big/\partial F_j$, there will correspond a distinct 2nd-LASS, having a distinct solution $\mathbf{a}^{(2)}(j;t) = \big[\mathbf{a}_1^{(2)}(j;t),\mathbf{a}_2^{(2)}(j;t)\big]^{\dagger}$, a fact that has been emphasized by using the index "j" in the list of arguments of this 2nd-level adjoint sensitivity function. Therefore, there will be as many 2nd-level adjoint functions as there are distinct first-order sensitivities $\partial R\big[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\big]\big/\partial F_j$, which is equivalent to the number of components F_j of the "feature-function" $\mathbf{F}(\boldsymbol{\theta})$. Notably, the integral operators on the left-sides of Eqs. (93) and (94) do not depend on the index "j", which means that the same left-hand side needs to be inverted for computing the 2nd-level adjoint function, regardless of the source term on the right-side (which corresponds to the particular component of the feature-function) of Eqs. (93) and (94). Therefore, if the inverses of the operators appearing on the left-sides of Eqs. (93) and (94) could be stored, they would not need to be inverted repeatedly, so the various 2nd-level adjoint functions would be computed most efficiently.

The second-order sensitivities of the decoder-response with respect to the optimal weights/parameters $\theta_k, k = 1,...,TW$, are obtained analytically by using the chain rule in conjunction with the expressions obtained in Eqs. (96) and (18), as follows:

$$\frac{\partial^2 R\big[\mathbf{F}(\boldsymbol{\theta})\big]}{\partial \theta_k \partial \theta_j} = \frac{\partial}{\partial \theta_k}\left\{\sum_{i=1}^{TF}\frac{\partial R\big[\mathbf{F}(\boldsymbol{\theta})\big]}{\partial F_i(\boldsymbol{\theta})}\frac{\partial F_i(\boldsymbol{\theta})}{\partial \theta_j}\right\}, \quad j,k = 1,...,TW. \tag{97}$$

When there are no feature functions but only individual model parameters, i.e., when $F_i(\boldsymbol{\theta}) \equiv \theta_i$ for all $i = 1,...,TF \triangleq TW$, the expression obtained in Eq. (96) yields directly the second-order sensitivities $\partial^2 R\big/\partial \theta_i \partial \theta_j$, for all $i,j = 1,...,TW$. In this case, the 2nd-LASS would need to be solved TW-times rather than just TF-times $(TF < TW)$, when TF feature functions can be constructed.

5. Illustrative Application of the 2nd-FASAM-NIE-V to Neutron Slowing Down in an Infinite Homogeneous Hydrogenous Medium

The 2nd-FASAM-NIE-V methodology developed in Section 4 will be applied in this Section to the illustrative model (considered in Section 3) that describes the energy distribution of neutrons in a homogeneous hydrogenous medium. As discussed in Section 4, the second-order sensitivities of

the decoder response, $R\left[C(E)\right]$, with respect to the feature functions $F(\boldsymbol{\theta})$, $\Sigma_d(\boldsymbol{\theta})$, and S, will be determined by considering the second-order sensitivities to be "the first-order sensitivities of the first-order sensitivities." Thus, the first-order sensitivities obtained in Eqs. (67)–(69) will play the role of "responses" in the application of the 2nd-FASAM-NIE-V methodology.

5.1. Computation of Second-Order Sensitivities Stemming from $\partial R/\partial F(\boldsymbol{\theta})$

The second-order sensitivities stemming from the first-order sensitivity $\partial R/\partial F(\boldsymbol{\theta})$ will be determined from the expression of the first-order G-variation $\delta\left\{\partial R/\partial F(\boldsymbol{\theta})\right\}$, which is by definition obtained from Eq. (67) as shown below:

$$\delta\left\{\frac{\partial R}{\partial F(\boldsymbol{\theta})}\right\} = \frac{d}{d\varepsilon}\left\{\frac{S^0 + \varepsilon\delta S}{E_s}\int_{E_\ell}^{E_s}\left[a^{(1,0)}(E) + \varepsilon\delta a^{(1)}(E)\right]\left(\frac{E_s}{E}\right)^{\left[F(\boldsymbol{\theta}^0) + \varepsilon\delta F(\boldsymbol{\theta})\right]}dE\right\}_{\varepsilon=0} \quad (98)$$

$$= \delta\left\{\frac{\partial R}{\partial F(\boldsymbol{\theta})}\right\}_{dir} + \delta\left\{\frac{\partial R}{\partial F(\boldsymbol{\theta})}\right\}_{ind},$$

where the direct-effect term $\delta\left\{\partial R/\partial F(\boldsymbol{\theta})\right\}_{dir}$ is defined as follows:

$$\delta\left\{\frac{\partial R}{\partial F(\boldsymbol{\theta})}\right\}_{dir} \triangleq \left\{\frac{(\delta S)}{E_s}\int_{E_\ell}^{E_s}a^{(1)}(E)\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}dE\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}$$

$$+\left\{\left[\delta F(\boldsymbol{\theta})\right]\frac{S}{E_s}\int_{E_\ell}^{E_s}a^{(1)}(E)\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}\ln\left(\frac{E_s}{E}\right)dE\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0},$$

$$(99)$$

while the indirect-effect term $\delta\left\{\partial R/\partial F(\boldsymbol{\theta})\right\}_{ind}$ is defined as follows:

$$\delta\left\{\frac{\partial R}{\partial F(\boldsymbol{\theta})}\right\}_{ind} \triangleq \left\{\frac{S}{E_s}\int_{E_\ell}^{E_s}\delta a^{(1)}(E)\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}dE\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}.$$

$$(100)$$

The direct-effect term can be computed at this time. On the other hand, the indirect-effect term can be computed after determining the variational function $\delta a^{(1)}(E)$, which is the solution of the G-differentiated 1st-LASS defined in Eq. (50). By definition, the G-differential of Eq. (50) is provided by the following equation:

$$\frac{d}{d\varepsilon}\left\{\left[a^{(1,0)}(E) + \varepsilon\delta a^{(1)}(E)\right]\right\}_{\varepsilon=0} - \frac{d}{d\varepsilon}\left\{\frac{F(\boldsymbol{\theta}^0) + \varepsilon\delta F(\boldsymbol{\theta})}{E}\int_{E_\ell}^{E}\left[a^{(1,0)}(e) + \varepsilon\delta a^{(1)}(e)\right]de\right\}_{\varepsilon=0} \quad (101)$$

$$= \frac{d}{d\varepsilon}\left\{\Sigma_d(\boldsymbol{\theta}^0) + \varepsilon\delta\Sigma_d(\boldsymbol{\theta})\right\}_{\varepsilon=0}.$$

Performing the operations indicated in Eq. (101) yields the following equation to be satisfied by the function $\delta a^{(1)}(E)$ at the nominal parameter values (the superscript "zero" will be omitted for notational simplicity):

$$\delta a^{(1)}(E) - \frac{F(\boldsymbol{\theta})}{E}\int_{E_\ell}^{E}\delta a^{(1)}(e)de = \frac{\delta F(\boldsymbol{\theta})}{E}\int_{E_\ell}^{E}a^{(1)}(e)de + \delta\Sigma_d(\boldsymbol{\theta}). \quad (102)$$

Evidently, Eq. (102) would need to be solved repeatedly, for every parameter variation, in order to compute the function $\delta a^{(1)}(E)$ that would correspond to the respective parameter variation. As was shown in Section 4, the need for repeatedly solving Eq. (102) can be avoided by deriving an

alternative expression for the indirect-effect term which does not involve $\delta a^{(1)}(E)$. This alternative expression is derived by introducing the 2nd-LASS, the solution of which would replace the function $\delta a^{(1)}(E)$ in the alternative expression for the indirect-effect term. Notably, Eq. (102) is independent of variations in the forward function, $\delta C(E)$; therefore, the 2nd-level adjoint sensitivity function will comprise a single component, denoted as $a^{(2)}(E;1)$, and the 2nd-LASS will comprise a single equation for this component.

Following the principles of the 2nd-FASAM-NIE-V, we use Eq. (45) to construct the inner product of Eq. (102) with a function $a^{(2)}(E;1)$, where the argument "1" indicates that this 2nd-level adjoint sensitivity function will correspond to the first-order sensitivity $\partial R[C(E)]/\partial F(\theta)$, to obtain the following relation:

$$\int_{E_\ell}^{E_s} a^{(2)}(E;1)\delta a^{(1)}(E)dE - F(\theta)\int_{E_\ell}^{E_s} a^{(2)}(E;1)\frac{dE}{E}\int_{E_\ell}^{E}\delta a^{(1)}(e)de$$
$$= \delta F(\theta)\int_{E_\ell}^{E_s} a^{(2)}(E;1)\frac{dE}{E}\int_{E_\ell}^{E} a^{(1)}(e)de + \delta\Sigma_d(\theta)\int_{E_\ell}^{E_s} a^{(2)}(E;1)dE. \tag{103}$$

The function $a^{(2)}(E;1)$ will be determined by requiring the left-side of Eq. (103) to represent the indirect-effect term defined in Eq. (100). For this purpose, the left-side will be recast using "integration by parts" into the following form:

$$\int_{E_\ell}^{E_s} a^{(2)}(E;1)\delta a^{(1)}(E)dE - F(\theta)\int_{E_\ell}^{E_s} a^{(2)}(E;1)\frac{dE}{E}\int_{E_\ell}^{E}\delta a^{(1)}(e)de$$
$$= \int_{E_\ell}^{E_s}\delta a^{(1)}(E)\left[a^{(2)}(E;1) - F(\theta)\int_{E}^{E_s} a^{(2)}(e;1)\frac{de}{e}\right]dE. \tag{104}$$

Requiring the right-side of Eq. (104) to represent the indirect-effect term defined in Eq. (100) yields the following Volterra-type 2nd-LASS to be satisfied by the 2nd-level adjoint sensitivity function $a^{(2)}(E;1)$:

$$a^{(2)}(E;1) - F(\theta)\int_{E}^{E_s} a^{(2)}(e;1)\frac{de}{e} = \frac{S}{E_s}\left(\frac{E_s}{E}\right)^{F(\theta)}. \tag{105}$$

The above 2nd-LASS is to be satisfied at the nominal parameter values θ^0. Notably, Eq. (105) is an "final-value problem" since the computation of $a^{(2)}(E;1)$ commences at the highest energy, where $a^{(2)}(E_s;1) = S/E_s$, and proceeds towards the lowest energy value, E_ℓ. For subsequent verification purposes, the closed-form explicit expression of $a^{(2)}(E;1)$ obtained by solving Eq.(105) is as follows:

$$a^{(2)}(E;1) = E^{-F(\theta)}\left[-SF(\theta)E_s^{F(\theta)-1}\ln E + SF(\theta)E_s^{F(\theta)-1}\ln E_s + SE_s^{F(\theta)-1}\right]. \tag{106}$$

It follows from Eqs. (103)–(105) that the indirect-effect term $\delta\{\partial R/\partial F(\theta)\}_{ind}$ is given by the following expression involving the 2nd-level adjoint sensitivity function $a^{(2)}(E;1)$:

$$\delta\left\{\frac{\partial R}{\partial F(\theta)}\right\}_{ind} = \delta F(\theta)\int_{E_\ell}^{E_s} a^{(2)}(E;1)\frac{dE}{E}\int_{E_\ell}^{E} a^{(1)}(e)de + \delta\Sigma_d(\theta)\int_{E_\ell}^{E_s} a^{(2)}(E;1)dE. \tag{107}$$

Adding the expression for the indirect-effect term $\delta\{\partial R/\partial F(\theta)\}_{ind}$ obtained in Eq. (107) to the expression for the direct-effect term $\delta\{\partial R/\partial F(\theta)\}_{dir}$ obtained in Eq. (99) yields the following expression for the G-differential $\delta\{\partial R/\partial F(\theta)\}$:

$$\delta\left\{\frac{\partial R}{\partial F(\boldsymbol{\theta})}\right\} = \delta F(\boldsymbol{\theta})\int_{E_\ell}^{E_s} a^{(2)}(E;1)\frac{dE}{E}\int_{E_\ell}^{E} a^{(1)}(e)\,de + \delta\Sigma_d(\boldsymbol{\theta})\int_{E_\ell}^{E_s} a^{(2)}(E;1)\,dE$$
$$+\frac{(\delta S)}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}dE + \left[\delta F(\boldsymbol{\theta})\right]\frac{S}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}\ln\left(\frac{E_s}{E}\right)dE. \tag{108}$$

The above expression is to be evaluated at the nominal parameter values $\boldsymbol{\theta}^0$.

It follows from the expression obtained in Eq. (108) that the second-order sensitivities (of the decoder response with respect to the feature functions) which stem from the first-order sensitivity $\partial R\left[C(E)\right]/\partial F(\boldsymbol{\theta})$ have the following expressions to be evaluated at the nominal parameter values $\boldsymbol{\theta}^0$:

$$\frac{\partial^2 R}{\partial^2 F(\boldsymbol{\theta})} = \int_{E_\ell}^{E_s} a^{(2)}(E;1)\frac{dE}{E}\int_{E_\ell}^{E} a^{(1)}(e)\,de + \frac{S}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}\ln\left(\frac{E_s}{E}\right)dE; \tag{109}$$

$$\frac{\partial^2 R}{\partial\Sigma_d(\boldsymbol{\theta})\partial F(\boldsymbol{\theta})} = \frac{1}{E_s}\int_{E_\ell}^{E_s} a^{(2)}(E;1)\,dE; \tag{110}$$

$$\frac{\partial^2 R}{\partial S\partial F(\boldsymbol{\theta})} = \frac{1}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}dE. \tag{111}$$

Since the 1$^{\text{st}}$-level adjoint sensitivity function $a^{(1)}(E)$ is already available, the sensitivity $\partial^2 R/\partial S\partial F(\boldsymbol{\theta})$ can be computed. The closed-form explicit expressions for the above second-order sensitivities are obtained by inserting the expression obtained in Eq. (106) for $a^{(2)}(E;1)$ and the expression obtained in Eq. (51) for $a^{(1)}(E)$, and performing the respective integrations. Carrying out these operations yields the following expressions:

$$\frac{\partial^2 R}{\partial F(\boldsymbol{\theta})\partial F(\boldsymbol{\theta})} = \left[S\Sigma_d(\boldsymbol{\theta})\right]\left\{-2\left[1-F(\boldsymbol{\theta})\right]^{-3}\left[1-\left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\right]\right.$$
$$\left. -2\left[1-F(\boldsymbol{\theta})\right]^{-2}\left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\ln\left(\frac{E_s}{E_\ell}\right) - \frac{F(\boldsymbol{\theta})}{1-F(\boldsymbol{\theta})}\left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\left[\ln\left(\frac{E_s}{E_\ell}\right)\right]^2\right\}; \tag{112}$$

$$\frac{\partial^2 R}{\partial\Sigma_d(\boldsymbol{\theta})\partial F(\boldsymbol{\theta})} = \frac{S}{1-F(\boldsymbol{\theta})}\left\{\frac{1}{1-F(\boldsymbol{\theta})}\left[1-\left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\right] - F(\boldsymbol{\theta})\left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\ln\left(\frac{E_s}{E_\ell}\right)\right\}; \tag{113}$$

$$\frac{\partial^2 R}{\partial S\partial F(\boldsymbol{\theta})} = \frac{\Sigma_d(\boldsymbol{\theta})}{1-F(\boldsymbol{\theta})}\left\{\frac{1}{1-F(\boldsymbol{\theta})}\left[1-\left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\right] - F(\boldsymbol{\theta})\left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1}\ln\left(\frac{E_s}{E_\ell}\right)\right\}. \tag{114}$$

5.2. Computation of Second-Order Sensitivities Stemming from $\partial R/\partial\Sigma_d(\boldsymbol{\theta})$

The second-order sensitivities stemming from the first-order sensitivity $\partial R/\partial\Sigma_d(\boldsymbol{\theta})$ will be determined from the expression of the first-order G-variation $\delta\left\{\partial R/\partial\Sigma_d(\boldsymbol{\theta})\right\}$, which is by definition obtained from Eq. (68) as shown below:

$$\delta\left\{\frac{\partial R}{\partial\Sigma_d(\boldsymbol{\theta})}\right\} = \frac{d}{d\varepsilon}\left\{\int_{E_\ell}^{E_s}\left[C^0(E)+\varepsilon\delta C(E)\right]dE\right\}_{\varepsilon=0} = \int_{E_\ell}^{E_s}\delta C(E)\,dE. \tag{115}$$

Comparing Eq. (115) with Eq. (35), it becomes apparent that the following relation holds:

$$\delta\left\{\partial R/\partial\Sigma_d\left(\boldsymbol{\theta}\right)\right\} = \left\{\delta R\left(C;\boldsymbol{\theta};\delta C\right)\right\}_{ind}\Big/\Sigma_d\left(\boldsymbol{\theta}\right).$$

(116)

Replacing the expression obtained for $\left\{\delta R\left(C;\boldsymbol{\theta};\delta C\right)\right\}_{ind}$ in Eq. (65) into Eq. (116) yields the following expression:

$$\delta\left\{\frac{\partial R}{\partial\Sigma_d\left(\boldsymbol{\theta}\right)}\right\} = \frac{F\left(\delta S\right)}{E_s\Sigma_d}\int_{E_\ell}^{E_s}a^{(1)}\left(E\right)dE + \frac{S\left(\delta F\right)}{E_s\Sigma_d}\int_{E_\ell}^{E_s}a^{(1)}\left(E\right)dE\left(\frac{E_s}{E}\right)^{F\left(\boldsymbol{\theta}\right)}.$$

(117)

It follows from Eq. (117) that the second-order sensitivities that stem from $\partial R/\partial\Sigma_d$ have the following expressions:

$$\frac{\partial^2 R}{\partial F\left(\boldsymbol{\theta}\right)\partial\Sigma_d\left(\boldsymbol{\theta}\right)} = \frac{S}{E_s\Sigma_d\left(\boldsymbol{\theta}\right)}\int_{E_\ell}^{E_s}a^{(1)}\left(E\right)dE\left(\frac{E_s}{E}\right)^{F\left(\boldsymbol{\theta}\right)};$$

(118)

$$\frac{\partial^2 R}{\delta\Sigma_d\left(\boldsymbol{\theta}\right)\partial\Sigma_d\left(\boldsymbol{\theta}\right)} = 0;$$

(119)

$$\frac{\partial^2 R}{\delta S\partial\Sigma_d\left(\boldsymbol{\theta}\right)} = \frac{F\left(\boldsymbol{\theta}\right)}{E_s\Sigma_d\left(\boldsymbol{\theta}\right)}\int_{E_\ell}^{E_s}a^{(1)}\left(E\right)dE.$$

(120)

The explicit closed-form expressions of the above second-order sensitivities are obtained by replacing the expression of $a^{(1)}\left(E\right)$ in Eqs. (118) and (120), respectively, and performing the respective integrations to obtain:

$$\frac{\partial^2 R}{\partial F\left(\boldsymbol{\theta}\right)\partial\Sigma_d\left(\boldsymbol{\theta}\right)} = \frac{S}{1-F\left(\boldsymbol{\theta}\right)}\left\{\frac{1}{1-F\left(\boldsymbol{\theta}\right)}\left[1-\left(\frac{E_s}{E_\ell}\right)^{F\left(\boldsymbol{\theta}\right)-1}\right] - F\left(\boldsymbol{\theta}\right)\left(\frac{E_s}{E_\ell}\right)^{F\left(\boldsymbol{\theta}\right)-1}\ln\left(\frac{E_s}{E_\ell}\right)\right\};$$

(121)

$$\frac{\partial^2 R}{\delta S\partial\Sigma_d\left(\boldsymbol{\theta}\right)} = \frac{F\left(\boldsymbol{\theta}\right)}{1-F\left(\boldsymbol{\theta}\right)}\left[1-\left(\frac{E_s}{E_\ell}\right)^{F\left(\boldsymbol{\theta}\right)-1}\right].$$

(122)

The expression for $\partial^2 R/\partial F\left(\boldsymbol{\theta}\right)\partial\Sigma_d\left(\boldsymbol{\theta}\right)$ in Eq. (118) must be equivalent to the expression for $\partial^2 R/\partial\Sigma_d\left(\boldsymbol{\theta}\right)\partial F\left(\boldsymbol{\theta}\right)$ in Eq. (110). Notably, therefore, these mixed second-order sensitivities are computed twice, using distinct expressions involving distinct adjoint functions, which provides an intrinsic mechanism for the stringent verification of the accuracy of the computations used for obtaining the numerical values of the respective adjoint functions.

5.3. Computation of Second-Order Sensitivities Stemming from $\partial R/\partial S$

The second-order sensitivities stemming from the first-order sensitivity $\partial R/\partial S$ will be determined from the expression of the first-order G-variation $\delta\left\{\partial R/\partial S\right\}$, which is by definition obtained from Eq. (69) as shown below:

$$\delta\left\{\frac{\partial R}{\partial S}\right\} = \frac{d}{d\varepsilon}\left\{\frac{\left[F\left(\boldsymbol{\theta}^0\right)+\varepsilon\delta F\left(\boldsymbol{\theta}\right)\right]}{E_s}\int_{E_\ell}^{E_s}\left[a^{(1,0)}\left(E\right)+\varepsilon\delta a^{(1)}\left(E\right)\right]dE\right\}_{\varepsilon=0}$$

$$= \delta\left\{\frac{\partial R}{\partial S}\right\}_{dir} + \delta\left\{\frac{\partial R}{\partial S}\right\}_{ind},$$

(123)

where the direct-effect term $\delta\{\partial R/\partial S\}_{dir}$ is defined as follows:

$$\delta\left\{\frac{\partial R}{\partial S}\right\}_{dir} \triangleq \left\{\frac{\delta F(\boldsymbol{\theta})}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)dE\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0},$$

(124)

while the indirect-effect term $\delta\{\partial R/\partial S\}_{ind}$ is defined as follows:

$$\delta\left\{\frac{\partial R}{\partial S}\right\}_{ind} \triangleq \left\{\frac{F(\boldsymbol{\theta})}{E_s}\int_{E_\ell}^{E_s} \delta a^{(1)}(E)dE\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}.$$

(125)

The direct-effect term $\delta\{\partial R/\partial S\}_{dir}$ can be computed at this time. To determine the indirect-effect term $\delta\{\partial R/\partial S\}_{ind}$ without needing to compute the variational function $\delta a^{(1)}(E)$ by solving repeatedly Eq. (102), the steps outlined in Section 5.1 are applied using a 2nd-order adjoint sensitivity function denoted as $a^{(2)}(E;3)$, where the argument "3" indicates that this function will correspond to the first-order sensitivity $\partial R/\partial S$. Thus, following the conceptual steps outlined in Eqs. (103)–(107) but for the function $a^{(2)}(E;3)$ as the counterpart of the function $a^{(2)}(E;1)$, and for the indirect-effect term $\delta\{\partial R/\partial S\}_{ind}$ as the counterpart of the indirect-effect term $\delta\{\partial R/\partial F(\boldsymbol{\theta})\}_{ind}$, leads to the following expression for the indirect-effect term $\delta\{\partial R/\partial S\}_{ind}$:

$$\delta\left\{\frac{\partial R}{\partial S}\right\}_{ind} \triangleq \left\{\delta F(\boldsymbol{\theta})\int_{E_\ell}^{E_s} a^{(2)}(E;3)\frac{dE}{E}\int_{E_\ell}^{E} a^{(1)}(e)de + \delta\Sigma_d(\boldsymbol{\theta})\int_{E_\ell}^{E_s} a^{(2)}(E;3)dE\right\}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0},\quad (126)$$

where the 2nd-order adjoint sensitivity function $a^{(2)}(E;3)$ is the solution of the following Volterra-type 2nd-LASS:

$$a^{(2)}(E;3) - F(\boldsymbol{\theta})\int_{E}^{E_s} a^{(2)}(e;3)\frac{de}{e} = \frac{F(\boldsymbol{\theta})}{E_s}.$$

(127)

For verification purposes, the explicit closed-form expression of the solution of Eq. (127) is provided below.

$$a^{(2)}(E;3) = \frac{F(\boldsymbol{\theta})}{E_s}\left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})}.$$

(128)

Adding the expressions obtained in Eqs. (126) and (124) yields the following expression for the first G-variation $\delta\{\partial R/\partial S\}$:

$$\delta\left\{\frac{\partial R}{\partial S}\right\} = \frac{\delta F(\boldsymbol{\theta})}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)dE + \delta F(\boldsymbol{\theta})\int_{E_\ell}^{E_s} a^{(2)}(E;3)\frac{dE}{E}\int_{E_\ell}^{E} a^{(1)}(e)de$$
$$+\delta\Sigma_d(\boldsymbol{\theta})\int_{E_\ell}^{E_s} a^{(2)}(E;3)dE.$$

(129)

It follows from Eq. (129) that the second-order sensitivities stemming from the first-order sensitivity $\partial R/\partial S$ have the following expressions:

$$\frac{\partial^2 R}{\partial F(\boldsymbol{\theta})\partial S} = \frac{1}{E_s}\int_{E_\ell}^{E_s} a^{(1)}(E)dE + \int_{E_\ell}^{E_s} a^{(2)}(E;3)\frac{dE}{E}\int_{E_\ell}^{E} a^{(1)}(e)de$$

(130)

$$\frac{\partial^2 R}{\partial\Sigma_d(\boldsymbol{\theta})\partial S} = \int_{E_\ell}^{E_s} a^{(2)}(E;3)dE.$$

(131)

$$\frac{\partial^2 R}{\partial S \partial S} = 0 .$$

$$(132)$$

Inserting the expressions of $a^{(2)}(E;3)$ and $a^{(1)}(E)$, respectively, into Eqs. (130) and (131), and performing the respective integrations, yields the following closed-form explicit expressions for the respective second-order sensitivities:

$$\frac{\partial R}{\partial F(\boldsymbol{\theta}) \partial S} = \frac{\Sigma_d(\boldsymbol{\theta})}{1 - F(\boldsymbol{\theta})} \left\{ \frac{1}{1 - F(\boldsymbol{\theta})} \left[1 - \left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1} \right] - F(\boldsymbol{\theta}) \left(\frac{E_s}{E_\ell}\right)^{F(\boldsymbol{\theta})-1} \ln\left(\frac{E_s}{E_\ell}\right) \right\}; \qquad (133)$$

$$\frac{\partial^2 R}{\partial \Sigma_d(\boldsymbol{\theta}) \partial S} = \frac{F(\boldsymbol{\theta})}{1 - F(\boldsymbol{\theta})} \left[1 - \left(\frac{E_s}{E}\right)^{F(\boldsymbol{\theta})-1} \right]. \qquad (134)$$

The expression for $\partial^2 R / \partial S F(\boldsymbol{\theta})$ in Eq. (111) must be equivalent to the expression for $\partial^2 R / \partial F(\boldsymbol{\theta}) \partial S$ in Eq. (130). The expression for $\partial^2 R / \partial S \partial \Sigma_d(\boldsymbol{\theta})$ in Eq. (122) must be equivalent to the expression for $\partial^2 R / \partial \Sigma_d(\boldsymbol{\theta}) \partial S$ in Eq. (131). The equivalences of these corresponding expressions provide stringent verification criteria for the accuracy of the computation of the respective adjoint functions.

5.4. Discussion: Direct Computation of Second-Order Sensitivities Versus Their Indirect Computation via Feature Functions

Notably, the 9 second-order sensitivities of the decoder response with respect to the 3 feature functions, $F(\boldsymbol{\theta})$, $\Sigma_d(\boldsymbol{\theta})$ and S were computed using 3 adjoint computations; each of these adjoint computations corresponds to one of the 3 first-order sensitivities of the decoder response with respect to the feature functions. Only 6 of these 9 second-order sensitivities have distinct values; the mixed sensitivities were computed twice, using different adjoint functions thus providing stringent verification criteria for the numerical computations of these functions. The second-order sensitivities of the decoder response with respect to the primary model parameters are obtained by applying the "chain-rule of differentiation" provided in Eq. (97) to the second-order sensitivities with respect to the feature functions.

On the other hand, the computation of the second-order sensitivities of the decoder response directly with respect to the primary model parameters would be performed by treating each of the first-order sensitivities defined in Eqs. (54)–(57) as a "decoder/model response." A shown in Section 3.1, there would be $TW+1$ first-order sensitivities in this case, which means that there would be $TW+1$ "2nd-level adjoint sensitivity systems" to be solved, each one having a source-term that would correspond to one of the $TW+1$ first-order sensitivities. Evidently, it is considerably more advantageous computationally to consider compute the second-order sensitivities via "feature functions," whenever possible, rather than directly with respect to the primary model parameters.

6. Discussion and Conclusions

This work has introduced the general mathematical framework of the 2nd-FASAM-NIE-V methodology. The acronym "2nd-FASAM-NIE-V" stands for "Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Volterra Type." The 2nd-FASAM-NIE-V encompasses the mathematical framework of the (first-order) 1st-FASAM-NIE-V methodology, which enables the most efficient computation of the exact expressions of all first-order sensitivities with respect to the feature functions and also with respect to the optimal values of the NIE-net's parameters/weights, after the respective NIE-Volterra-net was optimized to represent the underlying physical system. The 1st-FASAM-NIE-V methodology requires a single large-scale

computation for determining the first-level adjoint sensitivity function that is subsequently used for computing the sensitivities using conventional numerical quadrature formulas. The 2nd-FASAM-NIE-V requires as many large-scale computations (to solve the 2nd-Level Adjoint Sensitivity System) as there are first-order sensitivities of the decoder response with respect to the feature functions. Subsequently, the second-order sensitivities of the decoder response with respect to the primary model parameters are obtained by applying the "chain-rule of differentiation" to the second-order sensitivities with respect to the feature functions.

The application of the 1st-FASAM-NIE-V and the 2nd-FASAM-NIE-V methodologies has been illustrated by using a well-known model for neutron slowing down in a homogeneous hydrogenous medium. This model has been chosen because the application of the 1st-FASAM-NIE-V and the 2nd-FASAM-NIE-V yields tractable explicit exact expressions for all quantities of interest, including the various adjoint sensitivity functions and first- and second-order sensitivities of the decoder response with respect to all feature functions and also with respect to the primary model parameters. This illustrative application highlights the unsurpassed efficiency of the 1st-FASAM-NIE-V and the 2nd-FASAM-NIE-V for second-order sensitivity analysis of NIE-Volterra nets. Ongoing research aims at developing the "Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integro-Differential Equations," which will enable, in premiere, the exact computations of second-order sensitivities of decoder responses with respect to optimized weights/parameters, based on the NIDE-models introduced in [22].

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Chen, R.T.Q.; Rubanova, Y.; Bettencourt, J.; Duvenaud, D.K. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: New York, NY, USA, 2018; Volume 31, pp. 6571–6583. https://doi.org/10.48550/arXiv.1806.07366.
2. Ruthotto, L.; Haber, E. Deep neural networks motivated by partial differential equations. *J. Math. Imaging Vis.* **2018**, *62*, 352–364.
3. Lu, Y.; Zhong, A.; Li, Q.; Dong, B. Beyond finite layer neural networks: Bridging deep architectures and numerical differential equations. In Proceedings of the International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; pp. 3276–3285.
4. Dupont, E.; Doucet, A.; The, Y.W. Augmented neural odes. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; Volume 32, pp. 14–15.
5. Grathwohl, W.; Chen, R.T.Q.; Bettencourt, J.; Sutskever, I.; Duvenaud, D. Ffjord: Free-form continuous dynamics for scalable reversible generative models. In Proceedings of the International Conference on Learning Representations, New Orleans, LA, USA, 6–9 May 2019.
6. Zhong, Y.D.; Dey, B.; Chakraborty, A. Symplectic ode-net: Learning Hamiltonian dynamics with control. In Proceedings of the International Conference on Learning Representations, Addis Ababa, Ethiopia, 30 April 2020.
7. Kidger, P.; Morrill, J.; Foster, J.; Lyons, T. Neural controlled differential equations for irregular time series. In Proceedings of the Advances in Neural Information Processing Systems, Virtual, 6–12 December 2020; Volume 33, pp. 6696–6707.
8. Morrill, J.; Salvi, C.; Kidger, P.; Foster, J. Neural rough differential equations for long time series. In Proceedings of the International Conference on Machine Learning, Virtual, 18–24 July 2021; pp. 7829–7838.
9. Kidger, P. On Neural Differential Equations. *arXiv* **2022**, arXiv:2202.02435.
10. Rokhlin, V. Rapid solution of integral equations of classical potential theory. *J. Comput. Phys.* **1985**, *60*, 187–207.

11. Rokhlin, V. Rapid solution of integral equations of scattering theory in two dimensions. *J. Comput. Phys.* **1990**, *86*, 414–439.

12. Greengard, L.; Kropinski, M.C. An integral equation approach to the incompressible Navier-Stokes equations in two dimensions. *SIAM J. Sci. Comput.* **1998**, *20*, 318–336.

13. Effati, S.; Buzhabadi, R. A neural network approach for solving Fredholm integral equations of the second kind. *Neural Comput. Appl.* **2012**, *21*, 843–852.

14. Zappala, E.; de Oliveira Fonseca, A.H.; Caro, J.O.; van Dijk, D. Neural Integral Equations. *arXiv* **2023**, arXiv:2209.15190v4.

15. Xiong, Y.; Zeng, Z.; Chakraborty, R.; Tan, M.; Fung, G.; Li, Y.; Singh, V. Nystromformer: A nystrom-based algorithm for approximating self-attention. *Proc. AAAI Conf. Artif. Intell.* **2021**, *35*, 14138.

16. Cacuci, D.G. Introducing the nth-Order Features Adjoint Sensitivity Analysis Methodology for Nonlinear Systems (n^{th}-FASAM-N): I. Mathematical Framework. *Am. J. Comput. Math.* **2024**, *14*, 11–42. https://doi.org/10.4236/ajcm.2024.141002. See also: Dan Gabriel Cacuci, *The nth-Order Comprehensive Adjoint Sensitivity Analysis Methodology (n^{th}-CASAM): Overcoming the Curse of Dimensionality in Sensitivity and Uncertainty Analysis, Volume III: Nonlinear Systems*, 369 pages, Springer Nature Switzerland, Cham, 2023, https://doi.org/10.1007/978-3-031-22757-8.

17. Cacuci, D.G. Introducing the Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Ordinary Differential Equations. I: Mathematical Framework. *Processes* **2024**, *12*, 2660. https://doi.org/10.3390/pr12122660.

18. Cacuci, D.G., "Introducing the Second-Order Features Adjoint Sensitivity Analysis Methodology for Fredholm-Type Neural Integral Equations. *Mathematics* **2025**, *13*, 14. https://doi.org/10.3390/math13010014.

19. Weinberg, A. M., Wigner, E.P. *The Physical Theory of Neutron Chain Reactors*, The University of Chicago Press, Chicago, Il., USA, 1958.

20. Lamarsh, J. R., *Introduction to Nuclear Reactor Theory*, Addison-Wesley Publishing Company, Reading, Massachusetts, USA, 1966.

21. Duderstadt, J. J.; Hamilton, L. J., *Nuclear Reactor Analysis*, John Wiley & Sons, New York, USA, 1976.

22. Zappala, E.; de Oliveira Fonseca, A. H.; Moberly, A.H.; Higley, J.M.; Abdallah, C.; Cardin, J.; Van Dijk, D.; Neural Integro-Differential Equations, arXiv:2206.14282v1 [cs.LG] 28 Jun 2022.