Article

Not peer-reviewed version

Introducing the Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Fredholm-Type

Dan Gabriel Cacuci *

Posted Date: 5 December 2024

doi: 10.20944/preprints202412.0474.v1

Keywords: neural integral equations; nonlinear Fredholm-type neural integral equations; first-order features adjoint sensitivity analysis methodology; second-order features adjoint sensitivity analysis methodology

Article

Introducing the Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Fredholm-Type

Dan Gabriel Cacuci

University of South Carolina, Department of Mechanical Engineering; cacuci@cec.sc.edu

Abstract: This works presents the "First-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Fredholm-Type" (1st-FASAM-NIE-Fredholm) and the "Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Fredholm-Type" (2nd-FASAM-NIE-Fredholm). It is shown that the 1st-FASAM-NIE-Fredholm methodology enables the efficient computation of exactly-determined first-order sensitivities of decoder response with respect to the optimized NIE-parameters, requiring a single "large-scale" computation for solving the 1st-Level Adjoint Sensitivity System (1st-LASS), regardless of the number of weights/parameters underlying the NIE-net. The 2nd-FASAM-NIE-Fredholm methodology enables the computation, with unparalleled efficiency, of the second-order sensitivities of decoder responses with respect to the optimized/trained weights involved in the NIE's decoder, hidden layers, and encoder, requiring only as many "large-scale" computations as there are first-order sensitivities with respect to the feature functions. The application of both the 1st-FASAM-NIE-Fredholm and the 2nd-FASAM-NIE-Fredholm methodologies is illustrated by considering a system of nonlinear Fredholm-type NIE which admits analytical solutions, thereby facilitating the verification of the expressions obtained for the first- and second-order sensitivities of NIE-decoder responses with respect to the model parameters (weights) that characterize the respective NIE-net.

Keywords: neural integral equations; nonlinear Fredholm-type neural integral equations; first-order features adjoint sensitivity analysis methodology; second-order features adjoint sensitivity analysis methodology

1. Introduction

Ordinary differential equations are widely employed tools for modeling continuous dynamical systems. Neural Ordinary Differential Equations (NODE), formally introduced by Chen et al. [1], have enabled the use of deep learning for modeling discretely sampled dynamical systems. NODE provide an explicit connection between deep feed-forward neural networks and dynamical systems [2,3], while providing a bridge between modern deep learning and traditional numerical modelling. NODE provide a flexible trade-off between efficiency, memory costs and accuracy. The approximation capabilities [4,5] of NODE are particularly useful for modelling time-series [1,6,7], continuous normalizing flows [1,8], and for modeling and controlling physical environments [see, e.g., 9]. However, NODE are limited to describing systems that are instantaneous, each time-step being determined locally in time, without contributions from the state of the system at other times.

In contradistinction to NODE, integral equations (IE) model global "long-distance" spatiotemporal relations. Moreover, ordinary differential equations (ODE) and/or partial differential equations (PDE) can often be recast in integral-equation forms that can be solved more efficiently using IE solvers, since IE solvers often possess stability properties that are superior to those of ODE and PDE solvers, as exemplified in classical potential theory [10], scattering theory [11], fluid flow [12], and integral neutron and photon transport [13,14].

In practice, it is desired to model the system under consideration by learning the operator that can reproduce the system by using data sampled from the respective system. Typical operator learning problems are formulated on finite grids, using finite-difference methods that approximate the domain of the functions under investigation. Recovering the continuous limit is a very challenging problem, particularly since irregularly sampled data may alter the evaluation of the learned operator. Operator learning entails the formulation of the operator learning problem through an IE solver [see, e.g., 15] which samples the domain of integration continuously. Due to their highly non-local behavior, IE solvers are suitable for modeling complex dynamics. The problem of learning dynamics from data through integral equations has been addressed by Zappala et al [16], who have introduced the Neural Integral Equation (NIE) and the Attentional Neural Integral Equation (ANIE). The NIE and the ANIE can be used to generate dynamics and can also be used to infer the spatiotemporal relations that 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 [17]; this interpretation makes it possible to approximate the integral kernel of the model using self-attention, which enables efficient integration over higher dimensions.

Neural nets are trained by minimizing a “loss function” which is usually meant to represent the discrepancy between a “reference solution” and the output produced by the respective net’s decoder. The minimization procedure requires the computation of the gradients of the loss function with respect to the weights to be optimized. After the neural-net is optimized to reproduce the underlying physical system as closely as possible, the subsequent responses of interest become various functionals of the net’s decoder output rather than some “loss function.” Furthermore, the physical system modeled by a neural-net comprises parameters that stem from measurements and/or computations. Such parameters are not perfectly well known but are subject to uncertainties that stem from the experiments and/or computations that underly the origin of the respective parameters. Hence, it is important to quantify the uncertainties induced in the decoder’s output by the uncertainties that afflict the parameters/weights underlying the physical system modeled by the respective neural-net. The quantification of the uncertainties in the net’s decoder and derived results (called ”responses”) of interest require the computation of 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 of such scalar model parameters, including correlations, material properties, etc. It is convenient to refer to such scalar-valued functions as “features of primary model parameters.” Cacuci [18] has recently introduced 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.” Subsequently, 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/functions of parameters.” Particular forms of the n^{th}-FASAM-N enabled the development of specific sensitivity analysis methodologies for NODE-nets, as follows: the “First-Order Features Adjoint Sensitivity Analysis Methodology for Neural Ordinary Differential Equations (1^{st}-FASAM-NODE)” [19] and the “Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Ordinary Differential Equations (2^{nd}-FASAM-NODE)” [20]. The 1^{st}-FASAM-NODE and the 2^{nd}-FASAM-NODE are pioneering sensitivity analysis methodologies which enable the computation, with unparalleled efficiency, of exactly-determined first-order and, respectively, second-order sensitivities of decoder response with respect to the optimized/trained weights involved in the NODE’s decoder, hidden layers, and encoder. The applications of both the 1^{st}-FASAM-NODE and the 2^{nd}-FASAM-NODE methodologies were illustrated by performing first-and second-order sensitivity analyses of the heat and energy transfer processes in the Nordheim-Fuchs phenomenological model for reactor safety [19,20].

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

By applying the general concepts underlying the n^{th}-FASAM-N methodology [18], this works introduces the "First-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Fredholm-Type (1^{st}-FASAM-NIE-Fredholm)" and the "Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Fredholm-Type (2^{nd}-FASAM-NIE-Fredholm)." The 1^{st}-FASAM-NIE-Fredholm methodology is introduced in Section 2. Subsection 2.1 presents the particular case when there are no feature functions but only model parameters, in which case the FASAM-NIE-Fredholm reduces to the "First-Order Comprehensive Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Fredholm-Type" (1^{st}-CASAM-NIE-Fredholm). The application of the 1^{st}-FASAM-NIE-Fredholm Methodology is illustrated in Section 3 by considering a system of coupled nonlinear Fredholm-type NIE which admits analytical solutions, thereby facilitating the verification of the expressions obtained for the first-order sensitivities of NIE-decoder responses with respect to the model parameters (weights) that characterize the respective NIE-net.

The 2^{nd}-FASAM-NIE-Fredholm methodology is introduced in Section 4. When there are no feature functions but only individual model parameters, the 2^{nd}-FASAM-NIE-Fredholm reduces, as a particular case, to the "Second-Order Comprehensive Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Fredholm-Type" (2^{nd}-CASAM-NIE-Fredholm), which is presented in Subsection 4.1. The application of the 2^{nd}-FASAM-NIE-Fredholm methodology is illustrated in Section 5 by continuing the analysis of the model considered in Section 3. As generally illustrated in Section 4, the second-order sensitivities are conceptually determined by considering them to be the "first-order sensitivities of the first-order sensitivities."

The discussion presented in Section 6 concludes this work, noting that the 1^{st}-FASAM-NIE-Fredholm methodology enables the computation, with unparalleled efficiency, of exactly-determined first-order sensitivities of decoder response with respect to the NODE-parameters, requiring a single "large-scale" computation for solving the 1^{st}-Level Adjoint Sensitivity System (1^{st}-LASS), regardless of the number of weights/parameters underlying the NIE-net. When "feature functions of parameters "can be identified within the NIE structure, the number of quadratures for computing the first-order sensitivities is smaller than the number of quadratures needed for computing the first-order decoder-response sensitivities directly with respect to the parameters, since the latter can be computed analytically and exactly by using the first-order sensitivities with respect to the feature functions.

The 2^{nd}-FASAM-NIE-Fredholm methodology enables the computation (with unparalleled efficiency) of exactly-determined second-order sensitivities of decoder response with respect to the NODE-parameters. In order to compute all of the second-order sensitivities of a decoder-response with respect to the respective the feature functions, the 2^{nd}-FASAM-NIE-Fredholm methodology requires only as many "large-scale" computations as there are first-order sensitivities with respect to the feature functions. When no "feature functions" can be constructed, the 2^{nd}-FASAM-NIE-Fredholm methodology requires as many "large-scale" computations as there are first-order sensitivities of the decoder-response with respect to the model parameters.

As has already been mentioned, the 1^{st}-FASAM-NIE-Fredholm and the 2^{nd}-FASAM-NIE-Fredholm methodologies are applicable to many scientific fields, including classical potential theory [10], scattering theory [11], fluid flow [12], and integral neutron and photon transport [13,14]. Ongoing work aims at developing the Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Volterra-Type (2^{nd}-FASAM-NIE-Volterra). Subsequent work will aim at generalizing these developments to address the efficient computation of exact expressions of high-order sensitivities of systems that can be modeled using Neural Integro-Differential Equations.

2. First-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Fredholm-Type (1^{st}-FASAM-NIE-Fredholm)

A network of nonlinear "Neural Integral Equations (NIE)" of Fredholm-type and its decoder-response can be generally represented as follows:

$$\mathbf{h}(t,\mathbf{x})=\mathbf{g}\left[\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]+\int_{t_0}^{t_f}d\tau\int_{\Omega}\mathbf{G}\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F}(\boldsymbol{\theta});t,\tau;\mathbf{x},\mathbf{z}\right]d\mathbf{z}, \tag{1}$$

$$R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]=\int_{t_0}^{t_f}dt\int_{\Omega}D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]d\mathbf{x} \tag{2}$$

where:

(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. The initial value is denoted as t_0 (which can be considered to be an initial measurement time) while the stopping value is denoted as t_f (which can be considered to be the next measurement time). Often, the variable t is called the "global time" while the variable τ is called the "local time."

(ii) The TD-dimensional column vector $\mathbf{x} \triangleq [x_1,...,x_{TD}]^\dagger \in R^{TD}$ comprises, as components, the real-valued independent variables (spatial, energy, solid angle, etc.) that characterize the phase-space under consideration. The vector $\mathbf{x} \triangleq [x_1,...,x_{TD}]^\dagger$, where "$TD$" denotes the "total number of dimensions of the phase-space under consideration," is defined on a domain denoted as Ω. The vector $\mathbf{z} \triangleq [z_1,...,z_{TD}]^\dagger \in R^{TD}$ is defined in the same way as $\mathbf{x} \triangleq [x_1,...,x_{TD}]^\dagger \in R^{TD}$. 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 $\mathbf{h}(t,\mathbf{x}) \triangleq \left[h_1(t,\mathbf{x}),...,h_{TH}(t,\mathbf{x})\right]^\dagger$ represents the hidden/latent neural networks.

(iv) The TH-dimensional vector-valued function $\mathbf{G}\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F}(\boldsymbol{\theta});t,\tau;\mathbf{x},\mathbf{z}\right] \triangleq \left[G_1(\mathbf{h};\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}),...,G_{TH}(\mathbf{h};\mathbf{F};t,\tau;\mathbf{x},\mathbf{z})\right]^\dagger$ models the dynamics of the latent neurons. The components of the vector $\boldsymbol{\theta} \triangleq [\theta_1,...,\theta_{TW}]^\dagger$ represents 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 $\boldsymbol{\theta} \triangleq [\theta_1,...,\theta_{TW}]^\dagger$ 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. 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 general, $\mathbf{G}\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F}(\boldsymbol{\theta});t,\tau;\mathbf{x},\mathbf{z}\right]$ is a nonlinear function of $\mathbf{h}(t,\mathbf{x})$ and $\mathbf{F}(\boldsymbol{\theta})$.

(v) The scalar-valued quantity $R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]$ is a functional of $\mathbf{h}(t,\mathbf{x})$ and $\mathbf{F}(\boldsymbol{\theta})$, and represents the network's decoder-response. The function $D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]$ models the decoder and may contain Dirac-delta distributions, if the decoder-response is to be evaluated at some particular point in time and/or phase-space.

The NIE-Fredholm-net defined in Equation (1) has the following representation in component form,

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

$$h_i\left(t,\mathbf{x}\right)=g_i\left[\mathbf{F}\left(\boldsymbol{\theta}\right);t;\mathbf{x}\right]+\int_{t_0}^{t_f}d\tau\int_{\Omega}G_i\left[\mathbf{h}\left(\tau;\mathbf{z}\right);\mathbf{F}\left(\boldsymbol{\theta}\right);t,\tau;\mathbf{x},\mathbf{z}\right]d\mathbf{z};\ i=1,...,TH. \tag{3}$$

After the "training" of the NIE-net has been accomplished by using the "adjoint" or other methods to minimize the user-chosen "loss functional" meant to represent the discrepancy between a "reference solution" and the output produced by the NIE-decoder, the primary parameters ("weights") $\boldsymbol{\theta}\triangleq\left[\theta_1,...,\theta_{TW}\right]^{\dagger}$ will have been assigned "optimal" values which will have minimized the "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 $\mathbf{h}^0\left(t,\mathbf{x}\right)$, which will satisfy the following forms of Equations (1) and (2):

$$\mathbf{h}^0\left(t,\mathbf{x}\right)=\mathbf{g}\left[\mathbf{F}\left(\boldsymbol{\theta}^0\right);t;\mathbf{x}\right]+\int_{t_0}^{t_f}d\tau\int_{\Omega}\mathbf{G}\left[\mathbf{h}^0\left(\tau;\mathbf{z}\right);\mathbf{F}\left(\boldsymbol{\theta}^0\right);t,\tau;\mathbf{x},\mathbf{z}\right]d\mathbf{z}, \tag{4}$$

$$R\left[\mathbf{h}^0;\mathbf{F}\left(\boldsymbol{\theta}^0\right)\right]=\int_{t_0}^{t_f}dt\int_{\Omega}D\left[\mathbf{h}^0\left(t;\mathbf{x}\right);\mathbf{F}\left(\boldsymbol{\theta}^0\right);t;\mathbf{x}\right]d\mathbf{x}. \tag{5}$$

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" output, which can be generally represented by the functional $R\left[\mathbf{h};\mathbf{F}\left(\boldsymbol{\theta}\right)\right]$ defined in Equation (2). Since the physical system modeled by the NIE-net comprises parameters that stem from measurements and/or computations, the optimal weights/parameters are unavoidably afflicted by uncertainties that stem from the respective experiments and/or computations. Hence, it is important to quantify the uncertainties induced in the NIE-decoder output by the uncertainties that afflict the parameters/weights underlying the physical system modeled by the NIE-net. The quantification of the uncertainties in the NIE's decoder-responses of interest require the computation of the sensitivities of the NIE decoder-response with respect to the optimized NIE weights/parameters.

The known nominal values $\boldsymbol{\theta}^0$ of the primary model parameters ("weights") characterizing the IDE-net will differ from the true but unknown values $\boldsymbol{\theta}$ of the respective weights 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 $\delta\mathbf{F}\triangleq\mathbf{F}\left(\boldsymbol{\theta}\right)-\mathbf{F}^0$, $\mathbf{F}^0\triangleq\mathbf{F}\left(\boldsymbol{\theta}^0\right)$, in the feature functions and, variations $\mathbf{v}^{(1)}\left(t;\mathbf{x}\right)\triangleq\left[v_1^{(1)}\left(t;\mathbf{x}\right),\ldots,v_{TH}^{(1)}\left(t;\mathbf{x}\right)\right]^{\dagger}\triangleq\left[\delta h_1\left(t;\mathbf{x}\right),\ldots,\delta h_{TH}\left(t;\mathbf{x}\right)\right]^{\dagger}\triangleq\mathbf{h}\left(t;\mathbf{x}\right)-\mathbf{h}^0\left(t;\mathbf{x}\right)$ around the nominal/optimal functions $\mathbf{h}^0\left(t\right)$. In turn, the variations $\delta\mathbf{F}\triangleq\mathbf{F}\left(\boldsymbol{\theta}\right)-\mathbf{F}^0$ and $\mathbf{v}^{(1)}\left(t;\mathbf{x}\right)$ will induce variations of the form $\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)};\delta\mathbf{F};t;\mathbf{x}\right)\triangleq R\left[\mathbf{h};\mathbf{F}\left(\boldsymbol{\theta}\right);t;\mathbf{x}\right]-R\left[\mathbf{h}^0;\mathbf{F}\left(\boldsymbol{\theta}^0\right);t;\mathbf{x}\right]$ in the NIE decoder's response.

The 1st-FASAM-IDE 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 [18] methodology. The fundamental concept for defining the sensitivity of an operator-valued quantity $R\left(\mathbf{e}\right)$ with respect to variations $\delta\mathbf{e}$ in a neighborhood around the nominal values $\mathbf{e}^0$, has been shown by Cacuci [18] to be provided by the 1st-order Gateaux- (G-) variation $\delta R\left(\mathbf{e}^0;\delta\mathbf{e}\right)$ of $R\left(\mathbf{e}\right)$, which is defined as follows:

$$\delta R\left(\mathbf{e}^0;\delta\mathbf{e}\right)\triangleq\left\{\frac{d}{d\varepsilon}\left[R\left(\mathbf{e}^0+\varepsilon\delta\mathbf{e}\right)\right]\right\}_{\varepsilon=0}\triangleq\lim_{\varepsilon\to 0}\frac{R\left(\mathbf{e}^0+\varepsilon\delta\mathbf{e}\right)-R\left(\mathbf{e}^0\right)}{\varepsilon}, \tag{6}$$

for a scalar $\varepsilon\in\mathsf{F}$ and for all (i.e., arbitrary) vectors $\delta\mathbf{e}$ in a neighborhood $\left(\mathbf{e}^0+\varepsilon\delta\mathbf{e}\right)$ around $\mathbf{e}^0$. The G-variation $\delta R\left(\mathbf{e}^0;\delta\mathbf{e}\right)$ is an operator defined on the same domain as $R\left(\mathbf{e}\right)$, has the same range as $R\left(\mathbf{e}\right)$, and satisfies the relation: $R\left(\mathbf{e}^0+\varepsilon\delta\mathbf{e}\right)-R\left(\mathbf{e}^0\right)=\delta R\left(\mathbf{e}^0;\delta\mathbf{e}\right)+\Delta\left(\delta\mathbf{e}\right)$, with

$\lim_{\varepsilon\to 0}\left[\Delta(\varepsilon\delta\mathbf{e})\right]/\varepsilon=0$. When the G-variation $\delta R\left(\mathbf{e}^0;\delta\mathbf{e}\right)$ is linear in the variation $\delta\mathbf{e}$, it can be written in the form $\delta R\left(\mathbf{e}^0;\delta\mathbf{e}\right)=\{\partial R/\partial\mathbf{e}\}_{\mathbf{e}^0}\,\delta\mathbf{e}$, where $\{\partial R/\partial\mathbf{e}\}_{\mathbf{e}^0}$ denotes the first-order G-derivative of $R(\mathbf{e})$ with respect to $\mathbf{e}$ evaluated at $\mathbf{e}^0$. For the particular case of the decoder response defined by Equation (2), the vector $\mathbf{e}$ has the following particular form: $\mathbf{e}\triangleq\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]^\dagger$.

Applying the definition provided in Equation (6) to Equation (2) yields the following expression for the first-order G-variation $\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)};\delta\mathbf{F}\right)$ of the response $R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]$:

$$\begin{aligned}\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)};\delta\mathbf{F}\right)&=\left\{\frac{d}{d\varepsilon}\int_{t_0}^{t_f}dt\int_\Omega D\left[\mathbf{h}^0(t;\mathbf{x})+\varepsilon\mathbf{v}^{(1)}(t;\mathbf{x});\mathbf{F}^0+\varepsilon\delta\mathbf{F};t;\mathbf{x}\right]d\mathbf{x}\right\}_{\varepsilon=0}\\&=\left\{\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\delta\mathbf{F}\right)\right\}_{dir}+\left\{\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)}\right)\right\}_{ind},\end{aligned}\tag{7}$$

where the following definitions were used:

$$\begin{aligned}\left\{\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\delta\mathbf{F}\right)\right\}_{dir}&\triangleq\int_{t_0}^{t_f}dt\int_\Omega d\mathbf{x}\left\{\frac{\partial D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]}{\partial\mathbf{F}}\delta\mathbf{F}\right\}_{\left(\mathbf{h}^0;\mathbf{F}^0\right)}\\&=\sum_{i=1}^{TF}\int_{t_0}^{t_f}dt\int_\Omega d\mathbf{x}\left\{\frac{\partial D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]}{\partial F_i}\delta F_i\right\}_{\left(\mathbf{h}^0;\mathbf{F}^0\right)},\end{aligned}\tag{8}$$

$$\begin{aligned}\left\{\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)}\right)\right\}_{ind}&\triangleq\int_{t_0}^{t_f}dt\int_\Omega d\mathbf{x}\left\{\frac{\partial D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]}{\partial\mathbf{h}}\mathbf{v}^{(1)}(t;\mathbf{x})\right\}_{\left(\mathbf{h}^0;\mathbf{F}^0\right)}\\&=\sum_{i=1}^{TH}\int_{t_0}^{t_f}d\tau\int_\Omega d\mathbf{x}\left\{\frac{\partial D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]}{\partial h_i}v_i^{(1)}(t;\mathbf{x})\right\}_{\left(\mathbf{h}^0;\mathbf{F}^0\right)}.\end{aligned}\tag{9}$$

The quantity $\left\{\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\delta\mathbf{F}\right)\right\}_{dir}$ defined in Equation (8) denotes the partial G-differential of the response $R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]$ with respect to the "feature function" $\mathbf{F}(\boldsymbol{\theta})\triangleq\left[F_1(\boldsymbol{\theta}),\ldots,F_{TF}(\boldsymbol{\theta})\right]^\dagger$, evaluated at the nominal values $\left(\mathbf{h}^0;\mathbf{F}^0\right)$. The quantity $\left\{\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\delta\mathbf{F}\right)\right\}_{dir}$ is called the "direct-effect term" because it arises directly from variations $\delta\mathbf{F}$, which in turn stem from parameter variations $\delta\boldsymbol{\theta}$, and can be computed directly using the nominal values $\left(\mathbf{h}^0;\mathbf{F}^0\right)$. The quantity $\left\{\delta R\left(\mathbf{h}^0;\mathbf{F}^0;\mathbf{v}^{(1)}\right)\right\}_{ind}$ defined in Equation(9) denotes the partial G-differential of the response $R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]$ with respect to the hidden state functions $\mathbf{h}(t)$, and is called the "indirect-effect term" because it arises indirectly, through the variations $\mathbf{v}^{(1)}(t)$ in the hidden state functions $\mathbf{h}(t)$. The indirect-effect term can be quantified only after having determined the variations $\mathbf{v}^{(1)}(t)$, which are caused by the variations $\delta\mathbf{F}$ through the NIE-net defined in Equation (1).

The first-order relationship between the variations $\mathbf{v}^{(1)}(t)$ and $\delta\mathbf{F}$ is obtained from the first-order G-variation of Equation (1), which is defined as follows:

$$\begin{aligned}\left\{\frac{d}{d\varepsilon}\left(\mathbf{h}^0+\varepsilon\mathbf{v}^{(1)}\right)\right\}_{\varepsilon=0}&=\left\{\frac{d}{d\varepsilon}\mathbf{g}\left[\mathbf{F}\left(\boldsymbol{\theta}^0\right)+\varepsilon\delta\mathbf{F};t;\mathbf{x}\right]\right\}_{\varepsilon=0}\\&+\left\{\frac{d}{d\varepsilon}\int_{t_0}^{t_f}d\tau\int_\Omega\mathbf{G}\left[\mathbf{h}^0(\tau;\mathbf{z})+\varepsilon\mathbf{v}^{(1)}(\tau;\mathbf{z});\mathbf{F}\left(\boldsymbol{\theta}^0\right)+\varepsilon\delta\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]d\mathbf{z}\right\}_{\varepsilon=0}.\end{aligned}\tag{10}$$

Carrying out the operations indicated in Equation (10) yields the following NIE-net of Fredholm-type, which will be called the "1st-Level Variational Sensitivity System" (1st-LVSS)", for the "1st-level variational function" $\mathbf{v}^{(1)}(t;\mathbf{x})$:

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

$$\mathbf{v}^{(1)}(t;\mathbf{x})=\left\{\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\frac{\partial\mathbf{G}\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial\mathbf{h}(\tau;\mathbf{z})}\mathbf{v}^{(1)}(\tau;\mathbf{z})\right\}_{\left(\mathbf{h}^0,\mathbf{F}^0\right)}+\left\{\mathbf{Q}^{(1)}(\mathbf{h};\mathbf{F})\delta\mathbf{F}\right\}_{\left(\mathbf{h}^0,\mathbf{F}^0\right)},\quad(11)$$

where:

$$\mathbf{Q}^{(1)}(\mathbf{h};\mathbf{F})\triangleq\begin{pmatrix}q_{11}^{(1)}(\mathbf{h};\mathbf{F}) & \cdots & q_{1,TF}^{(1)}(\mathbf{h};\mathbf{F})\\ \vdots & \ddots & \vdots\\ q_{TH,1}^{(1)}(\mathbf{h};\mathbf{F}) & \cdots & q_{TF,TF}^{(1)}(\mathbf{h};\mathbf{F})\end{pmatrix}\triangleq\frac{\partial\mathbf{g}(\mathbf{F};t;\mathbf{x})}{\partial\mathbf{F}}+\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{x}'\frac{\partial\mathbf{G}(\mathbf{h};\mathbf{F};t,\tau;\mathbf{x},\mathbf{x}')}{\partial\mathbf{F}}\quad(12)$$

Equivalently, the 1st-LVSS defined by Equation (11) can be written in component form as follows:

$$\begin{aligned}&v_i^{(1)}(t;\mathbf{x})-\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\frac{\partial G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_j(\tau;\mathbf{z})}v_j^{(1)}(\tau;\mathbf{z})\\&=\sum_{j=1}^{TF}q_{ij}^{(1)}(\mathbf{h};\mathbf{F})\delta F_j;\ i=1,\ldots,TH.\end{aligned}\quad(13)$$

where:

$$q_{ij}^{(1)}(\mathbf{h};\mathbf{F})\triangleq\frac{\partial g_i(\mathbf{F};t;\mathbf{x})}{\partial F_j}+\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\frac{\partial G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial F_j};\ i=1,\ldots,TH;\ j=1,\ldots,TF.\quad(14)$$

The 1st-LVSS is to be computed at the nominal/optimal values for the respective model parameters, but this fact had not been indicated explicitly in Equation (13) in order to simplify the notation.

It is important to note that the 1st-LVSS is linear in the variational function $\mathbf{v}^{(1)}(t;\mathbf{x})$, although it remains nonlinear in $\mathbf{h}(t;\mathbf{x})$, unless the original NIE's were quadratic in the function $\mathbf{h}(t;\mathbf{x})$. Note also that the 1st-LVSS would need to be solved anew for each variation δF_j, $j=1,\ldots,TF$, which could become prohibitively expensive computationally if TF is a large number.

The need for repeatedly solving the 1st-LVSS can be avoided if the indirect-effect term defined in Equation (9) could be expressed in terms of a "right-hand side" that does not involve the variational function $\mathbf{v}^{(1)}(t;\mathbf{x})$. This goal can be achieved by expressing the right-side of Equation (9) in terms of the solutions of the "1st-Level Adjoint Sensitivity System (1st-LASS)," to be constructed next, the construction of which requires the introduction of adjoint operators. Adjoint operators can be defined in Banach spaces but are most useful in Hilbert spaces. For the NIE considered in this work, the appropriate Hilbert space having elements of the form $\mathbf{v}^{(1)}(t;\mathbf{x})\in\mathsf{H}_1(\Omega_{tx})$ will be defined on the domain $\Omega_{tx}\triangleq\left[(t_0,t_f)\cup\Omega\right]$ and will be denoted as $\mathsf{H}_1(\Omega_{tx})$. The inner product in $\mathsf{H}_1(\Omega_{tx})$, which will be denoted as $\left\langle\boldsymbol{\chi}^{(1)}(t;\mathbf{x}),\boldsymbol{\eta}^{(1)}(t;\mathbf{x})\right\rangle_1$, where $\boldsymbol{\chi}^{(1)}(t;\mathbf{x})\triangleq\left[\chi_1^{(1)}(t;\mathbf{x}),\ldots,\chi_{TH}^{(1)}(t;\mathbf{x})\right]^\dagger\in\mathsf{H}_1(\Omega_{tx})$, $\boldsymbol{\eta}^{(1)}(t;\mathbf{x})\triangleq\left[\eta_1^{(1)}(t;\mathbf{x}),\ldots,\eta_{TH}^{(1)}(t;\mathbf{x})\right]^\dagger\in\mathsf{H}_1(\Omega_{tx})$, is defined as follows:

$$\left\langle\boldsymbol{\chi}^{(1)}(t;\mathbf{x}),\boldsymbol{\eta}^{(1)}(t;\mathbf{x})\right\rangle_1\triangleq\int_{t_0}^{t_f}dt\int_{\Omega}\left[\boldsymbol{\chi}^{(1)}(t;\mathbf{x})\right]^\dagger\boldsymbol{\eta}^{(1)}(t;\mathbf{x})d\mathbf{x}=\sum_{j=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}\chi_i^{(1)}(t;\mathbf{x})\eta_i^{(1)}(t;\mathbf{x})d\mathbf{x}.\quad(15)$$

The inner product $\left\langle\boldsymbol{\chi}^{(1)}(t;\mathbf{x}),\boldsymbol{\eta}^{(1)}(t;\mathbf{x})\right\rangle_1$ 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 Equation (11) with a vector $\mathbf{a}^{(1)}(t;\mathbf{x})\triangleq\left[a_1^{(1)}(t;\mathbf{x}),\ldots,a_{TH}^{(1)}(t;\mathbf{x})\right]^\dagger\in\mathsf{H}_1(\Omega_{tx})$, where the superscript "(1)" indicates "1st-Level", to obtain the following relationship:

$$\begin{aligned}&\left\langle\mathbf{a}^{(1)}(t;\mathbf{x}),\ \mathbf{v}^{(1)}(t;\mathbf{x})-\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\frac{\partial\mathbf{G}\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial\mathbf{h}(\tau;\mathbf{z})}\mathbf{v}^{(1)}(\tau;\mathbf{z})\right\rangle_1\\&=\left\langle\mathbf{a}^{(1)}(t;\mathbf{x}),\ \mathbf{Q}^{(1)}(\mathbf{h};\mathbf{F})\delta\mathbf{F}\right\rangle_1.\end{aligned}\quad(16)$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

Using the definition of the adjoint operator in $\mathsf{H}_1(\Omega_{tx})$, the left-side of Equation (16) will be transformed so as to transfer the operations on the function $\mathbf{v}^{(1)}(t;\mathbf{x})$ to operations on the function $\mathbf{a}^{(1)}(t;\mathbf{x})$. These operations will involve reversing the orders of summations and integrations for the second term in the inner-product on the left-side of Equation (16), as follows:

$$
\begin{aligned}
&\left\langle \mathbf{a}^{(1)}(t;\mathbf{x}), \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \frac{\partial \mathbf{G}\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial \mathbf{h}(\tau;\mathbf{z})} \mathbf{v}^{(1)}(\tau;\mathbf{z}) \right\rangle_1 \\
&\triangleq \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x} \left[\mathbf{a}^{(1)}(t;\mathbf{x})\right]^{\dagger} \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \frac{\partial \mathbf{G}\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial \mathbf{h}(\tau;\mathbf{z})} \mathbf{v}^{(1)}(\tau;\mathbf{z}) \\
&= \sum_{i=1}^{TH} \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x}) \sum_{j=1}^{TH} \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \frac{\partial G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_j(\tau;\mathbf{z})} v_j^{(1)}(\tau;\mathbf{z}) \\
&= \sum_{j=1}^{TH} \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z}\, v_j^{(1)}(\tau;\mathbf{z}) \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x} \sum_{i=1}^{TH} a_i^{(1)}(t;\mathbf{x}) \frac{\partial G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_j(\tau;\mathbf{z})} \\
&= \sum_{i=1}^{TH} \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z}\, v_i^{(1)}(\tau;\mathbf{z}) \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x} \sum_{m=1}^{TH} a_m^{(1)}(t;\mathbf{x}) \frac{\partial G_m\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_i(\tau;\mathbf{z})} \\
&= \sum_{i=1}^{TH} \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, v_i^{(1)}(t;\mathbf{x}) \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \sum_{m=1}^{TH} a_m^{(1)}(\tau;\mathbf{z}) \frac{\partial G_m\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})} \\
&= \left\langle \mathbf{v}^{(1)}(t;\mathbf{x}), \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \left[\frac{\partial \mathbf{G}\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial \mathbf{h}(t;\mathbf{x})}\right]^{\dagger} \mathbf{a}^{(1)}(\tau;\mathbf{z}) \right\rangle_1 .
\end{aligned}
\qquad (17)
$$

Using the result obtained in Equation (17) in Equation (16) yields the following relation:

$$
\begin{aligned}
&\left\langle \mathbf{a}^{(1)}(t;\mathbf{x}),\ \mathbf{v}^{(1)}(t;\mathbf{x}) - \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \frac{\partial \mathbf{G}\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial \mathbf{h}(\tau;\mathbf{z})} \mathbf{v}^{(1)}(\tau;\mathbf{z}) \right\rangle_1 \\
&= \left\langle \mathbf{v}^{(1)}(t;\mathbf{x}), \mathbf{a}^{(1)}(t;\mathbf{x}) - \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \left[\frac{\partial \mathbf{G}\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial \mathbf{h}(t;\mathbf{x})}\right]^{\dagger} \mathbf{a}^{(1)}(\tau;\mathbf{z}) \right\rangle_1 .
\end{aligned}
\qquad (18)
$$

The term on the right-side of Equation (18) is now required to represent the "indirect-effect" term defined in Equation (9), which is achieved by requiring that the function $\mathbf{a}^{(1)}(t;\mathbf{x})$ satisfy the following relation written in NIE-format:

$$
\mathbf{a}^{(1)}(t;\mathbf{x}) - \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \left[\frac{\partial \mathbf{G}\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial \mathbf{h}(t;\mathbf{x})}\right]^{\dagger} \mathbf{a}^{(1)}(\tau;\mathbf{z}) = \frac{\partial D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]}{\partial \mathbf{h}} . \qquad (19)
$$

The Fredholm-like system obtained in Equation (19) will be called the "1st-Level Adjoint Sensitivity System" and its solution, $\mathbf{a}^{(1)}(t;\mathbf{x})$, 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 $\mathbf{h}(t;\mathbf{x})$ but this fact has not been explicitly indicated in order to simplify the notation. The 1st-LASS is linear in $\mathbf{a}^{(1)}(t;\mathbf{x})$ but is, in general, nonlinear in $\mathbf{h}(t;\mathbf{x})$. Most importantly, the 1st-LASS is independent of any parameter variations and needs to be solved once only to determine the 1st-level adjoint sensitivity function $\mathbf{a}^{(1)}(t;\mathbf{x})$.

In component form, the 1st-LASS has the following structure:

$$
a_i^{(1)}(t;\mathbf{x}) = \sum_{j=1}^{TH} \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \frac{\partial G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})} a_j^{(1)}(\tau;\mathbf{z}) + \frac{\partial D\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})}; i = 1,...,TH. \qquad (20)
$$

It follows from Equations (16)–(19) that the indirect-effect term defined in Equation (9) has the following expression in terms of the 1st-level adjoint sensitivity function $\mathbf{a}^{(1)}(t;\mathbf{x})$:

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

$$\left\{ \delta R\left(\mathbf{h}^{0};\mathbf{F}^{0};\mathbf{a}^{(1)}\right)\right\}_{ind} = \left\{ \left\langle \mathbf{a}^{(1)}(t;\mathbf{x}),\, \mathbf{Q}^{(1)}(\mathbf{h};\mathbf{F})\,\delta\mathbf{F}\right\rangle_{1}\right\}_{\left(\mathbf{h}^{0};\mathbf{F}^{0}\right)}. \tag{21}$$

Using the results obtained in Equations (21) and (8) in Equation (7) 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 $\delta\mathbf{F}$:

$$\begin{aligned} &\delta R\left(\mathbf{h}^{0};\mathbf{F}^{0};\mathbf{v}^{(1)};\delta\mathbf{F}\right) \\ &= \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x} \left\{ \frac{\partial D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F}(\boldsymbol{\theta})\right]}{\partial\mathbf{F}} + \left[\mathbf{a}^{(1)}(t;\mathbf{x})\right]^{\dagger} \mathbf{Q}^{(1)}(\mathbf{h};\mathbf{F}) \right\} \delta\mathbf{F} \triangleq \sum_{j=1}^{TF} \frac{\partial R}{\partial F_j}\delta F_j . \end{aligned} \tag{22}$$

The expression on the right-side of Equation (22) is to be satisfied at the nominal/optimal values for the respective model parameters, but this fact has not been indicated explicitly in order to simplify the notation. It follows from Equation (22) that 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 $\mathbf{F}(\boldsymbol{\theta})$ have the following expressions:

$$\begin{aligned} \frac{\partial R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]}{\partial F_j} &= \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x} \frac{\partial D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]}{\partial F_j} + \sum_{i=1}^{TH} \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x}) \frac{\partial g_i(\mathbf{F};t;\mathbf{x})}{\partial F_j} \\ &+ \sum_{i=1}^{TH} \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x}) \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \frac{\partial G_i(\mathbf{h};\mathbf{F};t,\tau;\mathbf{x},\mathbf{z})}{\partial F_j}; \quad j = 1,\dots,TF. \end{aligned} \tag{23}$$

The sensitivities with respect to the primary model parameters can be obtained by using the result in Equation (23) together with the "chain rule" of differentiating compound functions, as follows:

$$\frac{\partial R}{\partial \theta_j} = \sum_{i=1}^{TF} \frac{\partial R}{\partial F_i}\frac{\partial F_i}{\partial \theta_j}, \quad j = 1,\dots,TW. \tag{24}$$

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

When there are no feature functions but only model parameters, that is when $F_i(\boldsymbol{\theta}) \equiv \theta_i$ for all $i = 1,\dots,TF \triangleq TW$, the expression obtained in Equation (23) yields directly the first-order sensitivities $\partial R/\partial \theta_j$, for all $j = 1,\dots,TW$, taking on the following specific form:

$$\begin{aligned} \frac{\partial R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]}{\partial \theta_j} &= \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x} \frac{\partial D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]}{\partial \theta_j} + \sum_{i=1}^{TH} \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x}) \frac{\partial g_i(\mathbf{F};t;\mathbf{x})}{\partial \theta_j} \\ &+ \sum_{i=1}^{TH} \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x}) \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \frac{\partial G_i(\mathbf{h};\mathbf{F};t,\tau;\mathbf{x},\mathbf{z})}{\partial \theta_j}; \quad j = 1,\dots,TW. \end{aligned} \tag{25}$$

In Equation (25), the 1st-level adjoint sensitivity function $\mathbf{a}^{(1)}(t;\mathbf{x}) \triangleq \left[a_1^{(1)}(t;\mathbf{x}),\dots,a_{TH}^{(1)}(t;\mathbf{x})\right]^{\dagger}$ is still the solution of the 1st-LASS defined by Equation (20), since the 1st-LASS is independent of any parameter variations. In this case, however, all of the sensitivities $\partial R/\partial \theta_j$, for all $j = 1,\dots,TW$ would be obtained by computing integrals (using quadrature formulas). In contradistinction, when features of parameters can be established so that the sensitivities can be computed by using Equation (23) rather than Equation (25), only TF $(TF < TW)$ integrals would need to be computed (using quadrature formulas) to obtain the $\partial R/\partial F_j$, $j = 1,\dots,TF$; the sensitivities with respect to the model parameters would subsequently be obtained analytically using the chain-rule provided in Equation (24). Thus, when there are no feature functions of parameters, the 1st-FASAM-NODE reduces to the First-Order Comprehensive Adjoint Sensitivity Analysis Methodology [18] applied to Neural Integral Equations of Fredholm-Type (1st-CASAM-NIE-Fredholm).

3. Illustrative Application of the 1st-FASAM-NIE-Fredholm Methodology

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

The application of the 1st-FASAM-NIE-Fredholm Methodology will be illustrated in this Section by considering a system of nonlinear Fredholm-type NIE which admits analytical solutions, thereby facilitating the verification of the expressions that will be obtained for the first-order sensitivities of NIE-decoder responses with respect to the model parameters (weights) that characterize the NIE-net. The NIE-net comprises two uncoupled nonlinear Fredholm-type equations which will be coupled through the nonlinear decoder-response. The reason for choosing such a system is to show that, even though the initial equations(each describing an independent NIE-net) are uncoupled, their coupling through the decoder-response will necessitate solving coupled equations for determining the second (and higher-) order sensitivities. The illustrative NIE-model for the hidden/latent function $\mathbf{y} \triangleq \left[y_1(t), y_2(t)\right]^\dagger$ comprises the following uncoupled nonlinear Fredholm-type NIE:

$$y_1(t) = \frac{t^{\lambda_1}}{F_1(\boldsymbol{\alpha})} \int_a^b y_1^2(\tau)\, d\tau; \tag{26}$$

$$y_2(t) = \frac{t^{\lambda_2}}{F_2(\boldsymbol{\beta})} \int_a^b y_2^2(\tau)\, d\tau . \tag{27}$$

The NIE-decoder response will be considered to be provided by the following nonlinear functional of $\mathbf{y} \triangleq \left[y_1(t), y_2(t)\right]^\dagger$:

$$R\left[\mathbf{y};\mathbf{F}(\boldsymbol{\theta})\right] = F_3(\boldsymbol{\gamma}) \int_a^b y_1(t)\, y_2(t)\, dt . \tag{28}$$

In Equations (26)–(28), the scalar parameters $a, b, \lambda_1, \lambda_2$ are considered to be known exactly while the components of the model parameters $\boldsymbol{\alpha} \triangleq \left[\alpha_1,...,\alpha_{TA}\right]^\dagger$, $\boldsymbol{\beta} \triangleq \left[\beta_1,...,\beta_{TB}\right]^\dagger$, and $\boldsymbol{\gamma} \triangleq \left[\gamma_1,...,\gamma_{TG}\right]^\dagger$ are considered to be subject to uncertainties. The feature function $\mathbf{F}(\boldsymbol{\theta})$ is considered to comprise three components defined as follows:

$$\mathbf{F}(\boldsymbol{\theta}) = \left[F_1(\boldsymbol{\alpha}), F_2(\boldsymbol{\beta}), F_3(\boldsymbol{\gamma})\right]^\dagger;\ \boldsymbol{\theta} \triangleq \left[\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\right]^\dagger \triangleq \left[\theta_1,...,\theta_{TW}\right]^\dagger;\ TW \triangleq TA+TB+TG. \tag{29}$$

For subsequent verification purposes, the non-trivial exact analytical solutions of the Fredholm-type NIE defined by Equations (26) and (27) are provided below:

$$y_1(t) = c_1 F_1(\boldsymbol{\alpha}) t^{\lambda_1};\ y_2(t) = c_2 F_2(\boldsymbol{\beta}) t^{\lambda_2};\ c_i \triangleq \frac{2\lambda_i + 1}{b^{2\lambda_i+1} - a^{2\lambda_i+1}};\ i = 1,2. \tag{30}$$

Using in Equation (28) the results provided in Equation (30) yields the following closed-form expression for the response $R\left[\mathbf{y};\mathbf{F}(\boldsymbol{\theta})\right]$:

$$R\left[\mathbf{y};\mathbf{F}(\boldsymbol{\theta})\right] = c_1 c_2 k F_1(\boldsymbol{\alpha}) F_2(\boldsymbol{\beta}) F_3(\boldsymbol{\gamma});\ k \triangleq \frac{b^{\lambda_1+\lambda_2+1} - a^{\lambda_1+\lambda_2+1}}{\lambda_1+\lambda_2+1}. \tag{31}$$

The nominal/optimal parameter values, obtained after completing the training of the INE-net, will be denoted, as before, by using the superscript "zero", namely:

$$\boldsymbol{\alpha}^0 \triangleq \left[\alpha_1^0,...,\alpha_{TA}^0\right]^\dagger;\ \boldsymbol{\beta}^0 \triangleq \left[\beta_1^0,...,\beta_{TB}^0\right]^\dagger;\ \boldsymbol{\gamma}^0 \triangleq \left[\gamma_1^0,...,\gamma_{TG}^0\right]^\dagger; \tag{32}$$

$$\boldsymbol{\theta}^0 \triangleq \left[\boldsymbol{\alpha}^0,\boldsymbol{\beta}^0,\boldsymbol{\gamma}^0\right]^\dagger \triangleq \left[\theta_1^0,...,\theta_{TT}^0\right]^\dagger;\ TT \triangleq TA+TB+TG; \tag{33}$$

$$\mathbf{F}^0(\boldsymbol{\theta}) \triangleq \mathbf{F}(\boldsymbol{\theta}^0) \triangleq \left[F_1^0, F_2^0, F_3^0\right]^\dagger \triangleq \left[F_1(\boldsymbol{\alpha}^0), F_2(\boldsymbol{\beta}^0), F_3(\boldsymbol{\gamma}^0)\right]^\dagger . \tag{34}$$

Similarly, the value of the function $\mathbf{y} \triangleq \left[y_1(t), y_2(t)\right]^\dagger$ computed using the nominal/optimal parameter values will also be denoted using the superscript "zero", namely: $\mathbf{y}^0 \triangleq \left[y_1^0(t), y_2^0(t)\right]^\dagger$.

The first-order sensitivities of the response $R\left[\mathbf{y};\mathbf{F}(\boldsymbol{\theta})\right]$ with respect to the components of the feature function $\mathbf{F}(\boldsymbol{\theta})$ are obtained by applying the principle underlying Equation (7) to the expression provided in Equation (28), which yields the following expression:

$$\delta R\left(\mathbf{y}^0;F_3^0;\delta\mathbf{y};\delta F_3\right)\triangleq\left\{\frac{d}{d\varepsilon}\left(F_3^0+\varepsilon\delta F_3\right)\int_a^b\left(y_1^0+\varepsilon\delta y_1\right)\left(y_2^0+\varepsilon\delta y_2\right)dt\right\}_{\varepsilon=0} \tag{35}$$
$$\triangleq\left\{\delta R\left(\mathbf{y}^0;\delta F_3\right)\right\}_{dir}+\left\{\delta R\left(\mathbf{y}^0;F_3^0;\delta\mathbf{y}\right)\right\}_{ind},$$

where the following definitions were used:

$$\left\{\delta R\left(\mathbf{y}^0;\delta F_3\right)\right\}_{dir}\triangleq\left(\delta F_3\right)\left\{\int_a^b y_1(t)\,y_2(t)\,dt\right\}_{\left(\mathbf{y}^0;\,\mathbf{F}^0\right)}; \tag{36}$$

$$\left\{\delta R\left(\mathbf{y}^0;F_3^0;\delta\mathbf{y}\right)\right\}_{ind}\triangleq\left\{F_3(\boldsymbol{\gamma})\int_a^b\left[y_2(t)\,\delta y_1(t)+y_1(t)\,\delta y_2(t)\right]dt\right\}_{\left(\mathbf{y}^0;\mathbf{F}^0\right)}. \tag{37}$$

The "direct-effect term" $\left\{\delta R\left(\mathbf{y}^0;\delta F_3\right)\right\}_{dir}$ can be computed directly for every variation δF_3 at the nominal/optimal values $\left(\mathbf{y}^0;\mathbf{F}^0\right)$. The "indirect-effect term" $\left\{\delta R\left(\mathbf{y}^0;F_3^0;\delta\mathbf{y}\right)\right\}_{ind}$ can be quantified only after having determined the variations $\delta\mathbf{y}\triangleq\left[\delta y_1,\delta y_2\right]^{\dagger}$. The first-order relationship between the variations $\delta\mathbf{y}\triangleq\left[\delta y_1,\delta y_2\right]^{\dagger}$ and $\delta\mathbf{F}$ is obtained from the first-order G-variations of Equations (26) and (27), which yields the following relations:

$$\delta y_1(t)=\frac{2t^{\lambda_1}}{F_1(\boldsymbol{\alpha})}\int_a^b y_1(\tau)\,\delta y_1(\tau)\,d\tau-\left[\frac{t^{\lambda_1}}{F_1^2(\boldsymbol{\alpha})}\int_a^b y_1^2(\tau)\,d\tau\right]\delta F_1(\boldsymbol{\alpha}) \tag{38}$$

$$\delta y_2(t)=\frac{2t^{\lambda_2}}{F_2(\boldsymbol{\beta})}\int_a^b y_2(\tau)\,\delta y_2(\tau)\,d\tau-\left[\frac{t^{\lambda_2}}{F_2^2(\boldsymbol{\beta})}\int_a^b y_2^2(\tau)\,d\tau\right]\delta F_2(\boldsymbol{\beta}). \tag{39}$$

Using in Equations (38) and (39) the expressions obtained in Equation (30) for $y_1(t)$ and $y_2(t)$ yields the following relations:

$$\delta y_1(t)=2c_1t^{\lambda_1}\int_a^b\tau^{\lambda_1}\,\delta y_1(\tau)\,d\tau-c_1t^{\lambda_1}\,\delta F_1(\boldsymbol{\alpha}); \tag{40}$$

$$\delta y_2(t)=2c_2t^{\lambda_2}\int_a^b\tau^{\lambda_2}\,\delta y_2(\tau)\,d\tau-c_2t^{\lambda_2}\,\delta F_2(\boldsymbol{\beta}). \tag{41}$$

The Fredholm-type equations obtained Equations (40) and (41) constitute the 1st-Level Variational Sensitivity System (1st-LVSS) for the variational function $\delta\mathbf{y}\triangleq\left[\delta y_1,\delta y_2\right]^{\dagger}$. The 1st-LVSS is linear in variational function $\delta\mathbf{y}\triangleq\left[\delta y_1,\delta y_2\right]^{\dagger}$ and is to be satisfied at the nominal/optimal values for the respective model parameters, but this fact had not been indicated explicitly in order to simplify the notation. The 1st-LVSS defined by Equations (40) and (41) would need to be solved anew for each parameter variation $\delta\alpha_i,\ i=1,\dots,TA$, and $\delta\beta_i;\ i=1,\dots,TB$. This need for repeatedly solving the 1st-LVSS can be avoided if the indirect-effect term defined in Equation (37) could be expressed in terms of a "right-hand side" that does not involve the variational function $\delta\mathbf{y}\triangleq\left[\delta y_1,\delta y_2\right]^{\dagger}$. This goal can be achieved by expressing the right-side of Equation (37) in terms of the solutions of a "1st-Level Adjoint Sensitivity System (1st-LASS)" which will be constructed next by using adjoint operators defined in a Hilbert space denoted as $\mathsf{H}_1(a,b)$ that comprises as elements two-component vectors having the same structure as $\delta\mathbf{y}\triangleq\left[\delta y_1,\delta y_2\right]^{\dagger}$. The inner product between two vectors $\boldsymbol{\chi}^{(1)}(t)\triangleq\left[\chi_1^{(1)}(t),\chi_2^{(1)}(t)\right]^{\dagger}\in\mathsf{H}_1(a,b)$ and $\boldsymbol{\eta}^{(1)}(t)\triangleq\left[\eta_1^{(1)}(t),\eta_2^{(1)}(t)\right]^{\dagger}\in\mathsf{H}_1(a,b)$ will be denoted as $\left\langle\boldsymbol{\chi}^{(1)}(t),\boldsymbol{\eta}^{(1)}(t)\right\rangle_1$ and is defined as follows:

$$\left\langle\boldsymbol{\chi}^{(1)}(t),\boldsymbol{\eta}^{(1)}(t)\right\rangle_1\triangleq\int_a^b\left[\boldsymbol{\chi}^{(1)}(t)\right]^{\dagger}\boldsymbol{\eta}^{(1)}(t)\,dt=\sum_{j=1}^{2}\int_a^b\chi_i^{(1)}(t)\,\eta_i^{(1)}(t)\,dt. \tag{42}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

The next step is to form the inner product of Equations (40) and (41) with a vector $\mathbf{a}^{(1)}(t) \triangleq \left[a_1^{(1)}(t), a_2^{(1)}(t)\right]^\dagger \in \mathsf{H}_1(a,b)$, where the superscript "(1)" indicates "1st-Level", to obtain the following relationship:

$$\begin{aligned}
&\int_a^b a_1^{(1)}(t)\delta y_1(t)\,dt - 2c_1\int_a^b a_1^{(1)}(t)t^{\lambda_1}dt\int_a^b \tau^{\lambda_1}\delta y_1(\tau)\,d\tau + \int_a^b a_2^{(1)}(t)\delta y_2(t)\,dt \\
&-2c_2\int_a^b a_2^{(1)}(t)t^{\lambda_2}dt\int_a^b \tau^{\lambda_2}\delta y_2(\tau)\,d\tau = -\left[c_1\int_a^b a_1^{(1)}(t)t^{\lambda_1}dt\right]\delta F_1(\boldsymbol{\alpha}) - \left[c_2\int_a^b a_2^{(1)}(t)t^{\lambda_2}dt\right]\delta F_2(\boldsymbol{\beta}).
\end{aligned} \tag{43}$$

Next, the left-side of Equation (43) is transformed as follows:

$$\begin{aligned}
&\int_a^b a_1^{(1)}(t)\delta y_1(t)\,dt - 2c_1\int_a^b a_1^{(1)}(t)t^{\lambda_1}dt\int_a^b \tau^{\lambda_1}\delta y_1(\tau)\,d\tau + \int_a^b a_2^{(1)}(t)\delta y_2(t)\,dt \\
&-2c_2\int_a^b a_2^{(1)}(t)t^{\lambda_2}dt\int_a^b \tau^{\lambda_2}\delta y_2(\tau)\,d\tau = \int_a^b\left[a_1^{(1)}(t) - 2c_1t^{\lambda_1}\int_a^b a_1^{(1)}(\tau)\tau^{\lambda_1}d\tau\right]\delta y_1(t)\,dt \\
&+\int_a^b\left[a_2^{(1)}(t) - 2c_2t^{\lambda_2}\int_a^b a_2^{(1)}(\tau)\tau^{\lambda_2}d\tau\right]\delta y_2(t)\,dt.
\end{aligned} \tag{44}$$

The terms on the right-side of Equation (44) are now required to represent the "indirect-effect" term defined in Equation (37), which is achieved by requiring that the function $\mathbf{a}^{(1)}(t) \triangleq \left[a_1^{(1)}(t), a_2^{(1)}(t)\right]^\dagger$ satisfy the following relation:

$$a_1^{(1)}(t) - 2c_1t^{\lambda_1}\int_a^b \tau^{\lambda_1}a_1^{(1)}(\tau)\,d\tau = F_3(\boldsymbol{\gamma})y_2(t); \tag{45}$$

$$a_2^{(1)}(t) - 2c_2t^{\lambda_2}\int_a^b \tau^{\lambda_2}a_2^{(1)}(\tau)\,d\tau = F_3(\boldsymbol{\gamma})y_1(t). \tag{46}$$

The Fredholm-like system obtained in Equations (45) and (46) is called the "1st-Level Adjoint Sensitivity System" and its solution, $\mathbf{a}^{(1)}(t) \triangleq \left[a_1^{(1)}(t), a_2^{(1)}(t)\right]^\dagger$, is called the "1st-level adjoint sensitivity function." The 1st-LASS is to be satisfied at the nominal/optimal values for the respective model parameters, but this fact had not been indicated explicitly in order to simplify the notation. The 1st-LASS is linear in $\mathbf{a}^{(1)}(t)$ and is independent of all parameter variations, so it needs to be solved just once in order to determine $\mathbf{a}^{(1)}(t)$.

For subsequent verification purposes, the exact closed-form expression for the components of $\mathbf{a}^{(1)}(t) \triangleq \left[a_1^{(1)}(t), a_2^{(1)}(t)\right]^\dagger$ obtained by solving the above 1st-LASS are provided below:

$$a_1^{(1)}(t) = At^{\lambda_1} + Bt^{\lambda_2};\; A \triangleq -2c_1c_2kF_2(\boldsymbol{\beta})F_3(\boldsymbol{\gamma});\; B \triangleq c_2F_2(\boldsymbol{\beta})F_3(\boldsymbol{\gamma}); \tag{47}$$

$$a_2^{(1)}(t) = Ct^{\lambda_1} + Dt^{\lambda_2};\; C \triangleq c_1F_1(\boldsymbol{\alpha})F_3(\boldsymbol{\gamma});\; D \triangleq -2c_1c_2kF_1(\boldsymbol{\alpha})F_3(\boldsymbol{\gamma}). \tag{48}$$

Using Equations (43)–(46) yields the following expression for the indirect-effect term defined in Equation (37) in terms of the 1st-level adjoint sensitivity function $\mathbf{a}^{(1)}(t)$:

$$\left\{\delta R\left(\mathbf{y}^0; F_3^0; \delta\mathbf{y}\right)\right\}_{ind} = -\left[c_1\int_a^b a_1^{(1)}(t)t^{\lambda_1}dt\right]\delta F_1(\boldsymbol{\alpha}) - \left[c_2\int_a^b a_2^{(1)}(t)t^{\lambda_2}dt\right]\delta F_2(\boldsymbol{\beta}). \tag{49}$$

Adding the result obtained in Equation (49) to the expression given in Equation (36) for the direct-effect term yields the following expression for the G-variation $\delta R\left(y_1^0, y_2^0; F_3^0; \delta y_1, \delta y_2; \delta F_3\right)$:

$$\delta R = -\left[c_1\int_a^b a_1^{(1)}(t)t^{\lambda_1}dt\right]\delta F_1(\boldsymbol{\alpha}) - \left[c_2\int_a^b a_2^{(1)}(t)t^{\lambda_2}dt\right]\delta F_2(\boldsymbol{\beta}) + \left[\int_a^b y_1(t)y_2(t)\,dt\right]\delta F_3(\boldsymbol{\gamma}). \tag{50}$$

The expression on the right-side of Equation (50) is to be satisfied at the nominal/optimal values for the respective model parameters, but this fact had not been indicated explicitly in order to simplify the notation. It follows from Equation (50) that 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 $\mathbf{F}(\boldsymbol{\theta})$ have the following expressions:

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

$$\frac{\partial R}{\partial F_1} = -c_1 \int_a^b a_1^{(1)}(t) t^{\lambda_1} dt; \tag{51}$$

$$\frac{\partial R}{\partial F_2} = -c_2 \int_a^b a_2^{(1)}(t) t^{\lambda_2} dt; \tag{52}$$

$$\frac{\partial R}{\partial F_3} = \int_a^b y_1(t) y_2(t) dt. \tag{53}$$

The expressions of the sensitivities obtained in Equations(51)–(53) are to be evaluated at the nominal/optimal values for the respective model parameters, but this fact had not been indicated explicitly in order to simplify the notation.

Inserting into Equations (51)–(53) the expressions obtained for $\mathbf{y} \triangleq [y_1(t), y_2(t)]^\dagger$ in Equation (30), and the expressions obtained for $\mathbf{a}^{(1)}(t) \triangleq [a_1^{(1)}(t), a_2^{(1)}(t)]^\dagger$ in Equations (47) and (48) yields the following explicit expressions:

$$\frac{\partial R}{\partial F_1} = -c_1 \int_a^b \left(A t^{\lambda_1} + B t^{\lambda_2} \right) t^{\lambda_1} dt = c_1 c_2 k F_2(\boldsymbol{\beta}) F_3(\boldsymbol{\gamma}); \tag{54}$$

$$\frac{\partial R}{\partial F_2} = -c_2 \int_a^b a_2^{(1)}(t) t^{\lambda_2} dt = c_1 c_2 k F_1(\boldsymbol{\alpha}) F_3(\boldsymbol{\gamma}); \tag{55}$$

$$\frac{\partial R}{\partial F_3} = \int_a^b y_1(t) y_2(t) dt = c_1 c_2 k F_1(\boldsymbol{\alpha}) F_2(\boldsymbol{\beta}) \tag{56}$$

The validity of the expressions obtained in Equations (54)–(56) can be readily verified by using the expression of the decoder-response provided in Equation (31).

The sensitivities with respect to the primary model parameters can be obtained by using the results obtained in Equations(51)–(53) for the sensitivities with respect to the components of the feature function, together with the "chain rule" of differentiating compound functions, as shown in Equation (24).

3.1. Absence of Feature Functions: Comparison to 1st-CASAM-NIE-Fredholm

As has been discussed in Subsection 2.1, in the absence of feature functions of parameters, the 1st-FASAM-NIE-Fredholm reduces to the First-Order Comprehensive Adjoint Sensitivity Analysis Methodology [18] applied to Neural Integral Equations of Fredholm-Type (1st-CASAM-NIE-Fredholm). The first-order sensitivities of the NIE-decoder response with respect to the model's parameters are obtained by following the same steps as described in Equations (35)–(50), but with the following replacements: $\delta F_1(\boldsymbol{\alpha}) = \sum_{i=1}^{TA} \left[\partial F_1(\boldsymbol{\alpha}) / \partial \alpha_i \right] \delta \alpha_i$, $\delta F_2(\boldsymbol{\beta}) = \sum_{i=1}^{TB} \left[\partial F_2(\boldsymbol{\beta}) / \partial \beta_i \right] \delta \beta_i$, $\delta F_3(\boldsymbol{\gamma}) = \sum_{i=1}^{TB} \left[\partial F_3(\boldsymbol{\gamma}) / \partial \gamma_i \right] \delta \gamma_i$. With these replacements, Equation (50) takes on the following form:

$$\begin{aligned} \delta R = & -\left[c_1 \int_a^b a_1^{(1)}(t) t^{\lambda_1} dt \right] \sum_{i=1}^{TA} \left[\partial F_1(\boldsymbol{\alpha}) / \partial \alpha_i \right] \delta \alpha_i - \left[c_2 \int_a^b a_2^{(1)}(t) t^{\lambda_2} dt \right] \\ & \times \sum_{i=1}^{TB} \left[\partial F_2(\boldsymbol{\beta}) / \partial \beta_i \right] \delta \beta_i + \left[\int_a^b y_1(t) y_2(t) dt \right] \sum_{i=1}^{TG} \left[\partial F_3(\boldsymbol{\gamma}) / \partial \gamma_i \right] \delta \gamma_i . \end{aligned} \tag{57}$$

Similarly, Equations (51)–(53) will take on the following forms:

$$\frac{\partial R}{\partial \alpha_i} = -c_1 \frac{\partial F_1(\boldsymbol{\alpha})}{\partial \alpha_i} \int_a^b a_1^{(1)}(t) t^{\lambda_1} dt; \quad i = 1, ..., TA; \tag{58}$$

$$\frac{\partial R}{\partial \beta_i} = -c_2 \frac{\partial F_2(\boldsymbol{\beta})}{\partial \beta_i} \int_a^b a_2^{(1)}(t) t^{\lambda_2} dt; \quad i = 1, ..., TB; \tag{59}$$

$$\frac{\partial R}{\partial \gamma_i} = \frac{\partial F_3(\boldsymbol{\gamma})}{\partial \gamma_i} \int_a^b y_1(t) y_2(t) dt; \quad i = 1, ..., TG. \tag{60}$$

The 1st-level adjoint sensitivity function $\mathbf{a}^{(1)}(t) \triangleq \left[a_1^{(1)}(t), a_2^{(1)}(t)\right]^\dagger$ that appears in Equations (58)–(60) is the same as that which appears in Equations (51)–(53), since the 1st-LASS defined by Equations (45) and (46) is independent of any parameter variations. Comparing the results for the first-order sensitivities computed by applying the 1st-FASAM-NIE-Fredholm, if feature functions of parameters can be identified, with the application of the 1st-CASAM-NIE-Fredholm, to compute the sensitivities directly with respect to the model parameters in the absence of feature-functions reveals the following conclusions:

(i) Applying the 1st-FASAM-NIE-Fredholm requires the computations of as many integrals (which define the respective first-order sensitivities) as there are components of feature function (in the above example, there are 3 components); the sensitivities with respect to the model parameters are computed analytically, by applying the "chain-rule of differentiation" using the first-order sensitivities with respect to the components of the feature functions.

(ii) In the absence of feature functions of parameters, the application of the 1st-CASAM-NIE-Fredholm requires the computations of as many integrals (which define the respective first-order sensitivities) as there are parameters (in the above example, $TA+TB+TG$ integrals to be computed numerically using quadrature formulas).

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

The second-order sensitivities of the response $R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]$ defined in Equation (2) are obtained by computing the first-order sensitivities of the expression obtained in Equation (23), which represents the first-order sensitivities of the response $R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]$ with respect to the components of the feature function $\mathbf{F}(\boldsymbol{\theta}) \triangleq \left[F_1(\boldsymbol{\theta}),...,F_{TF}(\boldsymbol{\theta})\right]^\dagger$. In other words, the second-order sensitivities of $R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]$ will be computed by conceptually using their basic definitions as being the "first-order sensitivities of the first-order sensitivities."

4.1. Second-Order Sensitivities Stemming from $\partial R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]/\partial F_j$, $j=1,...,TF$.

The second-order sensitivities stemming from the first-order sensitivities $\partial R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]/\partial F_j$ are obtained from the first-order G-differential of Equation (23), for $j=1,...,TF$, as follows:

$$\begin{aligned}
\delta\left(\frac{\partial R}{\partial F_j}\right) \triangleq & \left\{\frac{d}{d\varepsilon}\left[\int_{t_0}^{t_f} dt \int_\Omega d\mathbf{x} \frac{\partial}{\partial F_j} D\left[\mathbf{h}^0+\varepsilon\mathbf{v}^{(1)};\mathbf{F}(\boldsymbol{\theta}^0)+\varepsilon\delta\mathbf{F};t;\mathbf{x}\right]\right]\right\}_{\varepsilon=0} \\
& +\left\{\sum_{i=1}^{TH}\frac{d}{d\varepsilon}\int_{t_0}^{t_f} dt \int_\Omega d\mathbf{x}\left[a_i^{(1,0)}(t;\mathbf{x})+\varepsilon\delta a_i^{(1)}(t;\mathbf{x})\right]\frac{\partial g_i\left(\mathbf{F}^0+\varepsilon\delta\mathbf{F};t;\mathbf{x}\right)}{\partial F_j}\right\}_{\varepsilon=0} \\
& +\left\{\sum_{i=1}^{TH}\frac{d}{d\varepsilon}\int_{t_0}^{t_f} dt \int_\Omega d\mathbf{x}\left(a_i^{(1,0)}+\varepsilon\delta a_i^{(1)}\right)\int_{t_0}^{t_f} d\tau \int_\Omega d\mathbf{z}\frac{\partial G_i\left(\mathbf{h}^0+\varepsilon\mathbf{v}^{(1)};\mathbf{F}^0+\varepsilon\delta\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right)}{\partial F_j}\right\}_{\varepsilon=0} \\
& \triangleq \delta\left(\frac{\partial R}{\partial F_j}\right)_{dir}+\delta\left(\frac{\partial R}{\partial F_j}\right)_{ind}; \quad j=1,...,TF.
\end{aligned} \tag{61}$$

In Equation (61), the expression of the direct-effect term $\delta\left(\partial R/\partial F_j\right)_{dir}$ 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:

$$
\begin{aligned}
\delta\left(\frac{\partial R}{\partial F_j}\right)_{dir} &\triangleq \int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x} \frac{\partial D\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]}{\partial \mathbf{F}\partial F_j}\delta\mathbf{F} + \sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x})\frac{\partial^2 g_i(\mathbf{h};\mathbf{F})}{\partial \mathbf{F}\partial F_j}\delta\mathbf{F} \\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x})\int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z}\frac{\partial^2 G_i(\mathbf{h};\mathbf{F};t,\tau;\mathbf{x},\mathbf{z})}{\partial \mathbf{F}\partial F_j}\delta\mathbf{F} \\
&=\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\sum_{k=1}^{TF}\frac{\partial D\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]}{\partial F_k\partial F_j}\delta F_k + \sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x})\sum_{k=1}^{TF}\frac{\partial^2 g_i(\mathbf{h};\mathbf{F})}{\partial F_k\partial F_j}\delta F_k \\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x})\int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z}\sum_{k=1}^{TF}\frac{\partial^2 G_i(\mathbf{h};\mathbf{F};t,\tau;\mathbf{x},\mathbf{z})}{\partial F_k\partial F_j}\delta F_k.
\end{aligned}
\tag{62}
$$

The expression of the indirect-effect term $\delta\left(\partial R/\partial F_j\right)_{ind}$ in Equation (61) is obtained after performing the operations with respect to the scalar ε and comprises the variations $\mathbf{v}^{(1)}(t;\mathbf{x})$ and $\delta\mathbf{a}^{(1)}(t;\mathbf{x}) \triangleq \left[\delta a_1^{(1)}(t;\mathbf{x}),\ldots,\delta a_{TH}^{(1)}(t;\mathbf{x})\right]^{\dagger}$, being defined as follows:

$$
\begin{aligned}
\delta\left(\frac{\partial R}{\partial F_j}\right)_{ind} &\triangleq \sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\frac{\partial^2 D\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})\partial F_j}v_i^{(1)}(t;\mathbf{x}) \\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x})\int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z}\sum_{m=1}^{TH}\frac{\partial^2 G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_m(\tau;\mathbf{z})\partial F_j}v_m^{(1)}(\tau;\mathbf{z}) \\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\,\delta a_i^{(1)}(t;\mathbf{x})\left\{\frac{\partial g_i(\mathbf{F};t;\mathbf{x})}{\partial F_j}+\int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z}\frac{\partial G_i(\mathbf{h};\mathbf{F};t,\tau;\mathbf{x},\mathbf{z})}{\partial F_j}\right\} \\
&=\sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, v_i^{(1)}(t;\mathbf{x})\left\{\frac{\partial^2 D\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})\partial F_j}\right. \\
&\left.+\sum_{m=1}^{TH}\int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z}\, a_m^{(1)}(\tau;\mathbf{z})\frac{\partial^2 G_m\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})\partial F_j}\right\} \\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\,\delta a_i^{(1)}(t;\mathbf{x})\left\{\frac{\partial g_i(\mathbf{F};t;\mathbf{x})}{\partial F_j}+\int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z}\frac{\partial G_i(\mathbf{h};\mathbf{F};t,\tau;\mathbf{x},\mathbf{z})}{\partial F_j}\right\}.
\end{aligned}
\tag{63}
$$

The expressions in Equations (62) and (63) 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 $\mathbf{v}^{(1)}(t;\mathbf{x})$ and $\delta\mathbf{a}^{(1)}(t;\mathbf{x})$. The variation $\mathbf{v}^{(1)}(t;\mathbf{x})$ is the solution of the 1st-LVSS defined by Equation (11). On the other hand, the variational function $\delta\mathbf{a}^{(1)}(t;\mathbf{x})$ is the solution of the system of equations obtained by G-differentiating the 1st-LASS. By definition, the G-differential of Equation (20) is obtained as follows, for $i=1,\ldots,TH$:

$$
\begin{aligned}
\left\{\frac{d}{d\varepsilon}\left[a_i^{(1,0)}(t;\mathbf{x})+\varepsilon\delta a_i^{(1)}(t;\mathbf{x})\right]\right\}_{\varepsilon=0} &= \left\{\frac{d}{d\varepsilon}\frac{\partial D\left[\mathbf{h}^0(t;\mathbf{x})+\varepsilon\mathbf{v}^{(1)}(t;\mathbf{x});\mathbf{F}^0+\varepsilon\delta\mathbf{F};t;\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})}\right\}_{\varepsilon=0} \\
&+\left\{\frac{d}{d\varepsilon}\sum_{j=1}^{TH}\int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z}\frac{\partial G_j\left(\mathbf{h}^0+\varepsilon\mathbf{v}^{(1)};\mathbf{F}^0+\varepsilon\delta\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right)}{\partial h_i(t;\mathbf{x})}\left[a_j^{(1,0)}(\tau;\mathbf{z})+\varepsilon\delta a_j^{(1)}(\tau;\mathbf{z})\right]\right\}_{\varepsilon=0}.
\end{aligned}
\tag{64}
$$

Performing the operations indicated in Equation (64) and rearranging the various terms yields the following relations for $i=1,\ldots,TH$:

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

$$
\begin{aligned}
&\delta a_i^{(1)}(t;\mathbf{x})-\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,\frac{\partial G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})}\delta a_j^{(1)}(\tau;\mathbf{z}) \\
&-\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_j^{(1)}(\tau;\mathbf{z})\sum_{k=1}^{TH}\frac{\partial^2 G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_k(t;\mathbf{x})\partial h_i(t;\mathbf{x})}v_k^{(1)}(t;\mathbf{x}) \\
&-\sum_{j=1}^{TH}\frac{\partial^2 D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};t;\mathbf{x}\right]}{\partial h_j(t;\mathbf{x})\partial h_i(t;\mathbf{x})}v_j^{(1)}(t;\mathbf{x})=\sum_{k=1}^{TF}\frac{\partial^2 D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};t;\mathbf{x}\right]}{\partial F_k\partial h_i(t;\mathbf{x})}\delta F_k \\
&+\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_j^{(1)}(\tau;\mathbf{z})\sum_{k=1}^{TF}\frac{\partial^2 G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial F_k\partial h_i(t;\mathbf{x})}\delta F_k .
\end{aligned}
\tag{65}
$$

As indicated by the result obtained in Equation (65), the variations $\delta\mathbf{a}^{(1)}(t;\mathbf{x})$ are coupled to the variations $\mathbf{v}^{(1)}(t;\mathbf{x})$. Therefore, they can be obtained by simultaneously solving Equations (65) and (13), which together will be called the "2nd-Level Variational Sensitivity System (2nd-LVSS)" and its solution, namely the vector $\mathbf{v}^{(2)}(t;\mathbf{x})\triangleq\left[\delta\mathbf{a}^{(1)}(t;\mathbf{x}),\mathbf{v}^{(1)}(t;\mathbf{x})\right]^{\dagger}$ 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 [18], which considers the function $\mathbf{v}^{(2)}(t;\mathbf{x})\triangleq\left[\delta\mathbf{a}^{(1)}(t;\mathbf{x}),\mathbf{v}^{(1)}(t;\mathbf{x})\right]^{\dagger}$ to be an element in a Hilbert space denoted as $\mathsf{H}_2(\Omega_{tx})$. The Hilbert space denoted as $\mathsf{H}_2(\Omega_{tx})$ 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 $\boldsymbol{\chi}^{(2)}(t;\mathbf{x})\triangleq\left[\boldsymbol{\chi}_1^{(2)}(t;\mathbf{x}),\boldsymbol{\chi}_2^{(2)}(t;\mathbf{x})\right]^{\dagger}\in\mathsf{H}_2(\Omega_{tx})$, $\boldsymbol{\eta}^{(2)}(t;\mathbf{x})=\left[\boldsymbol{\eta}_1^{(2)}(t;\mathbf{x}),\boldsymbol{\eta}_2^{(2)}(t;\mathbf{x})\right]^{\dagger}\in\mathsf{H}_2(\Omega_{tx})$, with $\boldsymbol{\eta}_1^{(2)}(t;\mathbf{x})\triangleq\left[\eta_{1,1}^{(2)}(t;\mathbf{x}),\ldots,\eta_{1,TH}^{(2)}(t;\mathbf{x})\right]^{\dagger}$, $\boldsymbol{\eta}_2^{(2)}(t;\mathbf{x})\triangleq\left[\eta_{2,1}^{(2)}(t;\mathbf{x}),\ldots,\eta_{2,TH}^{(2)}(t;\mathbf{x})\right]^{\dagger}$, which is defined as follows:

$$
\begin{aligned}
&\left\langle\boldsymbol{\chi}^{(2)},\boldsymbol{\eta}^{(2)}\right\rangle_2\triangleq\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{x}\left[\boldsymbol{\chi}^{(2)}(t;\mathbf{x})\right]^{\dagger}\boldsymbol{\eta}^{(2)}(t;\mathbf{x})=\left\langle\boldsymbol{\chi}_1^{(2)},\boldsymbol{\eta}_1^{(2)}\right\rangle_1+\left\langle\boldsymbol{\chi}_2^{(2)},\boldsymbol{\eta}_2^{(2)}\right\rangle_1 \\
&=\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{x}\,\chi_{1,j}^{(2)}(t;\mathbf{x})\eta_{1,j}^{(2)}+\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{x}\,\chi_{2,j}^{(2)}(t;\mathbf{x})\eta_{2,j}^{(2)} .
\end{aligned}
\tag{66}
$$

Using the definition provided in Equation (66), for the inner product of Equations (13) and (65) with a vector $\mathbf{a}^{(2)}(t;\mathbf{x})=\left[\mathbf{a}_1^{(2)}(t;\mathbf{x}),\mathbf{a}_2^{(2)}(t;\mathbf{x})\right]^{\dagger}\in\mathsf{H}_2(\Omega_{tx})$, with components $\mathbf{a}_1^{(2)}(t;\mathbf{x})\triangleq\left[a_{1,1}^{(2)}(t;\mathbf{x}),\ldots,a_{1,TH}^{(2)}(t;\mathbf{x})\right]^{\dagger}$, $\mathbf{a}_2^{(2)}(t;\mathbf{x})\triangleq\left[a_{2,1}^{(2)}(t;\mathbf{x}),\ldots,a_{2,TH}^{(2)}(t;\mathbf{x})\right]^{\dagger}$, to obtain the following relation:

$$
\left\langle\mathbf{a}^{(2)}(t;\mathbf{x}),2^{nd}\,LASS\right\rangle_2=\left\langle\mathbf{a}_1^{(2)},Eq.(13)\right\rangle_1+\left\langle\mathbf{a}_2^{(2)},Eq.(66)\right\rangle_1 . \tag{67}
$$

Performing the inner-product indicated in Equation (67) yields the following relation:

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

$$\begin{aligned}
&\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{1,i}^{(2)}(t;\mathbf{x})v_i^{(1)}(t;\mathbf{x})-\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{1,i}^{(2)}(t;\mathbf{x})\\
&\times\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,\frac{\partial G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_j(\tau;\mathbf{z})}v_j^{(1)}(\tau;\mathbf{z})+\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(t;\mathbf{x})\delta a_i^{(1)}(t;\mathbf{x})\\
&-\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(t;\mathbf{x})\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,\frac{\partial G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})}\delta a_j^{(1)}(\tau;\mathbf{z})\\
&-\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(t;\mathbf{x})\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_j^{(1)}(\tau;\mathbf{z})\sum_{k=1}^{TH}\frac{\partial^2 G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_k(t;\mathbf{x})\partial h_i(t;\mathbf{x})}v_k^{(1)}(t;\mathbf{x})\\
&-\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(t;\mathbf{x})\sum_{j=1}^{TH}\frac{\partial^2 D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};t;\mathbf{x}\right]}{\partial h_j(t;\mathbf{x})\partial h_i(t;\mathbf{x})}v_j^{(1)}(t;\mathbf{x})=Q^{(2)}(\delta\mathbf{F}),
\end{aligned}\tag{68}$$

where:

$$\begin{aligned}
&Q^{(2)}(\delta\mathbf{F})\triangleq\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(t;\mathbf{x})\sum_{k=1}^{TF}\frac{\partial^2 D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};t;\mathbf{x}\right]}{\partial F_k\partial h_i(t;\mathbf{x})}\delta F_k\\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{1,i}^{(2)}(t;\mathbf{x})\sum_{k=1}^{TF}\left[\frac{\partial g_i(\mathbf{F};t;\mathbf{x})}{\partial F_k}+\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,\frac{\partial G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial F_k}\right]\delta F_k\\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(t;\mathbf{x})\sum_{m=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_m^{(1)}(\tau;\mathbf{z})\sum_{k=1}^{TF}\frac{\partial^2 G_m\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial F_k\partial h_i(t;\mathbf{x})}\delta F_k.
\end{aligned}\tag{69}$$

Interchanging the order of summations and integrations, and relabeling the dummy index of summation and variables of integration enables the recasting of the second term on the left-side of Equation (68) as follows:

$$\begin{aligned}
&\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{1,i}^{(2)}(t;\mathbf{x})\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,\frac{\partial G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_j(\tau;\mathbf{z})}v_j^{(1)}(\tau;\mathbf{z})\\
&=\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,v_j^{(1)}(\tau;\mathbf{z})\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{1,i}^{(2)}(t;\mathbf{x})\frac{\partial G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_j(\tau;\mathbf{z})}\\
&=\sum_{i=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,v_i^{(1)}(\tau;\mathbf{z})\sum_{j=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{1,j}^{(2)}(t;\mathbf{x})\frac{\partial G_j\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_i(\tau;\mathbf{z})}\\
&=\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,v_i^{(1)}(t;\mathbf{x})\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_{1,j}^{(2)}(\tau;\mathbf{z})\frac{\partial G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})}.
\end{aligned}\tag{70}$$

Interchanging the order of summations and integrations, and relabeling the dummy index of summation and variables of integration enables the recasting of the fourth term on the left-side of Equation (68) as follows:

$$\begin{aligned}
&\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(t;\mathbf{x})\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,\frac{\partial G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})}\delta a_j^{(1)}(\tau;\mathbf{z})\\
&=\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,\delta a_j^{(1)}(\tau;\mathbf{z})\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(t;\mathbf{x})\frac{\partial G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})}\\
&=\sum_{i=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,\delta a_i^{(1)}(\tau;\mathbf{z})\sum_{j=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,j}^{(2)}(t;\mathbf{x})\frac{\partial G_i\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_j(t;\mathbf{x})}\\
&=\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,\delta a_i^{(1)}(t;\mathbf{x})\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_{2,j}^{(2)}(\tau;\mathbf{z})\frac{\partial G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_j(\tau;\mathbf{z})}.
\end{aligned}\tag{71}$$

Interchanging the order of summations and integrations, and relabeling the dummy index of summation and variables of integration enables the recasting of the fifth term on the left-side of Equation (68) as follows:

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

$$\begin{aligned}&\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(t;\mathbf{x})\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_j^{(1)}(\tau;\mathbf{z})\sum_{k=1}^{TH}\frac{\partial^2 G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_k(t;\mathbf{x})\partial h_i(t;\mathbf{x})}v_k^{(1)}(t;\mathbf{x})\\&=\sum_{k=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,v_k^{(1)}(t;\mathbf{x})\left\{\sum_{i=1}^{TH}a_{2,i}^{(2)}(t;\mathbf{x})\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_j^{(1)}(\tau;\mathbf{z})\frac{\partial^2 G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_k(t;\mathbf{x})\partial h_i(t;\mathbf{x})}\right\}.\end{aligned}\tag{72}$$

Interchanging the order of summations and relabeling the dummy index of summation enables the recasting of the sixth term on the left-side of Equation (68) as follows:

$$\begin{aligned}&\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(t;\mathbf{x})\sum_{j=1}^{TH}\frac{\partial^2 D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};t;\mathbf{x}\right]}{\partial h_j(t;\mathbf{x})\partial h_i(t;\mathbf{x})}v_j^{(1)}(t;\mathbf{x})\\&=\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,v_i^{(1)}(t;\mathbf{x})\sum_{j=1}^{TH}a_{2,j}^{(2)}(t;\mathbf{x})\frac{\partial^2 D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};t;\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})\partial h_j(t;\mathbf{x})}.\end{aligned}\tag{73}$$

The results obtained in Equations (70)–(73) are now used to recast the relation provided in Equation (68) as follows:

$$\begin{aligned}&\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{1,i}^{(2)}(t;\mathbf{x})v_i^{(1)}(t;\mathbf{x})-\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,v_i^{(1)}(t;\mathbf{x})\\&\times\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_{1,j}^{(2)}(\tau;\mathbf{z})\frac{\partial G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})}+\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(t;\mathbf{x})\delta a_i^{(1)}(t;\mathbf{x})\\&-\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,\delta a_i^{(1)}(t;\mathbf{x})\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_{2,j}^{(2)}(\tau;\mathbf{z})\frac{\partial G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_j(\tau;\mathbf{z})}\\&-\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,v_i^{(1)}(t;\mathbf{x})\left\{\sum_{j=1}^{TH}a_{2,j}^{(2)}(t;\mathbf{x})\sum_{k=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_k^{(1)}(\tau;\mathbf{z})\frac{\partial^2 G_k\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})\partial h_j(t;\mathbf{x})}\right\}\\&-\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,v_i^{(1)}(t;\mathbf{x})\sum_{j=1}^{TH}a_{2,j}^{(2)}(t;\mathbf{x})\frac{\partial^2 D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};t;\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})\partial h_j(t;\mathbf{x})}=Q^{(2)}(\delta\mathbf{F}).\end{aligned}\tag{74}$$

The terms on the left side of Equation (74) can be grouped as follows:

$$\begin{aligned}&\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,v_i^{(1)}(t;\mathbf{x})\left\{a_{1,i}^{(2)}(t;\mathbf{x})-\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_{1,j}^{(2)}(\tau;\mathbf{z})\frac{\partial G_j\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})}\right.\\&-\sum_{j=1}^{TH}a_{2,j}^{(2)}(t;\mathbf{x})\sum_{k=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_k^{(1)}(\tau;\mathbf{z})\frac{\partial^2 G_k\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})\partial h_j(t;\mathbf{x})}\\&\left.-\sum_{j=1}^{TH}a_{2,j}^{(2)}(t;\mathbf{x})\frac{\partial^2 D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};t;\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})\partial h_j(t;\mathbf{x})}\right\}\\&+\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,\delta a_i^{(1)}(t;\mathbf{x})\left\{a_{2,i}^{(2)}(t;\mathbf{x})-\sum_{j=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_{2,j}^{(2)}(\tau;\mathbf{z})\frac{\partial G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_j(\tau;\mathbf{z})}\right\}\\&=Q^{(2)}(\delta\mathbf{F}).\end{aligned}\tag{75}$$

The left-side of Equation (75) can now be required to represent the indirect-effect term defined in Equation (63), by imposing the requirement that the hitherto arbitrary function $\mathbf{a}^{(2)}(t;\mathbf{x})=\left[\mathbf{a}_1^{(2)}(t;\mathbf{x}),\mathbf{a}_2^{(2)}(t;\mathbf{x})\right]^{\dagger}\in\mathsf{H}_2(\Omega_{tx})$ be the solution of the following NIE-like systems of equations:

$$\begin{aligned}
&a_{1,i}^{(2)}(j;t;\mathbf{x})-\sum_{m=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_{1,m}^{(2)}(j;\tau;\mathbf{z})\frac{\partial G_m\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})}\\
&-\sum_{m=1}^{TH}a_{2,m}^{(2)}(j;t;\mathbf{x})\sum_{k=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_k^{(1)}(\tau;\mathbf{z})\frac{\partial^2 G_k\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})\partial h_m(t;\mathbf{x})}\\
&-\sum_{m=1}^{TH}a_{2,m}^{(2)}(j;t;\mathbf{x})\frac{\partial^2 D\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};t;\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})\partial h_m(t;\mathbf{x})}=\frac{\partial^2 D\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta});t;\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})\partial F_j}\\
&+\sum_{m=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_m^{(1)}(\tau;\mathbf{z})\frac{\partial^2 G_m\left[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}\right]}{\partial h_i(t;\mathbf{x})\partial F_j};\ i=1,\ldots,TH;\ \ j=1,\ldots,TF.
\end{aligned}\tag{76}$$

$$\begin{aligned}
&a_{2,i}^{(2)}(j;t;\mathbf{x})-\sum_{m=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_{2,m}^{(2)}(j;\tau;\mathbf{z})\frac{\partial G_i\left[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}\right]}{\partial h_m(\tau;\mathbf{z})}\\
&=\frac{\partial g_i(\mathbf{F};t;\mathbf{x})}{\partial F_j}+\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\frac{\partial G_i(\mathbf{h};\mathbf{F};t,\tau;\mathbf{x},\mathbf{z})}{\partial F_j};\ i=1,\ldots,TH;\ \ j=1,\ldots,TF.
\end{aligned}\tag{77}$$

The above NIE-system will be called the "2nd-Level Adjoint Sensitivity System (2nd-LASS)" and its solution, $\mathbf{a}^{(2)}(j;t;\mathbf{x})=\left[\mathbf{a}_1^{(2)}(j;t;\mathbf{x}),\mathbf{a}_2^{(2)}(j;t;\mathbf{x})\right]^{\dagger}$, $j=1,\ldots,TF$, will be called the "2nd-level adjoint sensitivity function." It is important to note that the sources on the right-sides of Equations (76) and (77), respectively, stem from the first-order sensitivities $\partial R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]/\partial F_j$, $j=1,\ldots,TF$, and are hence dependent on the index "j". This means that for each first-order sensitivity $\partial R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]/\partial F_j$, there will correspond a distinct 2nd-LASS, with a distinct solution $\mathbf{a}^{(2)}(j;t;\mathbf{x})=\left[\mathbf{a}_1^{(2)}(j;t;\mathbf{x}),\mathbf{a}_2^{(2)}(j;t;\mathbf{x})\right]^{\dagger}$, a fact that has been emphasized by adding the index "j" to the list of arguments of the 2nd-level adjoint sensitivity function. Therefore, there will be as many 2nd-level adjoint functions as there are distinct first-order sensitivities $\partial R\left[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})\right]/\partial F_j$, which is equivalent to the number of components F_j of the "feature-function" $\mathbf{F}(\boldsymbol{\theta})$. It is also very important to note that the integral operators on the left-sides of Equations (76) and (77) are independent of the index "j", which means that the same left-hand sides need to be inverted for computing the 2nd-level adjoint function, regardless of the right-sides (which corresponds to the particular component of the feature-function) of Equations (76) and (77). Therefore, if the inverses of the operators appearing on the left-sides of Equations (76) and (77) could be stored, they would not need to be inverted repeatedly, so the various 2nd-level adjoint functions would be computed most efficiently.

It follows from Equations (75)–(77) that the indirect-effect term $\delta\left(\partial R/\partial F_j\right)_{ind}$ defined by Equation (63) can be expressed in terms of the 2nd-level adjoint sensitivity function as follows:

$$\delta\left(\partial R/\partial F_j\right)_{ind}=Q^{(2)}(\delta\mathbf{F}).\tag{78}$$

Using the results obtained in Equations (78), (69), and (62) in Equation (61) yields the following expression for the (second-order) G-differential $\delta\left\{\partial R/\partial F_j\right\}$ for $j=1,\ldots,TF$:

$$\begin{aligned}
\delta\left(\frac{\partial R}{\partial F_j}\right) \square \sum_{k=1}^{TF}\frac{\partial^2 R}{\partial F_k \partial F_j}\delta F_k &= \sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x}) \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \sum_{k=1}^{TF}\frac{\partial^2 G_i(\mathbf{h};\mathbf{F};t,\tau;\mathbf{x},\mathbf{z})}{\partial F_k \partial F_j}\delta F_k \\
&+\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x} \sum_{k=1}^{TF}\frac{\partial D[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})]}{\partial F_k \partial F_j}\delta F_k + \sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x}) \sum_{k=1}^{TF}\frac{\partial^2 g_i(\mathbf{h};\mathbf{F})}{\partial F_k \partial F_j}\delta F_k \\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_{2,i}^{(2)}(j;t;\mathbf{x}) \sum_{k=1}^{TF}\frac{\partial^2 D[\mathbf{h}(t;\mathbf{x});\mathbf{F};t;\mathbf{x}]}{\partial F_k \partial h_i(t;\mathbf{x})}\delta F_k \qquad (79)\\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_{1,i}^{(2)}(j;t;\mathbf{x}) \sum_{k=1}^{TF}\left[\frac{\partial g_i(\mathbf{F};t;\mathbf{x})}{\partial F_k} + \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \frac{\partial G_i[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}]}{\partial F_k}\right]\delta F_k \\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_{2,i}^{(2)}(j;t;\mathbf{x}) \sum_{m=1}^{TH}\int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z}\, a_m^{(1)}(\tau;\mathbf{z}) \sum_{k=1}^{TF}\frac{\partial^2 G_m[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}]}{\partial F_k \partial h_i(t;\mathbf{x})}\delta F_k .
\end{aligned}$$

It follows from Equation (79) that the second-order sensitivities $\partial^2 R[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})]/\partial F_i \partial F_j$ of the decoder-response with respect to the components of the feature function have the following expressions for $i,j=1,\ldots,TF$:

$$\begin{aligned}
\frac{\partial^2 R}{\partial F_k \partial F_j} &= \sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x}) \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \frac{\partial^2 G_i(\mathbf{h};\mathbf{F};t,\tau;\mathbf{x},\mathbf{z})}{\partial F_k \partial F_j} \\
&+\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x} \frac{\partial D[\mathbf{h};\mathbf{F}(\boldsymbol{\theta})]}{\partial F_k \partial F_j} + \sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_i^{(1)}(t;\mathbf{x}) \frac{\partial^2 g_i(\mathbf{h};\mathbf{F})}{\partial F_k \partial F_j} \\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_{2,i}^{(2)}(j;t;\mathbf{x}) \frac{\partial^2 D[\mathbf{h}(t;\mathbf{x});\mathbf{F};t;\mathbf{x}]}{\partial F_k \partial h_i(t;\mathbf{x})} \qquad (80)\\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_{1,i}^{(2)}(j;t;\mathbf{x}) \left[\frac{\partial g_i(\mathbf{F};t;\mathbf{x})}{\partial F_k} + \int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z} \frac{\partial G_i[\mathbf{h}(\tau;\mathbf{z});\mathbf{F};t,\tau;\mathbf{x},\mathbf{z}]}{\partial F_k}\right] \\
&+\sum_{i=1}^{TH}\int_{t_0}^{t_f} dt \int_{\Omega} d\mathbf{x}\, a_{2,i}^{(2)}(j;t;\mathbf{x}) \sum_{m=1}^{TH}\int_{t_0}^{t_f} d\tau \int_{\Omega} d\mathbf{z}\, a_m^{(1)}(\tau;\mathbf{z}) \frac{\partial^2 G_m[\mathbf{h}(t;\mathbf{x});\mathbf{F};\tau,t;\mathbf{z},\mathbf{x}]}{\partial F_k \partial h_i(t;\mathbf{x})}.
\end{aligned}$$

The second-order sensitivities of the decoder-response with respect to the optimal weights/parameters $\theta_k, k=1,\ldots,TW$, are obtained analytically by using the chain rule in conjunction with the expressions obtained in Equations (23) and (80), as follows:

$$\frac{\partial^2 R[\mathbf{F}(\boldsymbol{\theta})]}{\partial \theta_k \partial \theta_j} = \frac{\partial}{\partial \theta_k}\left\{\sum_{i=1}^{TF}\frac{\partial R[\mathbf{F}(\boldsymbol{\theta})]}{\partial F_i(\boldsymbol{\theta})}\frac{\partial F_i(\boldsymbol{\theta})}{\partial \theta_j}\right\}, \quad j,k=1,\ldots,TW. \qquad (81)$$

4.2. Particular Case: The Second-Order Comprehensive Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Fredholm-Type (2nd-CASAM-NIE-Fredholm).

When there are no feature functions but only individual model parameters, i.e., when $F_i(\boldsymbol{\theta}) \equiv \theta_i$ for all $i=1,\ldots,TF \square TW$, the expression obtained in Equation (80) yields directly the second-order sensitivities $\partial^2 R/\partial \theta_i \partial \theta_j$, for all $i,j=1,\ldots,TW$, taking on the following form:

$$\begin{aligned}&\frac{\partial^2 R}{\partial\theta_k\partial\theta_j}=\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_i^{(1)}(t;\mathbf{x})\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\frac{\partial^2 G_i(\mathbf{h};\boldsymbol{\theta};t,\tau;\mathbf{x},\mathbf{z})}{\partial\theta_k\partial\theta_j}\\&+\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\frac{\partial D[\mathbf{h};\boldsymbol{\theta}]}{\partial\theta_k\partial\theta_j}+\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_i^{(1)}(t;\mathbf{x})\frac{\partial^2 g_i(\mathbf{h};\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j}\\&+\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(j;t;\mathbf{x})\frac{\partial^2 D[\mathbf{h}(t;\mathbf{x});\boldsymbol{\theta};t;\mathbf{x}]}{\partial\theta_k\partial h_i(t;\mathbf{x})}\\&+\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{1,i}^{(2)}(j;t;\mathbf{x})\left[\frac{\partial g_i(\boldsymbol{\theta};t;\mathbf{x})}{\partial\theta_k}+\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\frac{\partial G_i[\mathbf{h}(\tau;\mathbf{z});\boldsymbol{\theta};t,\tau;\mathbf{x},\mathbf{z}]}{\partial\theta_k}\right]\\&+\sum_{i=1}^{TH}\int_{t_0}^{t_f}dt\int_{\Omega}d\mathbf{x}\,a_{2,i}^{(2)}(j;t;\mathbf{x})\sum_{m=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_m^{(1)}(\tau;\mathbf{z})\frac{\partial^2 G_m[\mathbf{h}(t;\mathbf{x});\boldsymbol{\theta};\tau,t;\mathbf{z},\mathbf{x}]}{\partial\theta_k\partial h_i(t;\mathbf{x})}.\end{aligned}\tag{82}$$

where the 2nd-level adjoint sensitivity function $\mathbf{a}^{(2)}(j;t;\mathbf{x})=\left[\mathbf{a}_1^{(2)}(j;t;\mathbf{x}),\mathbf{a}_2^{(2)}(j;t;\mathbf{x})\right]^{\dagger}$ will be the solution of the following form of the 2nd-LASS:

$$\begin{aligned}&a_{1,i}^{(2)}(j;t;\mathbf{x})-\sum_{m=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_{1,m}^{(2)}(j;\tau;\mathbf{z})\frac{\partial G_m[\mathbf{h}(t;\mathbf{x});\boldsymbol{\theta};\tau,t;\mathbf{z},\mathbf{x}]}{\partial h_i(t;\mathbf{x})}\\&-\sum_{m=1}^{TH}a_{2,m}^{(2)}(j;t;\mathbf{x})\sum_{k=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_k^{(1)}(\tau;\mathbf{z})\frac{\partial^2 G_k[\mathbf{h}(t;\mathbf{x});\boldsymbol{\theta};\tau,t;\mathbf{z},\mathbf{x}]}{\partial h_i(t;\mathbf{x})\partial h_m(t;\mathbf{x})}\\&-\sum_{m=1}^{TH}a_{2,m}^{(2)}(j;t;\mathbf{x})\frac{\partial^2 D[\mathbf{h}(t;\mathbf{x});\boldsymbol{\theta};t;\mathbf{x}]}{\partial h_i(t;\mathbf{x})\partial h_m(t;\mathbf{x})}=\frac{\partial^2 D[\mathbf{h};\boldsymbol{\theta};t;\mathbf{x}]}{\partial h_i(t;\mathbf{x})\partial\theta_j}\\&+\sum_{m=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_m^{(1)}(\tau;\mathbf{z})\frac{\partial^2 G_m[\mathbf{h}(t;\mathbf{x});\boldsymbol{\theta};\tau,t;\mathbf{z},\mathbf{x}]}{\partial h_i(t;\mathbf{x})\partial\theta_j};\ i=1,\dots,TH;\ \ j=1,\dots,TW.\end{aligned}\tag{83}$$

$$\begin{aligned}&a_{2,i}^{(2)}(j;t;\mathbf{x})-\sum_{m=1}^{TH}\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\,a_{2,m}^{(2)}(j;\tau;\mathbf{z})\frac{\partial G_i[\mathbf{h}(\tau;\mathbf{z});\boldsymbol{\theta};t,\tau;\mathbf{x},\mathbf{z}]}{\partial h_m(\tau;\mathbf{z})}\\&=\frac{\partial g_i(\boldsymbol{\theta};t;\mathbf{x})}{\partial\theta_j}+\int_{t_0}^{t_f}d\tau\int_{\Omega}d\mathbf{z}\frac{\partial G_i(\mathbf{h};\boldsymbol{\theta};t,\tau;\mathbf{x},\mathbf{z})}{\partial\theta_j};\ i=1,\dots,TH;\ \ j=1,\dots,TW.\end{aligned}\tag{84}$$

As indicated by Equations (83) and (84), the 2nd-LASS would need to be solved TW -times when no "feature functions" can be constructed. In contradistinction, when TF feature functions can be constructed, the 2nd-LASS would need to be solved only TF -times $(TF<TW)$. Thus, when there are no feature functions of parameters, the 2nd-FASAM-NODE reduces to the "Second-Order Comprehensive Adjoint Sensitivity Analysis Methodology" [18] applied to Neural Integral Equations of Fredholm-Type (2nd-CASAM-NIE-Fredholm).

5. Illustrative Application of the 2nd-FASAM-NIE-Fredholm Methodology

The application of the 2nd-FASAM-NIE-Fredholm methodology will be illustrated in this Section by continuing the analysis of the model considered in Section 3. As generally illustrated in Section 4, the second-order sensitivities are conceptually determined by considering them to be the "first-order sensitivities of the first-order sensitivities." Thus, the application of the 2nd-FASAM-NIE-Fredholm methodology commences by considering the first-order sensitivities obtained in Equations (51)–(53)

5.1. Determining the Second-Order Sensitivities Stemming from $\partial R/\partial F_1$ using Equation(51)

The second-order sensitivities stemming from the first-order sensitivity $\partial R/\partial F_1$ are obtained by G-differentiating Equation (51), which by definition yields the following expression:

$$\delta\left(\frac{\partial R}{\partial F_1}\right)=-\left\{c_1\frac{d}{d\varepsilon}\int_a^b\left[a_1^{(1)}(t)+\varepsilon\delta a_1^{(1)}\right]t^{\lambda_1}dt\right\}_{\varepsilon=0}=-c_1^0\int_a^b\delta a_1^{(1)}(t)t^{\lambda_1}dt,\tag{85}$$

 doi:10.20944/preprints202412.0474.v1

where the superscript "zero" indicates, as usual, the nominal/optimal values for the respective quantities. The variational function $\delta a_1^{(1)}(t)$ is the solution of the equations obtained by differentiating the 1st-LASS defined by Equation (45) and (46). The G-differential of the 1st-LASS is obtained as follows:

$$\left\{\frac{d}{d\varepsilon}\left[a_1^{(1,0)}(t)+\varepsilon\delta a_1^{(1)}\right]\right\}_{\varepsilon=0} - \left\{2t^{\lambda_1}\frac{d}{d\varepsilon}\int_a^b \tau^{\lambda_1}\left[a_1^{(1,0)}(\tau)+\varepsilon\delta a_1^{(1)}(\tau)\right]d\tau\right\}_{\varepsilon=0}$$
$$=\left\{\frac{d}{d\varepsilon}\left[F_3^0(\boldsymbol{\gamma})+\varepsilon\delta F_3\right]\left[y_2^0(t)+\varepsilon\delta y_2(t)\right]\right\}_{\varepsilon=0}; \qquad (86)$$

$$\left\{\frac{d}{d\varepsilon}\left[a_2^{(1,0)}(t)+\varepsilon\delta a_2^{(1)}(t)\right]\right\}_{\varepsilon=0} - \left\{2t^{\lambda_2}\frac{d}{d\varepsilon}\int_a^b \tau^{\lambda_2}\left[a_2^{(1,0)}(\tau)+\varepsilon\delta a_2^{(1)}(\tau)\right]d\tau\right\}_{\varepsilon=0}$$
$$=\left\{\frac{d}{d\varepsilon}\left[F_3^0(\boldsymbol{\gamma})+\varepsilon\delta F_3\right]\left[y_1^0(t)+\varepsilon\delta y_1(t)\right]\right\}_{\varepsilon=0}. \qquad (87)$$

Performing in Equations (86) and (87) the operations involving the scalar quantity ε yields the following system of equations:

$$\delta a_1^{(1)}(t) - 2c_1^0 t^{\lambda_1}\int_a^b \tau^{\lambda_1}\delta a_1^{(1)}(\tau)d\tau - F_3^0(\boldsymbol{\gamma})\delta y_2(t) = (\delta F_3)y_2^0(t); \qquad (88)$$

$$\delta a_2^{(1)}(t) - 2c_2^0 t^{\lambda_2}\int_a^b \tau^{\lambda_2}\delta a_2^{(1)}(\tau)d\tau - F_3^0(\boldsymbol{\gamma})\delta y_1(t) = (\delta F_3)y_1^0(t). \qquad (89)$$

It is evident from Equations (88) and (89) that the variational function $\delta\mathbf{a}^{(1)}(t)\triangleq\left[\delta a_1^{(1)}(t),\delta a_2^{(1)}(t)\right]^\dagger$ is coupled to the variational function $\delta\mathbf{y}(t)\triangleq\left[\delta y_1(t),\delta y_2(t)\right]^\dagger$. Therefore, the 2nd-level variational function $\mathbf{v}^{(1)}(t)\triangleq\left[\delta\mathbf{y}(t),\delta\mathbf{a}^{(1)}(t)\right]^\dagger$, constructed by concatenating the functions $\delta\mathbf{y}(t)$ and $\delta\mathbf{a}^{(1)}(t)$ will be the solution of the "2nd-Level Variational Sensitivity System (2nd-LVSS)" constructed by concatenating Equations (88), (89) together with the 1st-LVSS comprising Equations (40) and (41). It is evident that the 2nd-LVSS depends on the parameter variations and would need to be solved repeatedly for each parameter variation of interest. The need for repeatedly solving the 2nd-LVSS can be alleviated by applying the 2nd-FASAM-NIE-Fredholm methodology presented in Section 4 to construct a 2nd-LASS that would correspond to the 2nd-LVSS defined by Equations (88), (89), (40) and (41). This 2nd-LASS will be constructed by considering a Hilbert space denoted as $\mathsf{H}_2(a,b)$ and endowed with the following inner product, denoted as $\left\langle\boldsymbol{\chi}^{(2)}(t),\boldsymbol{\eta}^{(2)}(t)\right\rangle_2$, between two vectors $\boldsymbol{\chi}^{(2)}(t)\triangleq\left[\boldsymbol{\chi}_1^{(2)}(t),\boldsymbol{\chi}_2^{(2)}(t)\right]^\dagger\in\mathsf{H}_2(a,b)$, $\boldsymbol{\chi}_i^{(2)}(t)\triangleq\left[\chi_{i,1}^{(2)}(t),\chi_{i,2}^{(2)}(t)\right]^\dagger$, $i=1,2$, and $\boldsymbol{\eta}^{(2)}(t)\triangleq\left[\boldsymbol{\eta}_1^{(2)}(t),\boldsymbol{\eta}_2^{(2)}(t)\right]^\dagger\in\mathsf{H}_2(a,b)$, $\boldsymbol{\eta}_i^{(2)}(t)\triangleq\left[\eta_{i,1}^{(2)}(t),\eta_{i,2}^{(2)}(t)\right]^\dagger$, $i=1,2$:

$$\left\langle\boldsymbol{\chi}^{(1)}(t),\boldsymbol{\eta}^{(1)}(t)\right\rangle_2\triangleq\int_a^b\left[\boldsymbol{\chi}^{(1)}(t)\right]^\dagger\boldsymbol{\eta}^{(1)}(t)dt=\sum_{i=1}^{2}\int_a^b\boldsymbol{\chi}_i^{(2)}(t)\boldsymbol{\eta}_i^{(2)}(t)dt$$
$$=\sum_{i=1}^{2}\int_a^b\chi_{i,1}^{(2)}(t)\eta_{i,1}^{(2)}(t)dt+\sum_{j=1}^{2}\int_a^b\chi_{i,2}^{(2)}(t)\eta_{i,2}^{(2)}(t)dt. \qquad (90)$$

Using the inner-product defined in Equation (90), construct the inner-product of the 2nd-LVSS with a function denoted as $\mathbf{a}^{(2)}(1;t)=\left[\mathbf{a}_1^{(2)}(1;t),\mathbf{a}_2^{(2)}(1;t)\right]^\dagger$, $\mathbf{a}_i^{(2)}(1;t)\triangleq\left[a_{i,1}^{(2)}(1;t),a_{i,2}^{(2)}(1;t)\right]^\dagger$, $i=1,2$, to obtain the following relation:

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

$$\begin{aligned}
Q^{(2)}(\delta\mathbf{F}) &= \int_a^b a_{1,1}^{(2)}(1;t)\,\delta y_1(t)\,dt - 2c_1\int_a^b a_{1,1}^{(2)}(1;t)\,t^{\lambda_1}dt\int_a^b \tau^{\lambda_1}\delta y_1(\tau)\,d\tau \\
&+\int_a^b a_{1,2}^{(2)}(1;t)\,\delta y_2(t)\,dt - 2c_2\int_a^b a_{1,2}^{(2)}(1;t)\,t^{\lambda_2}dt\int_a^b \tau^{\lambda_2}\delta y_2(\tau)\,d\tau \\
&+\int_a^b a_{2,1}^{(2)}(1;t)\,\delta a_1^{(1)}(t)\,dt - 2c_1\int_a^b a_{2,1}^{(2)}(1;t)\,t^{\lambda_1}dt\int_a^b \tau^{\lambda_1}\delta a_1^{(1)}(\tau)\,d\tau \\
&-F_3\int_a^b a_{2,1}^{(2)}(1;t)\,\delta y_2(t)\,dt + \int_a^b a_{2,2}^{(2)}(1;t)\,\delta a_2^{(1)}(t)\,dt \\
&-2c_2\int_a^b a_{2,2}^{(2)}(1;t)\,t^{\lambda_2}dt\int_a^b \tau^{\lambda_2}\delta a_2^{(1)}(\tau)\,d\tau - F_3\int_a^b a_{2,2}^{(2)}(1;t)\,\delta y_1(t)\,dt,
\end{aligned} \tag{91}$$

where:

$$\begin{aligned}
Q^{(2)}(\delta\mathbf{F}) &\triangleq -c_1(\delta F_1)\int_a^b t^{\lambda_1}a_{1,1}^{(2)}(1;t)\,dt - c_2(\delta F_2)\int_a^b a_{1,2}^{(2)}(1;t)\,t^{\lambda_2}dt \\
&+(\delta F_3)\int_a^b a_{2,1}^{(2)}(1;t)\,y_2(t)\,dt + (\delta F_3)\int_a^b a_{2,2}^{(2)}(1;t)\,y_1(t)\,dt.
\end{aligned} \tag{92}$$

The argument "1" in the list of arguments of $\mathbf{a}^{(2)}(1;t)=\left[\mathbf{a}_1^{(2)}(1;t),\mathbf{a}_2^{(2)}(1;t)\right]^\dagger$ indicates that this 2nd-level adjoint function corresponds to the first-order sensitivity $\partial R/\partial F_1$ of the response with respect to the "first" component, $F_1(\boldsymbol{\alpha})$, of the "feature function $\mathbf{F}(\boldsymbol{\theta})$.

The right-side of Equation (91) is transformed as follows:

$$\begin{aligned}
Q^{(2)}(\delta\mathbf{F}) &= \int_a^b \delta y_1(t)\left\{a_{1,1}^{(2)}(1;t) - 2c_1 t^{\lambda_1}\int_a^b a_{1,1}^{(2)}(1;\tau)\,\tau^{\lambda_1}dt - F_3 a_{2,2}^{(2)}(1;t)\right\}dt \\
&+\int_a^b \delta y_2(t)\left\{a_{1,2}^{(2)}(1;t) - 2c_2 t^{\lambda_2}\int_a^b a_{1,2}^{(2)}(1;\tau)\,\tau^{\lambda_2}d\tau - F_3 a_{2,1}^{(2)}(1;t)\right\}dt \\
&+\int_a^b \delta a_1^{(1)}(t)\left\{a_{2,1}^{(2)}(1;t) - 2c_1 t^{\lambda_1}\int_a^b a_{2,1}^{(2)}(1;\tau)\,\tau^{\lambda_1}dt\right\}dt \\
&+\int_a^b \delta a_2^{(1)}(t)\left\{a_{2,2}^{(2)}(1;t) - 2c_2 t^{\lambda_2}\int_a^b a_{2,2}^{(2)}(1;\tau)\,\tau^{\lambda_2}dt\right\}dt.
\end{aligned} \tag{93}$$

The right-side of Equation (93) is now required to represent the (second-order) G-differential $\delta(\partial R/\partial F_1)$ defined in Equation (85). This requirement is accomplished by imposing the following relationships on the components of $\mathbf{a}^{(2)}(1;t)=\left[\mathbf{a}_1^{(2)}(1;t),\mathbf{a}_2^{(2)}(1;t)\right]^\dagger$:

$$a_{1,1}^{(2)}(1;t) - 2c_1 t^{\lambda_1}\int_a^b a_{1,1}^{(2)}(1;\tau)\,\tau^{\lambda_1}dt - F_3 a_{2,2}^{(2)}(1;t) = 0; \tag{94}$$

$$a_{1,2}^{(2)}(1;t) - 2c_2 t^{\lambda_2}\int_a^b a_{1,2}^{(2)}(1;\tau)\,\tau^{\lambda_2}d\tau - F_3 a_{2,1}^{(2)}(1;t) = 0; \tag{95}$$

$$a_{2,1}^{(2)}(1;t) - 2c_1 t^{\lambda_1}\int_a^b a_{2,1}^{(2)}(1;\tau)\,\tau^{\lambda_1}dt = -c_1 t^{\lambda_1}; \tag{96}$$

$$a_{2,2}^{(2)}(1;t) - 2c_2 t^{\lambda_2}\int_a^b a_{2,2}^{(2)}(1;\tau)\,\tau^{\lambda_2}dt = 0. \tag{97}$$

The system of equations represented by Equations (94)–(97) constitutes the "2nd-Level Adjoint Sensitivity System (2nd-LASS)" for the 2nd-level adjoint sensitivity function $\mathbf{a}^{(2)}(1;t)=\left[\mathbf{a}_1^{(2)}(1;t),\mathbf{a}_2^{(2)}(1;t)\right]^\dagger$. This 2nd-LASS is linear in $\mathbf{a}^{(2)}(1;t)$ and is independent of any parameter variations, so it needs to be solved just once to determine $\mathbf{a}^{(2)}(1;t)$. It follows from

 doi:10.20944/preprints202412.0474.v1

Equations (93)–(97) that the second-order G-differential $\delta\left(\partial R/\partial F_1\right)$ has the following expression in terms of the components of the 2nd-level adjoint sensitivity function $\mathbf{a}^{(2)}(1;t)=\left[\mathbf{a}_1^{(2)}(1;t),\mathbf{a}_2^{(2)}(1;t)\right]^{\dagger}$:

$$\delta\left(\partial R/\partial F_1\right)=Q^{(2)}\left(\delta\mathbf{F}\right)=-c_1\left(\delta F_1\right)\int_a^b t^{\lambda_1}a_{1,1}^{(2)}(1;t)\,dt-c_2\left(\delta F_2\right)\int_a^b a_{1,2}^{(2)}(1;t)\,t^{\lambda_2}dt$$
$$+\left(\delta F_3\right)\int_a^b a_{2,1}^{(2)}(1;t)\,y_2(t)\,dt+\left(\delta F_3\right)\int_a^b a_{2,2}^{(2)}(1;t)\,y_1(t)\,dt \triangleq \sum_{i=1}^{3}\frac{\partial^2 R}{\partial F_i\partial F_1}\partial F_i\,. \tag{98}$$

It follows from Equation (98) that the second-order sensitivities stemming from the first-order sensitivity have the following expressions:

$$\frac{\partial^2 R}{\partial F_1\partial F_1}=-c_1\int_a^b t^{\lambda_1}a_{1,1}^{(2)}(1;t)\,dt\,; \tag{99}$$

$$\frac{\partial^2 R}{\partial F_2\partial F_1}=-c_2\int_a^b a_{1,2}^{(2)}(1;t)\,t^{\lambda_2}dt\,; \tag{100}$$

$$\frac{\partial^2 R}{\partial F_3\partial F_1}=\int_a^b a_{2,1}^{(2)}(1;t)\,y_2(t)\,dt+\int_a^b a_{2,2}^{(2)}(1;t)\,y_1(t)\,dt\,. \tag{101}$$

The 2nd-LASS comprising Equations (94)–(97) can be readily solved to obtain the following expressions for the components of the 2nd-level adjoint sensitivity function $\mathbf{a}^{(2)}(1;t)$:

$$a_{1,1}^{(2)}(1;t)=0;\;\; a_{1,2}^{(2)}(1;t)=F_3\left(\boldsymbol{\gamma}\right)c_1\left(t^{\lambda_1}-2c_2kt^{\lambda_2}\right);\;\; a_{2,1}^{(2)}(1;t)=c_1t^{\lambda_1};\;\; a_{2,2}^{(2)}(1;t)=0\,. \tag{102}$$

Using in Equations (99)–(101) the results obtained in Equation (102) yields the following closed-form analytical expressions:

$$\frac{\partial^2 R}{\partial F_1\partial F_1}=0;\;\; \frac{\partial^2 R}{\partial F_2\partial F_1}=c_1c_2kF_3\left(\boldsymbol{\gamma}\right);\;\; \frac{\partial^2 R}{\partial F_3\partial F_1}=c_1c_2kF_2\left(\boldsymbol{\beta}\right). \tag{103}$$

The correctness of the results obtained in Equation (103) can be readily verified by using the expression of the first-order sensitivities obtained in Equations (54)–(56).

5.2. Determining the Second-Order Sensitivities Stemming from $\partial R/\partial F_2$ using Equation(52)

The second-order sensitivities stemming from the first-order sensitivity $\partial R/\partial F_1$ are obtained by G-differentiating Equation (52), which by definition yields the following expression:

$$\delta\left(\frac{\partial R}{\partial F_2}\right)=-\left\{c_2\frac{d}{d\varepsilon}\int_a^b\left[a_2^{(1,0)}(t)+\varepsilon\delta a_2^{(1)}\right]t^{\lambda_2}dt\right\}_{\varepsilon=0}=-c_2^0\int_a^b\delta a_2^{(1)}(t)\,t^{\lambda_2}dt\,, \tag{104}$$

where the superscript "zero" indicates, as usual, the nominal/optimal values for the respective quantities. The variational function $\delta a_2^{(1)}(t)$ which appears in Equation (104) is the solution of the "2nd-Level Variational Sensitivity System (2nd-LVSS)" defined by concatenating Equations (88), (89), (40) and (41). Following the same procedure as detailed in Subsection 5.1, the need for repeatedly solving the 2nd-LVSS for every parameter variation is alleviated by constructing the 2nd-LASS that will correspond to the (second-order) G-differential $\delta\left(\partial R/\partial F_2\right)$ obtained in Equation (104). These detailed derivations, which mirror those presented in Subsection 5.1, will be omitted here for the sake of brevity, but the final form of the 2nd-LASS corresponding to $\delta\left(\partial R/\partial F_2\right)$ is presented below for the 2nd-level adjoint sensitivity function $\mathbf{a}^{(2)}(2;t)=\left[\mathbf{a}_1^{(2)}(2;t),\mathbf{a}_2^{(2)}(2;t)\right]^{\dagger}$, $\mathbf{a}_i^{(2)}(2;t)\triangleq\left[a_{i,1}^{(2)}(2;t),a_{i,2}^{(2)}(2;t)\right]^{\dagger}$, $i=1,2$:

$$a_{1,1}^{(2)}(2;t)-2c_1t^{\lambda_1}\int_a^b a_{1,1}^{(2)}(2;\tau)\,\tau^{\lambda_1}dt-F_3a_{2,2}^{(2)}(2;t)=0; \tag{105}$$

$$a_{1,2}^{(2)}(2;t)-2c_2t^{\lambda_2}\int_a^b a_{1,2}^{(2)}(2;\tau)\,\tau^{\lambda_2}d\tau-F_3a_{2,1}^{(2)}(2;t)=0; \tag{106}$$

$$a_{2,1}^{(2)}(2;t)-2c_1t^{\lambda_1}\int_a^b a_{2,1}^{(2)}(2;\tau)\,\tau^{\lambda_1}dt=0; \tag{107}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

$$a_{2,2}^{(2)}(2;t) - 2c_2 t^{\lambda_2} \int_a^b a_{2,2}^{(2)}(2;\tau)\tau^{\lambda_2} dt = -c_2 t^{\lambda_2} . \tag{108}$$

The argument "2" in the list of arguments of $\mathbf{a}^{(2)}(2;t) = \left[\mathbf{a}_1^{(2)}(2;t), \mathbf{a}_2^{(2)}(2;t)\right]^{\dagger}$ indicates that this 2nd-level adjoint function corresponds to the first-order sensitivity $\partial R/\partial F_2$ of the response with respect to the "second" component, $F_2(\boldsymbol{\beta})$, of the "feature function $\mathbf{F}(\boldsymbol{\theta})$. In terms of the 2nd-level adjoint function $\mathbf{a}^{(2)}(2;t) = \left[\mathbf{a}_1^{(2)}(2;t), \mathbf{a}_2^{(2)}(2;t)\right]^{\dagger}$, the expression of the second-order G-differential $\delta(\partial R/\partial F_2)$ is as follows:

$$\begin{aligned} \delta(\partial R/\partial F_2) &= -c_1(\delta F_1)\int_a^b t^{\lambda_1} a_{1,1}^{(2)}(2;t)\,dt - c_2(\delta F_2)\int_a^b a_{1,2}^{(2)}(2;t)\,t^{\lambda_2}\,dt \\ &+ (\delta F_3)\int_a^b a_{2,1}^{(2)}(2;t)\,y_2(t)\,dt + (\delta F_3)\int_a^b a_{2,2}^{(2)}(2;t)\,y_1(t)\,dt \triangleq \sum_{i=1}^{3} \frac{\partial^2 R}{\partial F_i \partial F_2}\partial F_i . \end{aligned} \tag{109}$$

It follows from Equation (98) that the second-order sensitivities stemming from the first-order sensitivity $\partial R/\partial F_2$ have the following expressions:

$$\frac{\partial^2 R}{\partial F_1 \partial F_2} = -c_1 \int_a^b t^{\lambda_1} a_{1,1}^{(2)}(2;t)\,dt ; \tag{110}$$

$$\frac{\partial^2 R}{\partial F_2 \partial F_2} = -c_2 \int_a^b a_{1,2}^{(2)}(2;t)\,t^{\lambda_2}\,dt ; \tag{111}$$

$$\frac{\partial^2 R}{\partial F_3 \partial F_2} = \int_a^b a_{2,1}^{(2)}(2;t)\,y_2(t)\,dt + \int_a^b a_{2,2}^{(2)}(2;t)\,y_1(t)\,dt . \tag{112}$$

Solving the 2nd-LASS defined by Equations (105)–(108) yields the following closed-form expressions for the components of the 2nd-level adjoint function $\mathbf{a}^{(2)}(2;t) = \left[\mathbf{a}_1^{(2)}(2;t), \mathbf{a}_2^{(2)}(2;t)\right]^{\dagger}$:

$$a_{1,1}^{(2)}(2;t) = F_3(\boldsymbol{\gamma})c_2\left(t^{\lambda_2} - 2c_1 k t^{\lambda_1}\right);\; a_{1,2}^{(2)}(2;t) = 0;\; a_{2,1}^{(2)}(2;t) = 0;\; a_{2,2}^{(2)}(2;t) = c_2 t^{\lambda_2} . \tag{113}$$

Using in Equations (110)–(112) the results obtained in Equation (113) yields the following closed-form analytical expressions:

$$\frac{\partial^2 R}{\partial F_1 \partial F_2} = c_1 c_2 k F_3(\boldsymbol{\gamma});\; \frac{\partial^2 R}{\partial F_2 \partial F_2} = 0;\; \frac{\partial^2 R}{\partial F_3 \partial F_2} = c_1 c_2 k F_1(\boldsymbol{\alpha}). \tag{114}$$

The correctness of the results obtained in Equation (114) can be readily verified by using the expression of the first-order sensitivities obtained in Equations (54)–(56).

5.3. Determining the Second-Order Sensitivities Stemming from $\partial R/\partial F_3$ using Equation(53)

The second-order sensitivities stemming from the first-order sensitivity $\partial R/\partial F_1$ are obtained by G-differentiating Equation (53), which by definition yields the following expression:

$$\begin{aligned} \delta\left(\frac{\partial R}{\partial F_3}\right) &= \left\{\frac{d}{d\varepsilon}\int_a^b \left[y_1^{(0)}(t) + \varepsilon\delta y_1(t)\right]\left[y_2^{(0)}(t) + \varepsilon\delta y_2(t)\right]dt\right\}_{\varepsilon=0} \\ &= \int_a^b \delta y_1(t)\,y_2^{(0)}(t)\,dt + \int_a^b y_1^{(0)}(t)\,\delta y_2(t)\,dt , \end{aligned} \tag{115}$$

where the superscript "zero" indicates, as usual, the nominal/optimal values for the respective quantities. The variational functions $\delta y_1(t)$ and $\delta y_2(t)$ are the solutions of the 1st-LVSS defined by Equations (40) and (41). Notably, the second-order G-differential $\delta(\partial R/\partial F_3)$ does not depend on the variational function $\delta \mathbf{a}^{(1)}(t) \triangleq \left[\delta a_1^{(1)}(t), \delta a_2^{(1)}(t)\right]^{\dagger}$. Consequently, the 2nd-LASS which will correspond to $\delta(\partial R/\partial F_3)$ will be constructed by following the steps detailed in Section 3, which were employed for constructing the 1st-LASS obtained in Equations (45) and (46) for the 1st-level adjoint sensitivity function $\mathbf{a}^{(1)}(t) \triangleq \left[a_1^{(1)}(t), a_2^{(1)}(t)\right]^{\dagger}$. For the sake of brevity, these steps will not be

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 December 2024 doi:10.20944/preprints202412.0474.v1

repeated here. Performing these steps leads to the following 2nd-LASS for the 2nd-level adjoint sensitivity function $\mathbf{a}^{(2)}(3;t)=\left[a_1^{(2)}(3;t),a_2^{(2)}(3;t)\right]^{\dagger}$, where the argument "3" indicates that this 2nd-level adjoint sensitivity function corresponds to the first-order sensitivity $\partial R/\partial F_3$ of the response with respect to the "third" component, $F_3(\boldsymbol{\gamma})$, of the "feature function $\mathbf{F}(\boldsymbol{\theta})$:

$$a_1^{(2)}(3;t)-2c_1t^{\lambda_1}\int_a^b \tau^{\lambda_1}a_1^{(2)}(3;\tau)\,d\tau=y_2(t); \tag{116}$$

$$a_2^{(2)}(3;t)-2c_2t^{\lambda_2}\int_a^b \tau^{\lambda_2}a_2^{(2)}(3;\tau)\,d\tau=y_1(t). \tag{117}$$

In terms of the 2nd-level adjoint function $\mathbf{a}^{(2)}(3;t)=\left[a_1^{(2)}(3;t),a_2^{(2)}(3;t)\right]^{\dagger}$, the second-order G-differential $\delta(\partial R/\partial F_3)$ has the following expression:

$$\delta\left(\frac{\partial R}{\partial F_3}\right)=-\left[c_1\int_a^b a_1^{(2)}(3;t)t^{\lambda_1}dt\right]\delta F_1(\boldsymbol{\alpha})-\left[c_2\int_a^b a_2^{(2)}(3;t)t^{\lambda_2}dt\right]\delta F_2(\boldsymbol{\beta})\square\sum_{i=1}^{3}\frac{\partial^2 R}{\partial F_i\partial F_3}\partial F_i\,. \tag{118}$$

It follows from Equation (118) that the second-order sensitivities stemming from the first-order sensitivity $\partial R/\partial F_3$ have the following expressions:

$$\frac{\partial^2 R}{\partial F_1\partial F_3}=-c_1\int_a^b t^{\lambda_1}a_1^{(2)}(3;t)\,dt; \tag{119}$$

$$\frac{\partial^2 R}{\partial F_2\partial F_3}=-c_2\int_a^b a_2^{(2)}(3;t)t^{\lambda_2}dt; \tag{120}$$

$$\frac{\partial^2 R}{\partial F_3\partial F_3}=0. \tag{121}$$

Solving the 2nd-LASS represented by Equations (116) and (117) yields the following closed-form expressions for the components of the 2nd-level sensitivity function $\mathbf{a}^{(2)}(3;t)=\left[a_1^{(2)}(3;t),a_2^{(2)}(3;t)\right]^{\dagger}$:

$$a_1^{(2)}(3;t)=-2c_1c_2kF_2(\boldsymbol{\beta})t^{\lambda_1}+c_2F_2(\boldsymbol{\beta})t^{\lambda_2};\quad a_2^{(2)}(3;t)=c_1F_1(\boldsymbol{\alpha})t^{\lambda_1}-2c_1c_2kF_1(\boldsymbol{\alpha})t^{\lambda_2}. \tag{122}$$

Using in Equations (119)–(121) the results obtained in Equation (122) yields the following closed-form analytical expressions:

$$\frac{\partial^2 R}{\partial F_1\partial F_3}=c_1c_2kF_2(\boldsymbol{\beta});\ \frac{\partial^2 R}{\partial F_2\partial F_3}=c_1c_2kF_1(\boldsymbol{\alpha});\ \frac{\partial^2 R}{\partial F_3\partial F_2}=0. \tag{123}$$

The correctness of the results obtained in Equation (123) can be readily verified by using the expression of the first-order sensitivities obtained in Equations (54)–(56).

The following specific relations have been established in this Section:

(i) Equations (100) and (110) provide the following relation:

$$\frac{\partial^2 R}{\partial F_2\partial F_1}=-c_2\int_a^b a_{1,2}^{(2)}(1;t)t^{\lambda_2}dt=\frac{\partial^2 R}{\partial F_1\partial F_2}=-c_1\int_a^b t^{\lambda_1}a_{1,1}^{(2)}(2;t)\,dt. \tag{124}$$

(ii) Equations (101) and (119) provide the following relation:

$$\frac{\partial^2 R}{\partial F_3\partial F_1}=\int_a^b a_{2,1}^{(2)}(1;t)y_2(t)\,dt+\int_a^b a_{2,2}^{(2)}(1;t)y_1(t)\,dt=\frac{\partial^2 R}{\partial F_1\partial F_3}=-c_1\int_a^b t^{\lambda_1}a_1^{(2)}(3;t)\,dt. \tag{125}$$

(iii) Equations (112) and (120) provide the following relation:

$$\frac{\partial^2 R}{\partial F_3\partial F_2}=\int_a^b a_{2,1}^{(2)}(2;t)y_2(t)\,dt+\int_a^b a_{2,2}^{(2)}(2;t)y_1(t)\,dt=\frac{\partial^2 R}{\partial F_2\partial F_3}=-c_2\int_a^b a_2^{(2)}(3;t)t^{\lambda_2}dt. \tag{126}$$

The equalities presented in Equations (124)–(126) highlight the fact that the 2nd-order mixed sensitivities are computed twice, obtaining two distinct expressions involving two distinct 2nd-level adjoint sensitivity functions for each 2nd-order mixed sensitivity. This general property of the 2nd-FASAM-NIE Fredholm methodology enables a stringent verification of the accuracy of the computations of the various adjoint sensitivity functions when solving the respective 1st-LASS and 2nd-LASS.

 doi:10.20944/preprints202412.0474.v1

6. Discussion and Conclusions

This work has first introduced the First-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Fredholm-Type (1st-FASAM-NIE-Fredholm) for computing most efficiently the exact expressions of the first-order sensitivities of NIE decoder-responses with respect to the optimized NIE-weights/parameters. The computation of the first-order sensitivities of the decoder response requires a single "large-scale" computation for solving the 1st-Level Adjoint Sensitivity System (1st-LASS),regardless of the number of weights/parameters underlying the NIE-net. When "feature functions of parameters "can be identified within the NIE structure, the number of quadratures for computing the first-order sensitivities is smaller than the number of quadratures needed for computing the first-order decoder-response sensitivities directly with respect to the parameters, since the latter can be computed analytically and exactly by using the former (i.e., the first-order sensitivities with respect to the feature functions). The general principles underlying the 1st-FASAM-NIE-Fredholm methodology have been illustrated by using a nonlinear NIE-Fredholm net having many parameters that could be grouped within a feature function comprising three components and admitting analytical closed-form solutions for all of the illustrative results.

Building upon the 1st-FASAM-NIE-Fredholm methodology, this work has also presented the Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Fredholm-Type (2nd-FASAM-NIE-Fredholm), which enables the most efficient computation, using 2nd-Order Adjoint Sensitivity systems (2nd-LASS), of the exactly-derived expressions of the second-order sensitivities of NIE decoder-responses with respect to the optimized NIE-weights/parameters. When TF feature functions involving the TW weights underlying the NIE can be identified within the NIE-net, the 2nd-LASS needs to be solved only as many times (TF) as there are first-order sensitivities with respect to the feature functions in order to compute all of the second-order sensitivities of a decoder-response with respect to the respective the feature functions. When no "feature functions" can be constructed, the 2nd-LASS needs to be solved as many times (namely: TW -times) as there are first-order sensitivities of the decoder-response with respect to the model parameters $(TW > TF)$. The illustrative example presented in this work involved three feature functions namely $F_1(\boldsymbol{\alpha})$, $F_2(\boldsymbol{\beta})$, and $F_3(\boldsymbol{\gamma})$, but many (i.e., TW) model parameters. It was shown that only three "large-scale" computations were needed for solving the three 2nd-LASS which corresponded to these three feature functions. The second-order sensitivities with respect to the model parameters were obtained analytically, using the chain-rule of differentiation, from the sensitivities with respect to the components of the feature function. If the second-order sensitivities had been computed directly, TW large-scale computations would have been needed for their computation. In all cases, the 2nd-order mixed sensitivities are computed twice within the 2nd-FASAM-NIE-Fredholm methodology, obtaining two distinct expressions involving two distinct 2nd-level adjoint sensitivity functions for each 2nd-order mixed sensitivity. This general property of the 2nd-FASAM-NIE Fredholm methodology enables a stringent verification of the accuracy of the computations of the various adjoint sensitivity functions when solving the respective 1st-LASS and 2nd-LASS.

The 2nd-FASAM-NIE-Fredholm methodology is applicable to many scientific fields, including, in particular, the field of nuclear engineering, where the integral (Fredholm) form of the Boltzmann particle (neutron and gamma) transport equation plays a central role [13,14]. Ongoing work is aimed at developing the Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Integral Equations of Volterra-Type (2nd-FASAM-NIE-Volterra). Subsequent work will aim at generalizing these developments to address the efficient computation of exact expressions of high-order sensitivities of systems which can be modeled using Neural Integro-Differential Equations.

Funding: This research received no external funding.

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. 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; PMLR; pp. 3276–3285.
3. Ruthotto, L.; Haber, E. Deep neural networks motivated by partial differential equations. *J. Math. Imaging Vis.* **2018**, *62*, 352–364.
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. Kidger, P. On Neural Differential Equations. *arXiv* **2022**, arXiv:2202.02435.
6. 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.
7. 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; PMLR; pp. 7829–7838.
8. 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.
9. 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.
10. Rokhlin, V. Rapid solution of integral equations of classical potential theory. Journal of computational physics, 60(2):187-207, 1985.
11. Rokhlin, V. Rapid solution of integral equations of scattering theory in two dimensions. Journal of Computational Physics, 86(2):414-439, 1990.
12. Greengard, L; Kropinski M. C. An integral equation approach to the incompressible Navier-Stokes equations in two dimensions. SIAM Journal on Scientific Computing, 20(1):318-336, 1998.
13. Weinberg, A. M., Wigner, E.P. The Physical Theory of Neutron Chain Reactors, The University of Chicago Press, Chicago, Il., USA, 1958.
14. Prinja, A. K.; Larsen, E. W. General Principles of Neutron Transport. In Handbook of Nuclear Engineering, Cacuci, D. G. Ed., Ch. 5, Vol. 1, 543-642, Springer Science + Business Media, New York, NY, USA, 2010.
15. Sohrab Effati and Reza Buzhabadi. A neural network approach for solving Fredholm integral equations of the second kind. Neural Computing and Applications, 21(5):843-852, 2012.
16. Zappala, E.; de Oliveira Fonseca, A. H.; Caro, J. O.; van Dijk, D. Neural Integral Equations. arXiv:2209.15190v4 [cs.LG] 18 May 2023
17. Xiong, Y.; Zeng, Z.; Chakraborty, R.; Tan, M.; Fung, G.; Li, Y.; Singh, V. Nystromformer: A nystrom-based algorithm for approximating self-attention. In: Proceedings of the AAAI Conference on Artificial Intelligence. AAAI Conference on Artificial Intelligence, volume 35, page 14138. NIH Public Access, 2021.
18. Cacuci, D.G. Introducing the n^{th}-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.
19. Cacuci, D.G. First-Order Comprehensive Adjoint Sensitivity Analysis Methodology for Neural Ordinary Differential Equations: Mathematical Framework and Illustrative Application to the Nordheim–Fuchs Reactor Safety Model. J. Nucl. Eng. 2024, 5, 347–372. https://doi.org/10.3390/jne5030023.
20. Cacuci, D.G. Introducing the Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Ordinary Differential Equations. I: Mathematical Framework. https://doi.org/10.20944/preprints202410.1110.v1; Posted Date: 15 October 2024.
21. Cacuci, D.G. Introducing the Second-Order Features Adjoint Sensitivity Analysis Methodology for Neural Ordinary Differential Equations. II: Illustrative Application to Heat and Energy Transfer in the Nordheim-Fuchs Phenomenological Model for Reactor Safety. https://doi.org/10.20944/preprints202410.1118.v1. Posted Date: 15 October 2024.