

**Journal of  
Mechanics of  
Materials and Structures**

**Volume 8, No. 5-7**

**July–September 2013**



# JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES

msp.org/jomms

Founded by Charles R. Steele and Marie-Louise Steele

## EDITORIAL BOARD

ADAIR R. AGUIAR University of São Paulo at São Carlos, Brazil  
KATIA BERTOLDI Harvard University, USA  
DAVIDE BIGONI University of Trento, Italy  
IWONA JASIUK University of Illinois at Urbana-Champaign, USA  
THOMAS J. PENCE Michigan State University, USA  
YASUhide SHINDO Tohoku University, Japan  
DAVID STEIGMANN University of California at Berkeley

## ADVISORY BOARD

J. P. CARTER University of Sydney, Australia  
R. M. CHRISTENSEN Stanford University, USA  
G. M. L. GLADWELL University of Waterloo, Canada  
D. H. HODGES Georgia Institute of Technology, USA  
J. HUTCHINSON Harvard University, USA  
C. HWU National Cheng Kung University, Taiwan  
B. L. KARIHALOO University of Wales, UK  
Y. Y. KIM Seoul National University, Republic of Korea  
Z. MROZ Academy of Science, Poland  
D. PAMPLONA Universidade Católica do Rio de Janeiro, Brazil  
M. B. RUBIN Technion, Haifa, Israel  
A. N. SHUPIKOV Ukrainian Academy of Sciences, Ukraine  
T. TARNAI University Budapest, Hungary  
F. Y. M. WAN University of California, Irvine, USA  
P. WRIGGERS Universität Hannover, Germany  
W. YANG Tsinghua University, China  
F. ZIEGLER Technische Universität Wien, Austria

**PRODUCTION** production@msp.org

SILVIO LEVY Scientific Editor

Cover photo: Wikimedia Commons

---

See [msp.org/jomms](http://msp.org/jomms) for submission guidelines.


---

JoMMS (ISSN 1559-3959) at Mathematical Sciences Publishers, 798 Evans Hall #6840, c/o University of California, Berkeley, CA 94720-3840, is published in 10 issues a year. The subscription price for 2013 is US\$555/year for the electronic version, and \$705/year (+\$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues, and changes of address should be sent to MSP.

---

JoMMS peer-review and production is managed by EditFlow® from Mathematical Sciences Publishers.

PUBLISHED BY

 **mathematical sciences publishers**  
nonprofit scientific publishing

<http://msp.org/>

© 2013 Mathematical Sciences Publishers

## EFFICIENCIES OF ALGORITHMS FOR VIBRATION-BASED DELAMINATION DETECTION: A COMPARATIVE STUDY

OBINNA K. IHESIULOR, KRISHNA SHANKAR, ZHIFANG ZHANG AND TAPABRATA RAY

The need for efficient and low-cost techniques adequate for damage detection has become of great interest in engineering applications where structural health monitoring (SHM) is of paramount importance. Promising algorithms for SHM have to deliver results with very low computational and response time requirements and be trustworthy within a certain accuracy. Different algorithms (artificial neural networks (ANN), response surface methodology (RSM), and optimization techniques — gradient-based local search (GBLS) and nondominated sorting genetic algorithms (NSGA-II)) are proposed to fill this research gap. The concept of a surrogate model as a fast-executing model is also introduced. Because the objective of this paper is to concentrate on viable techniques suitable for damage detection using vibration methods with very low computational requirements, surrogates are therefore employed to curtail the computational expense. Particularly of interest among the proposed algorithms is RSM, the principle of which has proved successful in the pharmaceuticals industry over the years. However, RSM has not been so widely used in the field of structural engineering for delamination detection. In this paper, we have demonstrated that a fourth-order polynomial has the capability to detect delaminations in composite structures. In order to reduce the size of training data required to solve the inverse problem by the proposed algorithms, the idea of a suitable design space is brought to the limelight as the combination of all possible simulations that one is concerned about. Since the overall sum of design space is usually prohibitively large, we have used  $K$ -means clustering to effectively achieve this. This research concerns the application of ANN, RSM, and optimization techniques for delamination detection using changes in natural frequencies before and after damage. Efficiencies of algorithms (ANN, GBLS, and NSGA-II) are compared with the developed RSM models in terms of the accuracy of delamination detection and response time requirements. The methods have been shown to compete effectively for delamination detection and are accurate in detecting the size and locations of delaminations at midplanes. RSM has a unique feature in that it produces models with a small training dataset requirement and also generates mathematical models that are easy to interpret and implement. The optimization techniques, when integrated with surrogate models, require small training sets clustered through the entire design space. ANN, however, requires large training datasets to achieve its results. As such, the potential of these algorithms as tools for on-board damage detection when integrated into a SHM system is successfully demonstrated.

### 1. Introduction

The utilization of in-service composite structures can be affected by degradation resulting from prolonged use and exposure to harsh environmental conditions. This is because composite structures are susceptible

---

Obinna K. Ihesiulor is grateful to the University of New South Wales at the Australian Defense Force Academy for provision of tuition fee waiver and living allowance support.

*Keywords:* vibrations, delaminations, ANN,  $K$ -means clustering, surrogates, RSM, optimization techniques.

to hidden or barely visible damage caused by impacts, design errors, overheating, loading abrasion, and fatigue, that if unchecked results in fiber breakage, matrix cracking, and delaminations [Chakraborty 2005]. These factors can cause severe consequences for in-service structures with regards to higher life-cycle costs, low structural reliability, and loss of operational capability causing loss of lives and property [Kessler et al. 2005]. Hence, early and accurate detection and monitoring of structural failures such as delaminations is a principal concern.

Delamination detection has therefore gained much attention from the structural engineering community because unpredicted delamination damage may cause catastrophic failures [Zheng et al. 2011]. Hence, the need to avoid delamination failure by providing a reliable and effective nondestructive damage-identification technique is crucial to maintaining the safety and integrity of structures. Damage detection is of paramount importance because it is the most vital subsystem of structural health monitoring (SHM). Damage detection can be achieved via any of various methods such as visual inspection, nondestructive evaluation methods, and solution of inverse algorithms. In the field of medicine, for example, a doctor regularly monitors a patient's blood pressure to determine the health of the patient by observing deviations using sophisticated equipment. Similarly, engineers monitor the integrity of a structure by measuring changes in the responses of the structure which can lead to structural failure. Taking immediate responses in both scenarios can avoid catastrophic implications [Mufti 2001]. An effective SHM system involves the use of expertise in many disciplines, giving rise to solutions of the multidisciplinary problem involving damage detection modeling systems via finite element analysis, optimization methods, structures and materials, computers, communication and electronics, real-time controllers, intelligent processing, and so on. The aim of SHM is not just to detect structural failure, but also to provide an early indication of damage. The early warning provided by an SHM system can then be used to define remedial strategies before the structural damage leads to failure.

SHM is therefore the key to securing confidence in the utilization of fiber reinforced polymers (FRP) in engineering applications. The unique advantages of FRP composite materials have made them preferred in SHM applications. Previously, structures were monitored by carrying measuring devices to the site each time a set of readings was required. Nowadays, by using vibration measurements for structural health monitoring, structures can be monitored from time to time to ensure that they are in good condition by obtaining an extensive amount of processed data off-site. In this work, delamination detection in composite laminates is evaluated in the context of structural health monitoring, which essentially is a reliable system with the ability to detect and interpret adverse changes in a structure as a result of damage. The motivation driving SHM is that knowing the integrity of in-service structures on a continuous real-time basis is essential to structural engineers. Monitoring and evaluating the integrity, in-situ behavior, and health condition of a structure accurately and efficiently while it is in service optimizes resources for repair/replacement, reduces downtime while increasing productivity, and ensures the safety of lives and property. As a result, SHM enables avoidance of catastrophic failure through early detection of problems, optimal use of structures, prevention of regular shutdowns of in-service structures, and inspection of hard-to-reach places. SHM also replaces periodic maintenance with long-term maintenance schedules. This reduces maintenance labor and minimizes human involvement, thus improving safety and reliability [Kessler et al. 2005].

Most nondestructive testing techniques such as ultrasonic inspection, thermography, optical holography, and mechanical impedance for delamination identification in composite laminates [Buynak et al.

1989; Cantwell and Morton 1992; Broughton et al. 2000; Rao 2007] cannot be used for real-time and online damage detection [Doebbling et al. 1998]. Moreover, most of these techniques are mainly applicable for the inspection of limited areas of a structure locally and hence are labor-intensive, time-consuming, and cost-ineffective when considering large structures [Dackermann 2010]. Vibration-based monitoring is a viable method to overcome these limitations. This is because of the reliability of its measurements, ease of implementation, and relative cost competitiveness [Cawley and Adams 1979; Kim and Yiu 2004]. In addition, utilization of a vibration-based monitoring tool can provide fast in-situ and real-time monitoring [Doebbling et al. 1998]. SHM exploiting vibration measurements are global methods based on the principle that degradation due to damage in a structure changes its vibration parameters, namely, its natural frequencies, mode shapes, and damping characteristics. It is hence feasible to use any one measured vibration quantity to characterize and identify the presence of damage via an inverse modeling. The choice of the natural frequencies as one of the commonly used vibration parameters is attractive because the natural frequencies can be conveniently measured and determined easily from just a few accessible points on the structure and are usually less contaminated by experimental noise [Fang and Tang 2005]. Natural frequency-based methods use the natural frequency changes before and after damage as the basic feature for damage identification in solving the inverse problem.

Reviews of vibration-based health monitoring methods utilizing artificial neural networks (ANN) and optimization algorithms appear in previous works [Islam and Craig 1994; Okafor et al. 1996; Doebbling et al. 1998; Valoor and Chandrashekhara 2000; Harrison and Butler 2001; Nag et al. 2002; Watkins et al. 2002; Chen et al. 2004; Su et al. 2005; Addin et al. 2006; Zheng et al. 2011]. Although previous works have demonstrated the feasibility of ANN and optimization algorithms for delamination damage detection, some limitations are worth highlighting. Most of the damage detection methods that have been reviewed attempt to identify delamination by solving an inverse problem, which often requires the construction of numerical models. Numerical simulations have therefore become increasingly useful for studying the performance of engineering structures. However, because of the large scale of many damage-identification problems, it is often not feasible to conduct many experiments to explore all damage scenarios. Tremendous amounts of computing time are required to run these simulations and oftentimes computing power is simply not available to conduct such complex simulations due to budget and time considerations. This dependency on firsthand numerical models, which are computationally expensive, makes these approaches unpromising for SHM.

Using a small number of datasets, surrogate models can be built to efficiently explore the entire design space to determine areas of interest while lowering computational expense. Essentially, surrogates are models that will run in seconds on a single processor in contrast to the hours that it may take to run a more detailed analysis on a multiprocessor machine. On the other hand, one of the remarkable demerits of ANN is the requirement of large amounts of training data. ANN has, however, demonstrated undoubted efficiency for complex damage-detection schemes by seeking to discriminate between damaged and undamaged specimens, and has been widely employed for pattern recognition, classification, function approximation, signal processing, and damage identification [Addin et al. 2006]. ANN offers capabilities such as self-adaptiveness, generalization, abstraction capabilities, and suitability for real-time applications [Tsoukalas and Uhrig 1997]. These features make ANNs powerful tools for vibration-based damage identification. But the computational power and size of training data required to solve the inverse problem by ANN is usually exceedingly large. To fill this gap, the idea of *K*-means clustering algorithm

is introduced. Since the overall sum of design space is usually This algorithm smartly determines which numerical simulations should be run when resources are scarce and models are fit to these smartly chosen data points.

The content of this research focuses on SHM as a diagnostic tool for delamination damage detection via the utilization of efficient inverse algorithms with significant interest in their prediction-error quantification. In this aspect of SHM, the abstract data numbers are converted into quantities that relate directly to the responses of a structure, that is, the natural frequency measurements will be converted into quantities of delamination parameters via inverse algorithms. The key objective of the present study is therefore to provide a comparative analysis of different algorithms for delamination detection. Based on finite element simulation of a composite beam-type structure, a comparative study of several damage-identification algorithms is provided to illustrate the validity and effectiveness of the algorithms. The ANN adopted in this work is a response surface approximation method that is based on the concept of artificial intelligence. The ANN is to provide surrogate models to computationally expensive finite element (FE) models. The motivation for this is derived from the concept that constructing a surrogate model to approximate any expensive function can substantially reduce the computational cost for objective function evaluations during the course of optimization and improve the optimization search performance. Thus, once a neural network is effectively trained, it is capable of being used for future interpolation and approximation.

For the response surface methodology (RSM), a simple fourth-order polynomial model that associates input parameters to the output response is used for delamination prediction. RSM approximates the output of a given system as a function of some input variables (design variables) by solving a system of nonlinear equations. This method is effectively employed as an inexpensive low-order approximation model for delamination prediction and, because of the model form, minimal effort is required to build the model. The method also has relative flexibility in the range of problems it is able to model, unlike the ANN, which requires a huge effort to train its model.

Another advantage of RSM is the ease of implementation in damage-identification settings. The selection of the sampling points for building ANN surrogate models is vital and demanding because the prediction capabilities of an approximation function are highly influenced by the sampling points in the given design space [Kanungo et al. 2002]. The *K*-means clustering technique is used so as to ensure that the sampling points are evenly distributed over the design space. This method gives a systematic and efficient means of analyzing the complete design space. It explores the high-dimensional design space and screens the most clustered design points corresponding to the set of design variables.

To solve the optimization problems, local and global optimization algorithms are employed for delamination detection. Gradient-based optimization schemes are local optimizers that can get stuck at the first optimum obtained during the search process. When applied to continuous problems, this algorithm gives better performance than any other optimization scheme but is highly unsuitable when searching for global optima when there is a mixture of discrete and continuous variables [Schittkowski 1986]. Global optimizers like the nondominated sorting genetic algorithm (NSGA-II) are most promising in both discrete and continuous problems due to their robust and random nature of search but with a significantly high computing cost. The high computing cost involved in deducing the function objectives is greatly reduced using surrogate models, and hence taking advantage of their global optimization behavior.

## 2. Mathematical formulation

**2.1. Mathematical formulation of the optimization problem.** The goal of an optimization task is to ascertain a set of values that result in a maximum or minimum of a function called the objective function. The objective function is a mathematical expression describing a relationship of the optimization parameters. Optimization is an activity (which could be single or multiobjective) that aims at finding the best (that is, optimal) solution to a problem. Single-objective optimization is scalar valued with a single unique solution whereas when the objective is vector valued, the optimization process is referred to as multiobjective [Ray et al. 2001]. In this delamination detection optimization problem, the objective function is defined as a single objective function and two key components are effectively required to solve the optimization problem:

- the simulator (which essentially computes/simulates the natural frequencies) and
- the comparator (objective function).

The optimization objective is to compare and minimize the errors between measured (actual) and predicted natural frequencies. The percentage change in frequency caused by delaminations for  $i$ -th mode ( $dF_i$ ), where  $i = 1, \dots, n$ , is defined as

$$dF_i = \frac{F_{ui} - F_{di}}{F_{ui}} \times 100, \quad (1)$$

where  $F_{ui}$  and  $F_{di}$  are the numerically predicted natural frequencies of the undamaged and damaged composite beam, respectively.

The objective function (comparator) is the norm of the difference between the measured ( $dF_{Mi}$ ) natural frequencies for which the delamination parameters are to be determined and the numerically predicted natural frequencies ( $dF_i$ ). All the computational analysis for the delaminated natural frequencies (in Hz) is defined with respect to the beam in perfect condition and the deviation is expressed as a percentage change in the natural frequencies.

The error,  $E_i$ , between the measured and predicted shifts in frequencies due to the delamination for the  $i$ -th mode is given by

$$E_i = \left( \frac{dF_i - dF_{Mi}}{dF_{Mi}} \right)^2, \quad (2)$$

where  $dF_{Mi}$  is the measured percentage change in the natural frequencies for which the delamination parameters are to be ascertained.

The use of high-fidelity simulation tools to compute the natural frequencies of composite structures for the objective function always comes with an unavoidable computing time. For this reason, surrogate models are a key element to reduce the optimization cycle time by providing alternative function evaluations for the objective function. The objective function with surrogate,  $\text{Obj}_S$ , is therefore given by the sum of the errors

$$\text{Obj}_S = \sum_{i=1}^n E_i = E_1 + E_2 + E_3 + \dots + E_n. \quad (3)$$

Similarly, the objective function without surrogate,  $\text{Obj}$ , is given by

$$\text{Obj} = \sqrt{(dF_1 - dF_{M1})^2 + \dots + (dF_n - dF_{Mn})^2}, \quad (4)$$

where  $n$  is the maximum  $i$ -th mode.

In order to evaluate the objective function, the natural frequencies of the undelaminated and delaminated composite beams are simulated via finite element analysis (FEA). The objective function is evaluated using frequency shifts due to the delamination predicted using FEA and the measured frequency shifts. Very few simulations were used to build the database using  $K$ -means clustering and then surrogate-assisted optimization.

**2.2. Solution strategy.** An overview of the approach is shown in Figure 1.

A detailed description of the proposed methods to detect delaminations can be summarized as follows:

*Optimization without surrogates.* Step 1: Develop an FE model with ANSYS that computes the natural frequencies before and after damage. This is also known as the simulator. Step 2: Engage the simulator directly with the optimizer to evaluate the objective function to be minimized to determine the delamination parameters. The objective function, also known as the comparator, essentially minimizes the sum of errors between the simulated natural frequency changes before and after damage and the actual ones to determine the delamination parameters for any number of variables ( $N_d$ ).

*Surrogate-assisted optimization.* Step 1: Develop an FE model with ANSYS that computes the natural frequencies before and after damage. Step 2: Use the simulator to generate a database in the case of optimization with surrogates. Step 3: Reduce the size of the generated database, which is usually large, by a  $K$ -means clustering method. This ensures that a small number of well-clustered datasets within the entire design space is used for delamination prediction. Step 4: Use the  $K$ -means clustered datasets to create a surrogate model. Step 5: Engage the surrogate model directly with the optimizer to evaluate the objective function to be minimized to determine the delamination parameters for any number of variables ( $N_d$ ).

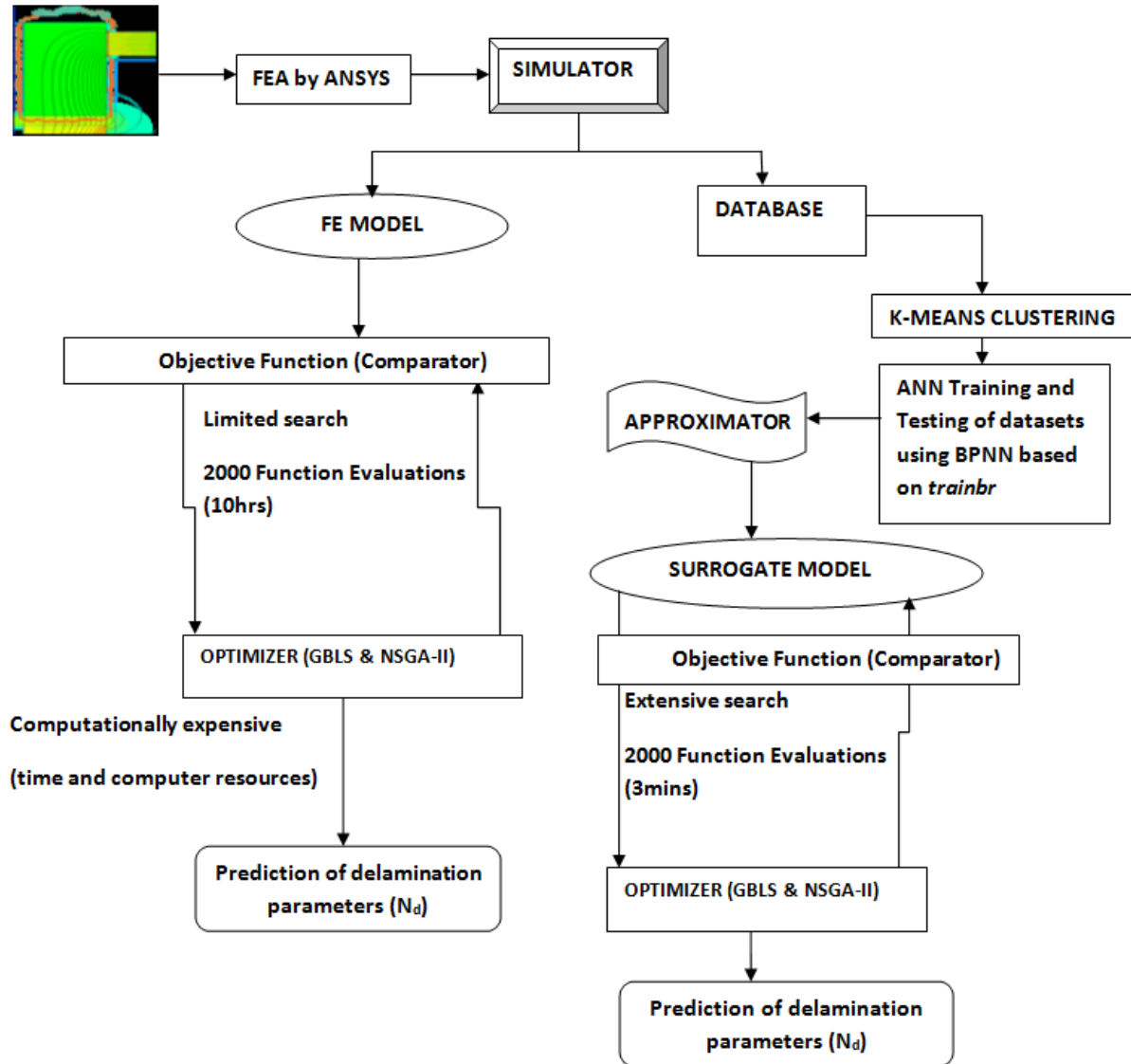
*Direct solution via ANN and RSM.* Step 1: Develop an FE model with ANSYS that computes a database of natural frequencies before and after damage. Step 2: Do a  $K$ -means clustering if the size of the database is to be reduced. Step 3: Use the database to train ANN and RSM models that give the delamination prediction for any number of variables ( $N_d$ ).

**2.3. Modeling of the laminated composite beam.** We studied the vibration behavior of eight-layer  $[0^\circ/90^\circ/90^\circ/0^\circ]$ s glass-epoxy laminated composite cantilever beams with and without delaminations. Glass fiber (E-glass) is used as reinforcement in the form of unidirectional fibers with epoxy resin as matrix for the composite beam. The laminates are reinforced unidirectionally. The material properties are given in Table 1 for the FE analysis and theoretical model. The composite beam has length  $L = 267$  mm, width  $W = 25.4$  mm, and thickness  $h = 1.778$  mm.

$E_1$	$E_2$	$G_{12}$	Poisson's ratio ( $\nu_{12}$ )	Density ( $\rho$ )
42.34 GPa	11.72 GPa	3.0025 GPa	0.27	1901.5 kg/m <sup>3</sup>

**Table 1.** Material properties of the composite beam laminates.





**Figure 1.** Schematic of solution methodology.

Considering through-width delaminations, delaminations are simulated extending through the width of the beam, with locations and sizes in the ranges  $0 < X \leq 70$  and  $0 < a \leq 58$ , respectively, satisfying the requirement that delamination must not extend outside the beam. For any delamination pattern  $([Z, X, a])$ ,  $Z$  denotes the interface, and the normalized delamination location is expressed as  $X = X_{\text{actual}}/L$ , where  $X_{\text{actual}}$  is the distance from the middle of the delamination to the fixed end of the beam and  $L$  is the total beam length. Similarly, the normalized delamination size is given as  $a = a_{\text{actual}}/L$ , where  $a_{\text{actual}}$  is the length of the delamination along the axis of the beam. Delaminations are simulated at different interfaces,  $Z = 1$  to  $4$ ;  $1$  being the midplane and  $4$  the outermost interface.

**2.3.1. Finite element (FE) analysis of the composite beam.** Herath et al. [2010] investigated the effects of delamination size and location in composite beams using a FEA model under quasistatic loading and an analytical Euler-beam model. Ramanamurthy and Chandrasekaran [2011] used FE modeling to develop a damage-detection method in a composite cantilever beam with an edge crack. Ishak et al. [2001] employed the application of the strip element method and adaptive multilayer perceptron networks (MLP) for inverse identification of interfacial delaminations in carbon/epoxy laminated composite beams. In this paper, FEA is used to solve the forward problem and generate data for frequency shifts for known delamination parameters. Numerical analysis is carried out using the commercial finite element program ANSYS 12.1 to build the FE models for both the undelaminated and delaminated glass fiber reinforced composite beams to investigate their vibration behavior. Analysis is carried out on a three-dimensional eight-node layered solid element (SOLID185) with three degrees of freedom at each node. The shell section is adopted to define the layer information.

In the FE model for the delaminated beam, the delaminated beam was modeled as two volumes, separated along the interface at which the delamination is located. The nodes situated along the interface of undelaminated segments were merged together while nodes in the interface of the delaminated area are left unmerged [Zhang et al. 2010]. Contact elements (TARGE170/CONTAC173) were introduced between the delaminated surfaces to prevent separation and interpenetration, so the upper and lower sublaminates had the same deflection, acting as two separate beams constrained to move together. To build the model, the number of elements required to provide acceptable levels of accuracy and also determine how fine the mesh should be in order to get convergence of the numerical results without excessive use of computational time was determined using a convergence analysis. The final FE mesh employed had 200 elements along the length, six elements across the width, and one element for each layer of the eight-ply laminate.

The natural frequencies obtained from the undelaminated model are consistent with those from theory. To determine the natural frequencies of the delaminated beam, ANSYS Workbench batch mode simulation was used to setup the eigenvalue modal analysis. Further, using the block Lanczos method, the natural frequencies of the first eight bending modes were extracted, discarding the torsion and in-plane bending modes. The natural frequencies of the undamaged and damaged glass fiber/epoxy composite beam were computed. The natural frequencies obtained from the FE model have been compared to theoretical results for validation purposes.

**2.3.2. Theoretical modeling of laminated composite beam.** The theoretical formulation of the vibration characteristics of the models of beams with and without delaminations were based on the pioneering works in this field. Ramkumar and Kulkarni [1979] were the first of these pioneers. They developed a simplified model to compute the free vibration frequencies of a cantilever laminated beam with a single through-the-width delamination at the interlaminar position. Their basic concept was to deduce mathematically the actual vibration properties of four different Timoshenko beams combined together by considering the delaminated and undelaminated portions of the beams. Their mathematical model of the eigenvalue problem fulfilled all the necessary boundary and continuity conditions between the adjoining beams. They also carried out experimental studies, which showed that the analytical computations of the natural frequencies were uniformly lower than the experimental ones. As a follow up, Wang and Liu [1982] employed the analytical model of [Ramkumar and Kulkarni 1979] using classical Bernoulli–Euler

[Z, X, a]	Mode	Theory ( $F_{ui}$ )	Theory ( $F_{di}$ )	FE ( $F_{ui}$ )	FE ( $F_{di}$ )	% diff. ( $F_{ui}$ )	% diff. ( $F_{di}$ )
[1, 54, 24]	1	16.06	15.97	16.04	15.94	0.159	0.14
	2	100.67	100.54	100.46	100.33	0.211	0.2
	3	281.9	252.38	281.04	251.49	0.349	0.3
	4	549.4	522.12	550.07	52.81	0.443	0.12
	5	913.09	780.59	907.91	776.82	0.482	0.57
	6	1364.04	1192.95	1353.75	1184.85	0.679	0.75
	7	1905.15	1636.83	1886.64	1624.22	0.771	0.97
	8	2536.44	2191.25	2505.46	2170.49	0.947	1.22
[1, 62, 18]	1	16.06	16.032	16.039	16.007	0.154	0.14
	2	100.67	100.333	100.458	100.120	0.213	0.2
	3	281.9	275.044	281.042	274.152	0.324	0.3
	4	549.4	504.202	550.068	501.887	0.459	0.119
	5	913.09	894.735	907.913	889.365	0.600	0.57
	6	1364.04	1199.253	1353.748	1191.622	0.636	0.75
	7	1905.15	1725.247	1886.636	1709.563	0.909	0.97
	8	2536.44	2291.992	2505.461	2271.168	0.909	1.22

**Table 2.** Quantitative comparison between the analytical and numerical natural frequencies (in Hz).

beam theory to obtain consistent results by incorporation of the coupling between flexural and axial vibrations of the delaminated sublaminates of the model. Then later on, Mujumdar and Suryanarayan [1988] applied a pressure distribution between two respective delaminated sections to impose a constraint between the two beam sections in order to obtain similar flexural deformation. This kind of proposed model was termed a “constrained model” in contrast with the so-called “free model” proposed in [Wang and Liu 1982]. Their analytical model, demonstrated for isotropic materials, was found to have the ability to determine the natural frequencies of a delaminated beam at any interface. Della and Shu [2005] extended the model of [Mujumdar and Suryanarayan 1988] to composite beams, by using the effective bending stiffness terms of composite laminates. In what could be termed a similar approach, Pardoen [1989] developed a constrained model to predict the natural frequencies of a simply supported composite beam considering only midplane delaminations.

In this analytical study, the model first adopted by [Ramkumar and Kulkarni 1979] is customized for composite laminates, to determine the changes in natural frequencies due to delaminations located at different interfaces.

**2.4. Validation of the FE analysis.** A comparison of the analytical and ANSYS numerical results for undelaminated ( $F_{ui}$ ) and delaminated ( $F_{di}$ ) beams is shown in Table 2. It is observed that the natural frequencies decrease as delaminations occur and the percentage errors increase for the higher modes. The objective of this exercise is to verify the results of the numerical model. The numerical results are consistent with the analytical results with errors less than 1%. Numerical modeling is preferred to analytical modeling particularly in complex systems because it is relatively easy to incorporate different boundary conditions, loading configurations, etc.

### 3. Algorithms for solution of the inverse problem

**3.1. Optimization search algorithms.** Global search methods are search schemes that are not prone to being trapped by local optima. They require huge amounts of computational effort. Local search methods, on the other hand, restrict candidate potential solutions to a confined design space starting at an initial guess [Shang et al. 2001]. Thus, a combination of these is an attractive choice, which has led to the development of the *memetic algorithm*.

**3.1.1. Gradient-based local search.** The gradient-based local search (GBLS) method is a kind of optimization algorithm that basically employs gradient information of the objective function for determining the direction of the subsequent search points from a given start point. This technique uses a function which seeks the minimization of a scalar (single) objective function of multiple design variables within a region specified by linear constraints and bounds using the sequential quadratic programming (SQP) algorithm.

The SQP is the most famous gradient-based algorithm. The SQP method is most successful specifically for problems of nonmultimodal behavior. The GLBS used in this work is based on SQP. SQP methods embody the state of the art in nonlinear programming methods. An overview of SQP is found in [Schittkowski 1986]. In this method, the function solves a quadratic programming subproblem at each iteration by obtaining and updating iteratively at every step an estimate or approximation of the Hessian of the Lagrangian function using a quasi-Newton updating method based on the Broyden–Fletcher–Goldfarb–Shanno formula [Schittkowski 1986]. The method uses iteration from an initial guess until it reaches a feasible local optimum. During the process of every iteration, quadratic nonlinear programming is used to solve the objective at that point. If the problem is unconstrained, then the method reduces to Newton’s method for finding a point where the gradient of the objective disappears. The solutions of the quadratic programming are used to initiate a search space towards a better solution until the optimum set is found.

However, the GLBS can be trapped in local optima and hence a reinitialization is used to avoid stagnation. Reinitialization is a process adopted to prevent to some extent the trapping of the technique at the local optimum by starting the optimization process from different start points. The final optimum may not be the global optimum, but this process effectively increases the optimization cycle time [Shang et al. 2001]. This technique has been implemented in MultiStart from the MATLAB Optimization Toolbox.

**3.1.2. Global optimizer based on evolutionary algorithms.** To ensure global optimum results and increase the accuracy of results while reducing the search time via surrogates, biologically inspired evolutionary algorithms (EA) are used. An EA is a kind of algorithm that encodes (both numerical and nonnumerical) design parameters to shuffle multiple candidates via parallel and interactive search. During the search process, there is firstly a selection performed based on survival of the fittest. To generate the next generation of possible candidate solutions, some parameter values are exchanged between two candidates (crossover) and new values introduced (mutation). EAs cannot be easily trapped in local minima or maxima as a result of crossover and mutation operations, which makes them an ideal method to effectively handle multimodal optimization problems.

EAs are therefore representatives of the class of stochastic (random) and robust optimization algorithms that do not require gradient information during the course of an optimization process but rather use an objective function value, which makes them more computationally expensive than the GBLS, which requires fewer iterations. However, they can handle a wide variety of problem characteristics such

as discrete and continuous design variables and multimodality, in contrast with the GBLs. Surrogates are used to contain their computationally expensive nature, hence exploring the advantage of their global optimization behavior.

The choice of an efficient optimization algorithm is based on the problem under study. For our optimization problem in which the variables are a combination of discrete and continuous, it is difficult to use conventional optimization algorithms such as the gradient-based method to obtain the global optimum, since they rely on the use of continuous variables. In an attempt to solve such discrete problems, a population-based stochastic algorithm, and in particular a nondominated sorting genetic algorithm (NSGA-II) [Deb et al. 2002], are found as an ideal choice to effectively manipulate the optimization task for detecting delaminations in composite laminates as they are the most promising in both discrete and continuous multimodal problems. While an evolutionary algorithm has been used in the current study, other forms of evolutionary algorithms such as particle swarms, differential evolution (DE), etc., can also be used. DE is peculiar to solving single objective optimization problems with continuous variables. Several DE variants exist such as the DE/rand/1/bin strategy and self-adaptive DE.

The parameters of NSGA-II and their values are listed in Table 3 while its flowchart is depicted in the authors' previous work [Ihesiulor et al. 2012a] (see Figure 1). The detailed description of NSGA-II processes includes:

- (1) Generation of initial population ( $G$ ): Generation of the first parent population of size  $N$ . This is randomly generated within the predefined feasible region (the upper and lower bounds of the design space).
- (2) Nondominated sort: Individual population are evaluated and sorted based on nondomination. A solution ( $s_1$ ) dominates (is preferred to) another solution ( $s_2$ ) if and only if  $s_1$  is better than  $s_2$  in the objective function. For every generation, fast nondominated sorting is applied to identify nondominated solutions to construct the nondominated front. This produces a set of candidate solutions that are nondominated by any individual in the population. These solutions are then discarded from the population temporarily until the next-best nondominated set is identified. This process goes on and on until all solutions are classified and assigned ranks equal to their nondomination level assuming fitness minimization.
- (3) Crowding distance: This is basically the niche safeguarding in the design objective space. The crowding distance assigned equals the front density in the neighborhood (the distance of each candidate solution from its nearest neighbors).
- (4) Selection by ranking: The initial population is sorted in ascending order according to fitness functions. Potentially better solutions are ranked higher than the worst solutions. Individuals are selected by the use of a binary tournament selection with the crowded comparison operator.
- (5) Genetic operators: A mating pool is formed by bonding of the parent and child populations. The simulated binary crossover operator and polynomial mutation are used to create an offspring of new population. Mutation occurs by random walk around individuals. The best population of parents and offspring with higher fitness is selected to reproduce the next generation.
- (6) Recombination and selection: Combination of the traits of the offspring population and parent population is done to reproduce the extended population of the next generation ( $2N$ ) [Ayala and dos

Maximum population size	$N = 200$
Crossover probability	$P_c = 0.9$
Mutation probability	$P_m = 0.1$
Maximum number of exact function evaluations	4000
Evolutionary operators (binary tournament selection)	$s = 2$
Evolutionary operators (polynomial mutation)	$\eta_m = 10$
Evolutionary operators (simulated binary crossover)	$\eta_c = 10$
Evolutionary operators (elitism, nondomination rank, crowding distance)	80
Number of independent runs (stochastic and can be run more than once)	$\geq 10$

**Table 3.** NSGA-II parameters.

Santos Coelho 2008]. This is done to ensure elitism and to keep diversity in generating subsequent successive populations [Sastry et al. 2006]. The replacement criteria keeps the best among the parents and offspring based on the fast nondominated sorting and maintains the best diversified individuals to provide a larger search space. The new population size with the initial population is filled with individuals from the sorting fronts starting from the best. The crowding distance method is recalled to maintain diversity if a front incompletely fills the next generation. This ensures that convergence in one direction does not take place. This process repeats itself to generate subsequent generations until a stopping criterion is reached.

**3.2. Database creation for ANN training using *K*-means clustering algorithm.** The ANN model and its results are highly dependent on the training dataset provided. It is hence necessary to maintain the diversity of the training set to obtain a good prediction model by ensuring that the training data is not clustered around one part of the design domain. The most vital strategy in the selection of the training datasets is to find the ideal set that is a true representation of all the possible samples in the total design space. This is done using *K*-means clustering.

*K*-means clustering is used for grouping large datasets into smaller sets called clusters. The number of *k*-groups or objects needed is specified. Each object is represented by some feature vector in *n*-dimensional space, *n* being the number of all characteristics used to describe the objects to cluster. The algorithm then randomly chooses *k*-points in that vector space. These points serve as the initial centers of the clusters and all objects are assigned to the center they are closest to [MacQueen 1967]. Basically, the *K*-means clustering algorithm finds a subset of *k*-groups known as centroids that minimizes the mean squared distance from each data point to its nearest center in an entire dimensional space of *n*-data points [Kanungo et al. 2002]. The algorithm is implemented in the following steps:

- Step 1: Specify *k*-points into the space represented by the objects that are being clustered. These points represent the initial group centroids.
- Step 2: Each object is allocated to the group with the closest centroid.
- Step 3: When all objects have been assigned, the positions of the *k*-centroids are recomputed.
- Steps 2 and 3 are iterated until the centroids no longer move. This introduces a separation of the objects into groups from which the metric to be minimized can be computed.

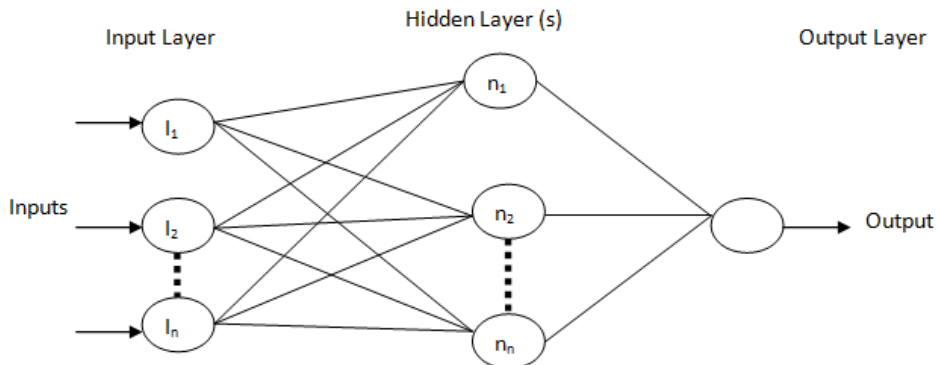
**3.3. Introduction to ANN as an inverse solver and surrogate creator.** Artificial neural networks (ANNs) have evolved as one of the promising artificial intelligence concepts used in real-world applications. ANN have been extensively used in structural engineering applications in the areas of failure prediction, delamination identification, crack detection, as well as others [Mahdi and El Kadi 2007]. ANN is a very powerful interpolator that can be used to map functions and derive a relationship between a set of input parameters and their output responses.

There are different types of ANN architectures, namely, multilayer perceptron (MLP), radial basis function (RBF), and so on. The MLP type [Delashmit and Manry 2005] is adopted in this study because it effectively provides a complex nonlinear mapping between the input and output variables. MLPs are feed-forward nets with one or more hidden layers between the input and output neurons as shown in Figure 2. ANN based on MLP is trained using a back propagation neural network (BPNN) algorithm, a gradient-based method that has emerged as successful in the training of multilayered neural nets using supervised learning. In supervised learning, the network learns using input and output data and provides an approximation of the functional mapping between the two.

A typical BPNN is based on the fact that a feed-forward neural net (FFNN) with at least one hidden layer can approximate any continuous nonlinear function with arbitrary accuracy the number of hidden neurons is sufficient. MLP is a FFNN which consists essentially of an input layer with several neurons (depending on the number of inputs,  $dF_i$  to  $dF_n$ ), a layer of output neurons and one or more layers of hidden neurons that wholly perform the application expected objective. The primary building blocks of MLP are the artificial neurons or processors. The neurons in each layer are fully interconnected to the preceding and subsequent layers. Every one of them is connected by adjustable associated weights to enable the network to map complex associations between the input and output data. The activation function ( $a$ ) for each neuron is defined as the summation of all the inputs multiplied by their connection weights and biases ( $wt_1, wt_2, \dots, wt_n$ ), given as

$$a = dF_1wt_1 + dF_2wt_2 + \dots + dF_nwt_n. \quad (5)$$

The activation function is transmitted through a link to other neurons via a feed-forward network design and then fed to a transfer function which could be linear or nonlinear to generate the output. The structure of the network is a function of the interaction between these neurons. Functions such as



**Figure 2.** A schematic framework of one-hidden-layer architecture.

the sigmoidal and radial (Gaussian) are used to build the neuron activity. A tangent-sigmoidal transfer function is used in the hidden layer and to avoid limiting the output to a small range, a linear transfer function is employed in the output layer. A Bayesian regularization (*trainbr*) back-propagation learning algorithm is utilized to speed up the convergence of the MLP model. Bayesian regularization is a robust iterative training algorithm that learns patterns based on input and output data; it essentially provides stronger and more efficient generalization ability by regularization. The ANN training process is stopped when a maximum number of 1000 epochs is reached. The error is measured based on the root-mean-square error (RMSE) between the predicted values and the output for all elements in the training and testing set. The networks are initialized to return neural network nets with weight and bias values updated according to the network initialization function.

**3.3.1. Pros of ANN.** The most vital advantages of ANN include:

- Its capability to model and map complex nonlinear systems (nonlinearity) by deriving a relationship between a set of input and output responses (input-output mapping).
- The ability to learn which allows the network to adapt to changes in the surrounding environment (adaptivity).
- Once ANNs are properly trained, damage identification is relatively fast and mathematical models do not need to be constructed.
- There are no limitations on the type of vibration parameters to be used as inputs for ANNs. The input and output parameters can be selected with much flexibility without increasing the complexity of network training.

**3.3.2. Cons of ANN.** A major demerit of ANNs is that the resulting weights and nets of the trained network are difficult to interpret (that is, an inability to obtain adequate solutions of complex problems with physical mathematical methods in contrast to RSM). Others include:

- It is difficult to find an appropriate network architecture.
- It usually suffers from the problems of underfitting (inaccurate approximation of the training data) resulting from too-small networks, and overfitting (inadequate generalization) due to too-large networks.
- It requires several networks of different architectures to be trained, and their performance compared on a separate set of test data to estimate their generalization properties.
- The training data of the database should be large enough to have a close relationship with the associated parameters (that is, sufficient training data for complex ANNs are necessary, requiring availability of a large database).

**3.3.3. Modifications to the basic ANN.** Extensive work has been done on the basic ANN to improve its generalization ability. The following are some contributions made to the ANN model to improve its generalization capability, accuracy of approximation, output variable handling, and training time.

- Single and ensemble nets: The capabilities of ANN to handle multidimensional outputs are well known. However, it is preferred to use individual ANN models to predict each output because it is not feasible to use ensemble (multiple) nets to predict all the output variables. Training single nets



enhances the prediction accuracy because only one output is predicted. This improves the handling capability of the optimizers when coupled with surrogates during the optimization process.

- The concept of the  $K$ -means clustering algorithm introduced enables one to effectively reduce the size of dataset required for ANN training and hence determine the minimum amount of datasets just sufficient for ANN training.
- The adoption of the *trainbr* training algorithm showed great effectiveness in the ANN results. Generalization enhancement by Bayesian regularization produces a network that performs well with the training data and exhibits smoother behavior when presented with new data.
- Increasing the number of hidden neurons and layers. Larger numbers of neurons in the hidden layer can give the network more flexibility because the network has more parameters it can optimize. However, the number of layers should be increased gradually because large hidden layers lead to undercharacterization of the network since the network must optimize more parameters than there are data vectors to constrain these parameters.
- Adoption of network architectures with tansig and purelin transfer functions yields better results for low RMSE values and high  $R^2$  values than other possible combinations tried.
- Preprocessing of the network inputs and targets improves the efficiency of the neural network training. It basically reduces the dimensions of the input vectors to increase network performance. This is essentially the normalization or scaling of inputs and targets so that they fall in the range  $[-1, 1]$ . However, it was seen that scaling had no great significant effect on the network performance for the proposed problem under consideration.

**3.3.4. Steps in designing ANN model.** The ANN training process is not an easy task and involves finding an appropriate ANN model for a given problem. Hence, the necessary requirements for a successful ANN development include: a sufficient database, careful selection of the parameters, network architecture (number of hidden layers, number of hidden and output nodes), transfer functions, and effective training and learning algorithms. These choices completely depend on the approximation function.

The process of building an ANN model can hence be outlined and summarized as follows:

- Selection of the number of input and output variables of the neural network and database creation (the dataset obtained from FEM analysis after a  $K$ -means clustering is divided into a training set and a testing set).
- Determination of network architecture by trial and error method (that is, number of hidden layers and number of hidden nodes — information processing occurs at many simple elements called neurons) required to generalize the design space.
- Selection of the training algorithm and functions and measure of its effectiveness based on RMSE performance.
- Invoking of a back-propagation algorithm to train multilayer feed-forward networks with differentiable transfer functions to perform function approximation. Training by back propagation is described as the process by which derivatives of network error with respect to network weights and biases are computed. This process is subdivided into:

- Feed forward of the input training data (input vector of set of design variables from the input layer is fed forward by propagation to the hidden layer and subsequently to the output layer) in a process termed the forward pass of the back propagation algorithm.
  - The computation and back propagation of the associated weight and errors. The predicted values of the output layer are compared with the target values. The error between the predicted and target values is calculated and propagated back toward the hidden layer in a process known as the backward pass of the back propagation algorithm. The error is used to update the weight matrices between the input-hidden layers and hidden-output layers.
- The mean square error of the network is computed by calculating the amount between the predicted and target values. Consequently, this error is minimized by a predetermined training algorithm using optimization algorithms based on a gradient-based back-propagation process which repeatedly changes the performance values depending on the network connection weights. The *trainbr* function is used for this purpose. This function trains the network by randomly initializing and updating the weights and bias values according to a Bayesian regularization algorithm.
  - Terminating the network training using 1000 maximum epochs to reduce the effect of random weights on training the network. This method stops the training when the maximum number of training cycles given is reached.
  - Computing the average performance of each of the network architecture with RMSE as the performance criterion and selecting the best network performance with the minimum RMSE.
  - The best model is applied to the test data to investigate the performance on an unseen set of data.

**3.3.5. ANN configuration to analyze the best network architectures.** Choosing the number of hidden layers and neurons in the hidden layer is also a demanding task and it is the principal limitation of ANN. Since the ANN configuration has a great influence on the predictive output, various arrangements have been considered. It is essential to designate a formula to describe the ANN configuration as  $(I-n_1-n_2-O)$ . For example, 2-20-1 means a one-hidden-layer ANN with two input parameters and one output parameter, with the hidden layers containing 20 elements (neurons); 3-10-10-2 denotes a three-input and two-output ANN, with ten neurons in two hidden layers. The MATLAB neural network toolbox has been used as the basis in which the networks can be configured in a very wide variety of architectures, and the training algorithms can be also chosen with ease.

Evaluation of the network performance is measured based on RMSE analysis and coefficient of determination ( $R^2$ ). The power of prediction can be quantified by the RMSE of the predicted output from test data. The smaller the RMSE of the test dataset is, the higher the predictive capability of the network.

The RMSE (that is, the root mean square of the differences between the actual and predicted values, which should be very close to zero) is expressed as

$$\text{RMSE} = \sqrt{\text{mean}((R_A - R_P)^2)}, \quad (6)$$

where  $R_A$  and  $R_P$  are the actual and predicted values from the network.

Similarly,  $R^2$  (that is, a measure of how well the variation in the output is explained by the targets; if

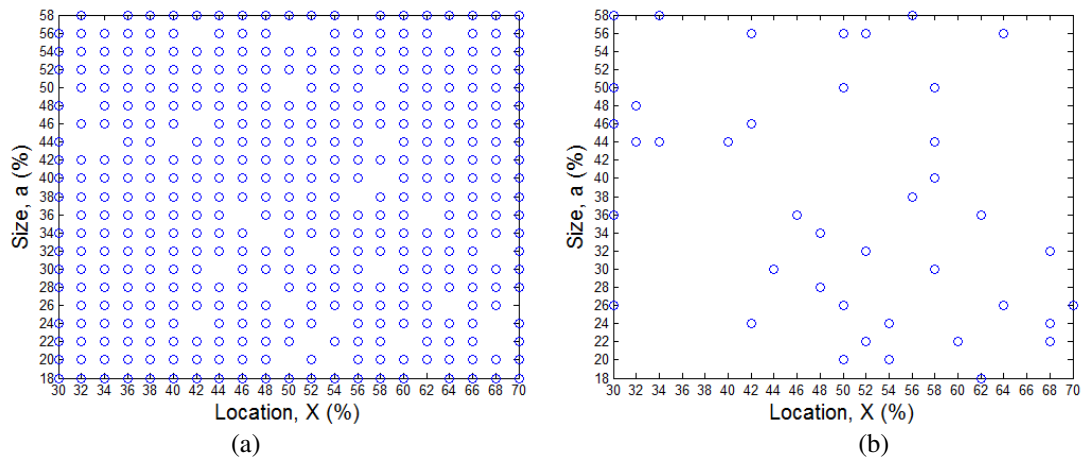
this number is equal to 1 there is a perfect correlation between the targets and outputs) is given by

$$R^2 = \frac{\sum(R_A - \bar{R}_A)(R_P - \bar{R}_P)}{\sqrt{\sum(R_A - \bar{R}_A)^2(\sum(R_P - \bar{R}_P))^2}}, \quad (7)$$

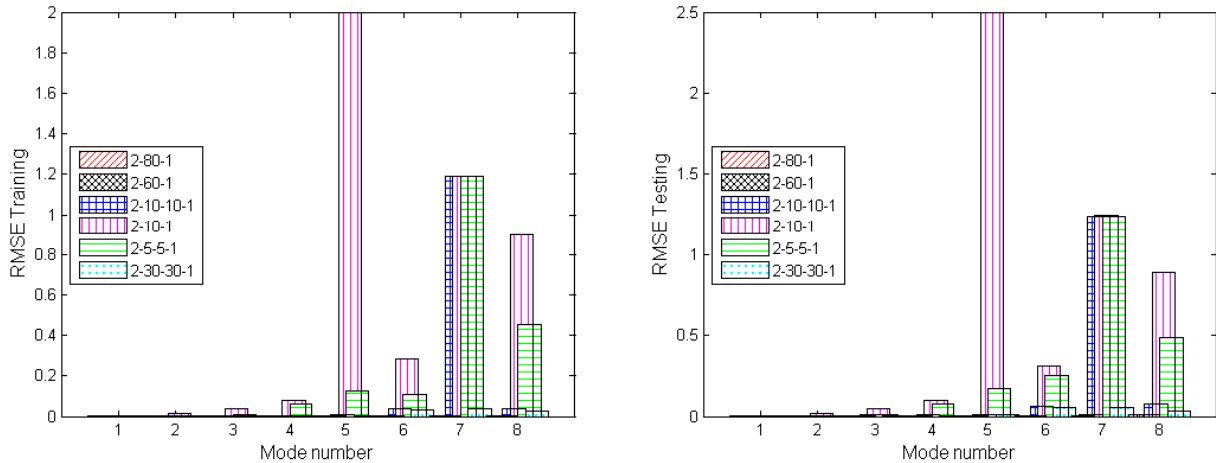
$\bar{R}_A$  and  $\bar{R}_P$  are defined as the means of the actual and predicted values from the network.

**3.3.6. Training performance of trial neural networks for the forward problem.** Several neural network architectures were tried since there is no laid down rule for choosing the optimal number of hidden layers and processing elements for each layer; it is basically problem dependent. However, incorporation of more processing elements enhances the generalization capability of the network for a larger number of training data points. The motivation was to determine the best network with the least RMSE and good coefficient of determination ( $R^2$ ) irrespective of the time. In other words, the length of computation time was not considered as a factor for choosing the best network. However, larger network architecture takes longer running time in all analysis, and smaller networks take a shorter time. It is shown that when enough training datasets have been used, the root mean square of the output error converges to zero. The neural network optimum network architecture in most cases consisted of one input layer, one or two hidden layers, and one output layer, with the circles representing processing elements or neurons in Figure 2. For the inverse problem, the eight inputs were the differences between the damaged beam and the undamaged beam for the first eight natural frequencies and the outputs were the interface, delamination size, and location, and vice versa for the forward problem. The hidden layers contained a maximum of 80 neurons.

For data generation, considering only midplane delaminations, 441 FE models equally spaced at gaps of 2% were run in a batch process for two hours assuming normalized delaminations located from 30% to 70% (30:2:70) of the total beam length and having normalized sizes ranging from 18% to 58% at gaps of 2% (18:2:58). The 441 generated datasets were randomized using  $K$ -means clustering. As shown in Figure 3(a), 90.7% (400) of the 441 simulations run in ANSYS were used for training a neural network



**Figure 3.** Design space for training and testing datasets. (a) 400 training datasets and (b) 41 testing datasets.



**Figure 4.** Comparison of performance of different network architectures on RMSE of training data (left) and testing data (right).

model and the remaining 9.3% (41) of the dataset were used for testing of the network as shown in Figure 3(b).

Then, consider the solution of a forward problem trained with ANN using different network architectures. Figure 4 demonstrates the trial performance of different network architectures with their training and testing RMSE values for the individual eight percentage frequency changes (individual nets) taken as output to the network with two inputs (location and size of delamination). It is shown that 2-80-1 is the best network with the least RMSE. It is evident that high network architectures give better results than low network architecture for both training and testing datasets.

**3.3.7. Training performance of trial neural networks for the inverse problem.** For the inverse problem where the eight natural frequencies are taken as input to the network and delamination location and size as network outputs, high network architectures are also found to give better results, in terms of low RMSE and high values of  $R^2$ , than small network architectures.

**3.4. Surrogate approach.** In the surrogate approach, ANN is used to build surrogate models for approximation of the computationally expensive FE models by solving the forward problem. This approach is solving the forward problem to approximate the output of the natural frequencies simulated using the FE models from the given corresponding delamination parameters. Later, these surrogates are used in the optimization loop instead of direct optimization via FE models as shown in Figure 1. The surrogate approach is extensively employed as an inexpensive approximation of the true function evaluations instead of the computationally expensive FEM simulations. Surrogates are especially interesting for expensive objective functions, since the necessary computational effort to build the surrogate is smaller than the effort of the objective function evaluation.

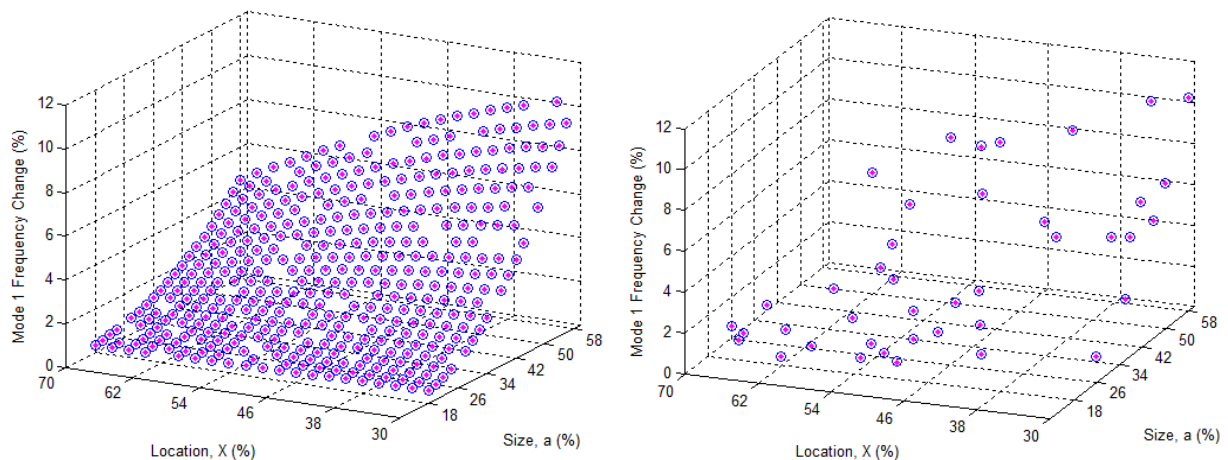
ANN, RBF, RSM, Kriging, and design and analysis of experiments are some examples of surrogate models [Sul et al. 2011]. The choice of surrogate models depends on the problem under consideration. In the current work ANN is to provide surrogate models to computationally expensive FE models and is found to be very effective. This justifies the need for surrogate models in reducing the optimization cycle

[Z, X, a]	Mode	FE	ANN	% error	[Z, X, a]	Mode	FE	ANN	% error
[1, 54, 24]	$dF_1$	0.8013	0.8026	0.1711	[1, 62, 18]	$dF_1$	0.4049	0.4044	0.1312
	$dF_2$	0.3182	0.3179	0.0831		$dF_2$	0.5266	0.5266	0.0075
	$dF_3$	10.6770	10.6769	0.0006		$dF_3$	2.6310	2.6310	0.0013
	$dF_4$	5.6542	5.6545	0.0050		$dF_4$	8.9072	8.9068	0.0039
	$dF_5$	14.5527	14.5530	0.0020		$dF_5$	2.1741	2.1740	0.0017
	$dF_6$	12.5625	12.5636	0.0083		$dF_6$	12.0654	12.0639	0.0126
	$dF_7$	13.9558	13.9579	0.0155		$dF_7$	9.4308	9.4367	0.0627
	$dF_8$	13.3759	13.3778	0.0134		$dF_8$	9.3618	9.3599	0.0199

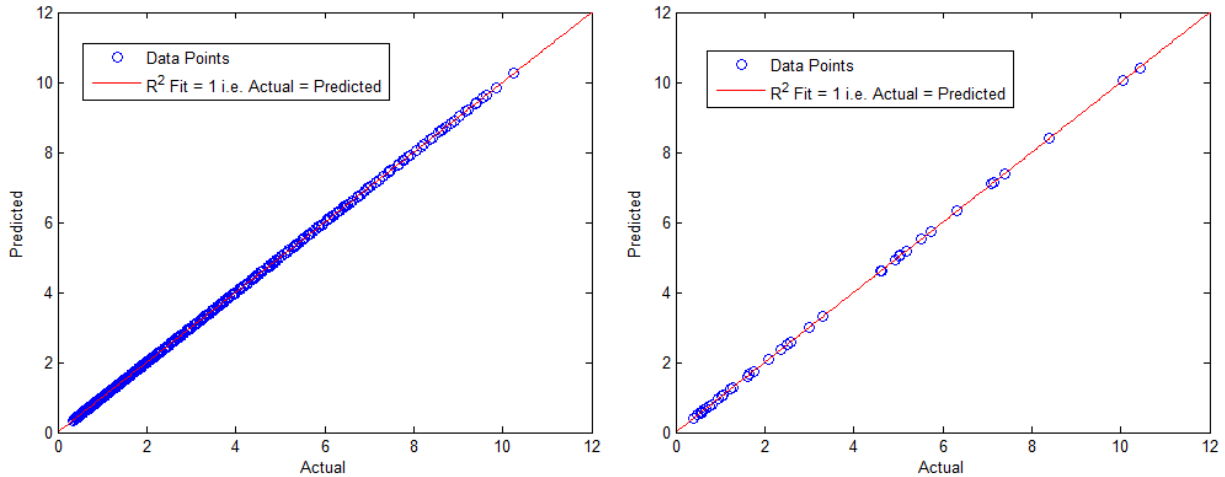
**Table 4.** Comparison of  $dF_i$  between FE simulations and ANN approximations.

time and exploring the complete design space with minimal computational cost. RSM was first tried, and failed to consistently provide a true approximation model due to high multimodality of the delamination detection problem in higher modes, say modes 4–8. Since ANN was found to perform very well for the solution of the problem under consideration, it was selected for the unique purpose of building surrogate models to expensive FE models [Ihesiulor et al. 2012b]. Other surrogate models were not tried.

**3.4.1. Validation of the surrogate model.** A validity check of our surrogate model is necessary to prevent misleading of the search by the optimizers due to poor approximations. To validate the performance of the trained ANN, a perfect match of true function evaluation by ANSYS and the approximated functions by ANN is shown in Table 4 for delamination signatures [1, 54, 24] and [1, 62, 18] for the first eight percentage changes in frequencies. The results are shown to have negligible error of not more than 0.17%. Also, perfect fits plot and regression ( $R^2$ ) plots for the 400 training and 41 test datasets between the simulated percentage change in frequencies and the predicted output from the neural network training for mode 1 are shown in Figures 5 and 6, respectively.



**Figure 5.** Simulated percentage change in natural frequencies: actual values (circles) versus predicted output (red stars) from the ANN model for mode-1 training (left) and testing (right) datasets.



**Figure 6.** Simulated percentage change in natural frequencies:  $R^2$  plot between actual and the predicted output from the ANN model for mode-1 training (left) and testing (right) datasets.

**3.5. Introduction to RSM.** Response surface methodology (RSM) can be defined as development of the mathematical and statistical techniques applied in the modeling and analysis of engineering problems in which the output of interest is governed by some input variables and the key objective is to optimize this output response [Montgomery 2005]. RSM is essentially a statistical method that employs quantitative data from appropriate simulations or experiments to determine and solve simultaneously multivariate equations.

RSM, also known as polynomial fitness function modeling, adopts regression curve fitting to obtain mathematical approximations of responses of a given system as functions of some input design variables [Todoroki 2001]. This method is widely employed as an inexpensive low-order approximation model instead of the more time-consuming but accurate calculations using FEM simulations. Response surfaces can easily be fitted to data by least-squares approach. The order of the polynomial is important; quadratic or cubic polynomials are mainly used, with quadratic polynomials best suited for continuous, unimodal problems. RSM models are widely used in polynomial approximation schemes due to their flexibility and ease of use. In the polynomial approximation method, the response surface model is a polynomial of  $n$ -th degree whose coefficients are determined from a linear system of equations. The linear system is solved using least-square minimization of the error between the predicted and actual values.

**3.5.1. Advantages of RSM.** In the current study, RSM is adopted as another inverse-problem solver because of the following benefits it enjoys over ANN:

- (1) It does not require too much computational effort and resources to generate its mathematical models (ease of calculations and use).
- (2) Its solution to the inverse problems can be approximately solved without the constraints of modeling.
- (3) Its approximation model can be easily validated through statistical means.

**3.5.2. Analysis of a first-order model response surface.** In RSM, the factors that are considered as most important are used to build a polynomial model in which the independent variable is response from experiments or numerical simulations. A first-order multiple regression model with  $N$  simulation runs carried out on  $k$  input variables ( $N > k$ ) and an output response  $R$  can be expressed as

$$R_i = C_0 + \sum_{j=1}^k C_j dF_{ij} + \epsilon_i \quad (i = 1, 2, \dots, N). \quad (8)$$

The response  $R_i$  is a function of the input variables  $dF_1, dF_2, \dots, dF_n$ , plus the error. The  $C_j$  are the regression coefficients, and  $dF_{ij}$  corresponds to the  $i$ -th sample and  $j$ -th independent variable.

Equation (8) can be expressed in matrix form as

$$R = dF * C + \epsilon, \quad (9)$$

$$\begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_N \end{bmatrix} = \begin{bmatrix} 1 & dF_{11} & dF_{12} & \cdots & dF_{1k} \\ 1 & dF_{21} & dF_{22} & \cdots & dF_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & dF_{N1} & dF_{N2} & \cdots & dF_{Nk} \end{bmatrix} \begin{bmatrix} C_0 \\ C_1 \\ \vdots \\ C_k \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix}.$$

$\underbrace{\hspace{10em}}_R \quad \underbrace{\hspace{10em}}_{dF} \quad \underbrace{\hspace{10em}}_C \quad \underbrace{\hspace{10em}}_\epsilon$

Invoking the least-square error method, the estimated coefficient  $\hat{C}$  of the coefficient vector ( $C$ ) can be given as

$$\hat{C} = (dF^T dF)^{-1} dF^T R. \quad (10)$$

**3.5.3. Analysis of a second-order model response surface.** The presence of high curvature in the response surface makes the first-order models unsuitable for complex problems. A second-order model becomes handy in approximating the true response surface. The second-order model accommodates all the terms in the first-order model, in addition to quadratic terms ( $C_{11} dF_{1i}^2$ ) and all cross-product terms ( $C_{13} dF_{1i} dF_{3j}$ ). The method of least squares can also be applied to estimate the coefficients in the model. The equation based on a second-order polynomial is given by

$$R_i = C_0 + \sum_{j=1}^k C_j dF_j + \sum_{j=1}^k C_{jj} dF_j^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^k C_{ij} dF_i dF_j + \epsilon. \quad (11)$$

Hence, to develop the relationship between the variations in the simulated natural frequencies and the corresponding size and location of delamination, response surfaces of fourth-degree polynomials were generally adequate in this study.

The RSM procedure described above is employed to fit a fourth-order polynomial equation using the simulation data where 400 FE model simulations are performed and 41 datasets used for testing. From the Minitab output, the fourth-order polynomial equation for predicting location and size is given below, where the output responses are the delamination location and size and the inputs are the first percentage changes in natural frequencies. The least-square error method is adopted to obtain the unknown coefficients of the polynomials. All the insignificant interaction terms are removed from the models using an ANOVA table.

**3.5.4. Validation and adequacy check of the developed models.** In order to ascertain the adequacy and goodness of the developed response surface approximation models, the coefficient of determination ( $R^2$ ) as already defined and the absolute average deviation (AAD) are adopted. The efficiency of the model in terms of its predictive power can be determined by both  $R^2$  and AAD, because  $R^2$  alone cannot be effectively used to measure the performance of the developed models. The  $R^2$  is basically a measure of how well the variation in the output is explained by the targets; if this number is equal to 1 there is a perfect correlation between targets and outputs. However, a large value of  $R^2$  does not necessarily imply that the regression model is a good one [Bas and Boyaci 2005]. Thus, it is possible for models that have large values of  $R^2$  to yield poor predictions of new observations or estimates of the mean response. Plotting actual results versus predicted results from the model gives a straight line passing the origin with an angle of  $45^\circ$ , but in practical cases the model fails to give accurate results to new data. This limitation of  $R^2$  is eliminated by using AAD analysis, which is a direct method for describing the deviations in the actual and predicted outputs by the models [Bas and Boyaci 2005]. Minitab was used to conduct all analysis.

The AAD is calculated by

$$\text{AAD} = \sum_{i=1}^N (|\text{abs}(R_A - R_P)/R_A|/N) * 100. \quad (12)$$

The expression for  $R^2$  is given in (7). Also, adjusted  $\bar{R}^2$  is defined as the improvement of  $R^2$  when the number of terms in a model is adjusted, which is and lower than the  $R^2$  value. It is given as

$$\bar{R}^2 = 1 - \frac{1 - R^2}{N - 1/N - k - 1} \quad (13)$$

where  $R_A$  and  $R_P$  are the actual and predicted responses, and  $N$  is the number of simulation runs.

Evaluation of the  $R^2$  and AAD values together was just adequate to check the accuracy of the developed models. The  $R^2$  must be close to 1 and the AAD between the predicted and actual output must be as small as possible tending towards 0. Acceptable values of  $R^2$  and AAD mean that the model equations define the true behavior of the system and they can be used for interpolation in the simulation design space.

From the equation shown in Table 5 for  $X$ , the  $R^2$  and adjusted  $R^2$  values for delamination location ( $X$ ) are calculated as 99.3% and 99.2%, respectively. The near-perfect prediction ( $R^2$  value) of the model is an indicator that the model generated has been perfected to fit the given data and thus is highly significant. The high  $R^2$  is a good indication of the predictive power of the developed model. Similarly, from the equation shown in Table 6 for  $a$ , the  $R^2$  and adjusted  $R^2$  values for the delamination size ( $a$ ) are deduced to be 100% and 100%, respectively. This shows that the developed model for delamination size has more predictive power in terms of accuracy of prediction results than the model for delamination location.

Also, from the equations shown in Tables 5 and 6, the calculated AAD for the delamination location and size are, respectively, obtained as 2.78% and 0.37%. This indicates that the fourth-order polynomial model for delamination size  $a$  is highly significant and adequate to represent the actual relationship between the response and the significant input variables, with very small AAD value (0.37%) and a satisfactory coefficient of determination ( $R^2 \approx 1$ ).



$$\begin{aligned}
&78.1 + 28.1 dF_1 - 29.6 dF_2 + 11.1 dF_3 + 16.5 dF_4 - 7.65 dF_5 - 1.74 dF_6 + 7.90 dF_7 + 6.22 dF_8 \\
&+ 6.62 dF_1 dF_2 + 12.2 dF_1 dF_3 - 0.647 dF_1 dF_4 - 0.289 dF_1 dF_5 + 1.79 dF_1 dF_6 - 1.70 dF_1 dF_7 \\
&+ 1.05 dF_1 dF_8 - 1.90 dF_2 dF_3 + 0.642 dF_2 dF_4 + 0.279 dF_2 dF_5 + 0.253 dF_2 dF_6 + 0.591 dF_2 dF_7 \\
&- 0.0537 dF_2 dF_8 - 0.0150 dF_3 dF_5 + 0.546 dF_3 dF_6 + 0.284 dF_3 dF_7 - 0.170 dF_3 dF_8 \\
&- 0.234 dF_4 dF_5 - 0.0908 dF_4 dF_6 - 0.150 dF_4 dF_7 - 0.0582 dF_4 dF_8 - 50.5 dF_1^2 \\
&- 0.015 dF_2^2 - 1.21 dF_3^2 + 0.666 dF_4^2 + 0.314 dF_5^2 + 0.274 dF_6^2 - 0.015 dF_7^2 \\
&- 0.087 dF_8^2 + 2.14 dF_1^3 + 0.0324 dF_2^3 + 0.0286 dF_3^3 - 0.00935 dF_4^3 - 0.0156 dF_5^3 \\
&- 0.0186 dF_6^3 + 0.00317 dF_7^3 + 0.0029 dF_8^3 - 0.0570 dF_1^4 - 0.00114 dF_2^4 - 0.000677 dF_3^4 \\
&+ 0.000030 dF_4^4 + 0.000114 dF_5^4 + 0.000264 dF_6^4 - 0.000075 dF_7^4 - 0.000055 dF_8^4
\end{aligned}$$

**Table 5.** Fourth-order polynomial equation for generated delamination location ( $X$ ) in terms of input percentage change in natural frequencies.

$$\begin{aligned}
&13.8 - 6.16 dF_1 + 1.23 dF_2 + 0.667 dF_3 + 0.214 dF_4 + 0.0664 dF_5 - 0.0667 dF_6 \\
&- 0.276 dF_7 + 0.541 dF_8 + 0.247 dF_1 dF_2 - 0.0635 dF_1 dF_3 + 0.0177 dF_1 dF_4 \\
&+ 0.0152 dF_1 dF_5 + 0.0530 dF_1 dF_6 - 0.0268 dF_1 dF_7 - 0.168 dF_1 dF_8 \\
&+ 0.00728 dF_2 dF_3 + 0.0328 dF_2 dF_4 - 0.0118 dF_2 dF_5 - 0.00396 dF_2 dF_6 \\
&+ 0.00520 dF_2 dF_7 + 0.0443 dF_2 dF_8 + 0.00523 dF_3 dF_5 - 0.00468 dF_3 dF_6 \\
&+ 0.00086 dF_3 dF_7 + 0.0228 dF_3 dF_8 - 0.00051 dF_4 dF_5 - 0.00669 dF_4 dF_6 \\
&+ 0.00139 dF_4 dF_7 - 0.00449 dF_4 dF_8 + 0.733 dF_1^2 - 0.213 dF_2^2 - 0.00885 dF_3^2 \\
&+ 0.0203 dF_4^2 + 0.0339 dF_5^2 + 0.0289 dF_6^2 + 0.0547 dF_7^2 - 0.0752 dF_8^2 - 0.0746 dF_1^3 \\
&+ 0.00782 dF_2^3 + 0.000083 dF_3^3 - 0.00118 dF_4^3 - 0.00184 dF_5^3 - 0.00128 dF_6^3 \\
&- 0.00289 dF_7^3 + 0.00389 dF_8^3 + 0.00197 dF_1^4 - 0.000149 dF_2^4 + 0.000020 dF_3^4 \\
&+ 0.000020 dF_4^4 + 0.000034 dF_5^4 + 0.000021 dF_6^4 + 0.000047 dF_7^4 - 0.000061 dF_8^4
\end{aligned}$$

**Table 6.** Fourth-order polynomial equation generated for delamination size ( $a$ ) in terms of input percentage change in natural frequencies.

For the response surface model for delamination location  $X$ , it was found that the low AAD was obtained due to high nonlinear curvature in the delamination location with respect to the percentage change in frequencies.

#### 4. Comparison of delamination prediction efficiencies of different algorithms

The approach or algorithm one uses will have a great effect on the accuracy of results. In solving the problem of delamination detection in composite laminates, different algorithms have been employed as seen in the previous sections and in [Ihesiulor et al. 2012a]. Hence, it is worthwhile to do a comparative analysis of these algorithms to ascertain the most effective and efficient in terms of accuracy of prediction. Efficiency of algorithms with respect to the minimization optimization algorithms can be measured as the minimum time needed to lower the error below a certain specified value associated with the value of the objective function after a given number of runs. The efficiency of the ANN and RSM algorithms can be defined as the measure of performance in terms of RMSE and AAD, respectively. Accordingly,

Algorithm	Name
1	Artificial neural network (ANN)
2	Response surface method (RSM)
3	GBLS with surrogates (GBLS <sub>WoS</sub> )
4	GBLS without surrogates (GBLS <sub>WS</sub> )
5	NSGA-II with surrogates (NSGA-II <sub>WS</sub> )
6	NSGA-II without surrogates (NSGA-II <sub>WoS</sub> )

**Table 7.** List of proposed algorithms.

one is only interested in the algorithm that gives the least prediction error regardless of the time, but the evaluation time should not be too large before expected results are evaluated for effective use in online SHM. With these factors under consideration, a comparative analysis is made in this section between four key algorithms (ANN, RSM, GBLS, and NSGA-II) under six different approaches (ANN, RSM, GBLS<sub>WS</sub>, NSGA-II<sub>WS</sub>, GBLS<sub>WoS</sub>, and NSGA-II<sub>WoS</sub>) as shown in Table 7, where GBLS<sub>WS</sub> and GBLS<sub>WoS</sub> are GBLS methods with and without surrogates, respectively. Similarly, NSGA-II<sub>WS</sub> and NSGA-II<sub>WoS</sub> are NSGA-II methods with and without surrogates, respectively.

So far we have shown that most algorithms are fast and contains several capabilities, such as:

- Ease of use and implementation and allowing a user to specify input data for which the delamination signature is to be ascertained easily (RSM).
- Generation of random results up to a significant number of runs and selection of the best result in terms of the minimum objective function value (NSGA-II).

For this study, in terms of dataset generation, the dataset scenario described earlier containing 400 training datasets and 41 test datasets was considered for this comparative analysis and two hours was required to generate the database. The training data used for this comparison was simulation data benchmarked at 400 datasets for all algorithms and ten damage cases, as shown in Table 8, out of the 41 reserved tests were selected for the performance comparative study. This comparison was based on the two-variable problem, that is, predicting delamination location and size at midplanes.

**4.1. Algorithm 1: ANN.** The ANN model used in the benchmark tests had 8 input nodes, 80 hidden nodes, and 2 output nodes (that is, a network architecture of 8-80-2). The training data contained 400 input and output vector pairs. The training time for the ANN algorithm was measured to be about 133 s. Table 9 gives the prediction results of the selected ten cases using ANN. It is seen that ANN gives for

S/N	Delamination signature $[X, a]$	S/N	Delamination signature $[X, a]$
1	[50, 26]	6	[58, 40]
2	[68, 24]	7	[48, 34]
3	[50, 20]	8	[58, 30]
4	[30, 36]	9	[48, 28]
5	[34, 44]	10	[62, 36]

**Table 8.** Selected test cases to ascertain the method with best performance.

S/N	Actual $[X, a]$	Predicted $[X, a]$	% ( $[(E - X), (E - a)]$ )
1	[50, 26]	[49.9946, 26.0000]	[0.0109, 0.0001]
2	[68, 24]	[68.0095, 23.9990]	[0.0139, 0.0042]
3	[50, 20]	[50.0068, 19.9999]	[0.0136, 0.0003]
4	[30, 36]	[30.0009, 36.0000]	[0.0031, 0.0001]
5	[34, 44]	[33.9994, 44.0000]	[0.0019, 0.0001]
6	[58, 40]	[57.9997, 40.0004]	[0.0006, 0.0011]
7	[48, 34]	[48.0056, 34.0001]	[0.0116, 0.0002]
8	[58, 30]	[57.9919, 30.0003]	[0.0139, 0.0011]
9	[48, 28]	[47.9980, 27.9999]	[0.0043, 0.0004]
10	[62, 36]	[61.9995, 36.0005]	[0.0008, 0.0013]
Total			[0.0745, 0.0088]

**Table 9.** Percentage errors of ten test cases using ANN modeling.

the ten cases total percentage errors of 0.07% and 0.009% in location and size, respectively. This shows that ANN is very unique in performance, giving a maximum error of 0.014% in all the cases under consideration.

**4.2. Algorithm 2: RSM.** Invoking the RSM regression models developed in the equations shown in Tables 5 and 6 using 400 datasets for delamination location and size, respectively, the prediction results of the ten selected cases are shown in Table 10. Results show that RSM gives adequate approximations as an inverse tool using the variations in natural frequencies for delamination detection. It is seen that RSM can be used to successfully predict delamination location and size. The test data helps to establish that a model that closely matches the actual values is developed and it is still a mathematical fit over the training data. It is seen that the results in predicting delamination size are more accurate than for delamination location because the percentage change in frequency increases monotonically with increase

S/N	Actual $[X, a]$	Predicted $[X, a]$	% ( $[(E - X), (E - a)]$ )
1	[50, 26]	[50.0724, 25.9779]	[0.1448, 0.0850]
2	[68, 24]	[69.0697, 23.9806]	[1.5731, 0.0808]
3	[50, 20]	[50.4786, 19.9550]	[0.9572, 0.2250]
4	[30, 36]	[29.9083, 36.1363]	[0.3057, 0.3786]
5	[34, 44]	[35.1090, 44.2064]	[3.2618, 0.4691]
6	[58, 40]	[59.3110, 40.0181]	[2.2603, 0.0452]
7	[48, 34]	[48.1630, 34.0942]	[0.3396, 0.2771]
8	[58, 30]	[58.4583, 29.9824]	[0.7902, 0.0587]
9	[48, 28]	[47.9457, 27.9744]	[0.1131, 0.0914]
10	[62, 36]	[62.8340, 35.9788]	[1.3452, 0.0589]
Total			[11.0909, 1.7698]

**Table 10.** Percentage errors of ten test cases using RSM modeling.

S/N	Actual $[X, a]$	Minimum <sub>Obj</sub>	Predicted $[X, a]$	% $([(E - X), (E - a)])$
1	[50, 26]	0.09550	[49.6446, 25.4675]	[0.7108, 2.0482]
2	[68, 24]	0.00553	[68.0706, 23.8523]	[0.1039, 0.6156]
3	[50, 20]	0.00852	[50.0634, 19.4827]	[0.1268, 2.5863]
4	[30, 36]	1.56888	[31.3309, 36.1582]	[4.4363, 0.4396]
5	[34, 44]	0.69202	[34.6365, 43.4115]	[1.8721, 1.3376]
6	[58, 40]	0.73237	[57.8042, 38.9887]	[0.3376, 2.5282]
7	[48, 34]	0.00528	[48.0291, 33.2507]	[0.0606, 2.2037]
8	[58, 30]	2.82593	[42.2571, 30.0895]	[27.1428, 0.2983]
9	[48, 28]	0.00543	[48.0293, 27.3201]	[0.0610, 2.4284]
10	[62, 36]	0.00481	[62.1006, 35.6956]	[0.1622, 0.8456]
Total				[35.0140, 15.3315]

**Table 11.** Percentage errors of ten test cases using NSGA-II<sub>woS</sub>.

in delamination size whereas for the delamination location versus percentage frequency change, the frequency changes increase with a very sharp curvature resulting in a high nonlinear complex function. Hence, it can be deduced that the total sums of errors for the ten test cases are given to be 11% and 2% for delamination location and size, respectively. However the maximum error for individual cases is shown to be 3% and 0.5% in delamination location and size, respectively. These results suggest that the performance of the RSM algorithm results is quite satisfactory. This approach also requires less time for computation, less than 5 s. The only significant time involved in this approach is the time required to create the 400 data points used for the surface fitting.

**4.3. Algorithm 3: NSGA-II<sub>woS</sub>.** The NSGA-II<sub>woS</sub> approach is basically the minimization of the objective function directly from the FE models. This approach is very time-consuming, or in other words, computationally expensive. To be consistent with the amount of datasets used for database creation only 400 function evaluations of the FE models are allowed for about two hours. Because of this significant amount of time required to produce 400 function evaluations, only one test run was allowed. Over this limit of function evaluations, Table 11 shows that prediction results are highly unsatisfactory, with the total sum of errors for the ten test cases given as 35% and 15% in location and size, respectively. The individual maximum errors in location and size predictions are seen to be 27% and 2.6%, respectively.

**4.4. Algorithm 4: NSGA-II<sub>WS</sub>.** The NSGA-II<sub>WS</sub> approach is essentially the minimization of the objective function via the surrogate models instead of the FE models. This approach is very time saving and increases optimization performance and results. At 400 function evaluations, Table 12 shows the prediction results of the best and mean predicted values over ten runs for each test case. The results are highly satisfactory with the maximum total sum of errors for the ten test cases given as 1% and 2% in location and size, respectively. The individual maximum errors in location and size predictions are 0.9% and 1%, respectively. The mean and best predicted results over the ten runs are also tabulated. It takes 107 s by the surrogate model to get to the lowest minimum objective function value at  $1.95 \times 10^{-8}$ . The standard deviation (std) over the ten runs for each case is found to be reasonable enough, confirming the correlation of results. The total time taken for this approach is calculated as the time taken for building

S/N	Actual $[X, a]$	Best $[X, a]$	Mean $[X, a]$
1	[50, 26]	[49.9291, 25.8786]	[49.6955, 25.5635]
2	[68, 24]	[68.0011, 24.0027]	[68.0426, 24.0560]
3	[50, 20]	[50.1820, 20.0756]	[49.9154, 19.8753]
4	[30, 36]	[30.0322, 36.0383]	[30.2427, 36.0153]
5	[34, 44]	[33.9957, 44.0029]	[34.0788, 43.8958]
6	[58, 40]	[57.9630, 40.0507]	[58.0399, 40.0705]
7	[48, 34]	[47.8901, 34.0011]	[48.1582, 34.1024]
8	[58, 30]	[58.0443, 29.8654]	[57.7149, 29.9079]
9	[48, 28]	[47.9878, 28.0775]	[47.7617, 27.6020]
10	[62, 36]	[61.9884, 36.0114]	[62.0230, 36.0974]
S/N	Best $\min_{\text{Obj}S}$	Best % $([(E-X), (E-a)])$	Std.
1	$1.03 \times 10^{-4}$	[0.1419, 0.4668]	[0.5757, 0.6222]
2	$1.36 \times 10^{-6}$	[0.0016, 0.0111]	[0.2205, 0.1768]
3	$4.41 \times 10^{-4}$	[0.3640, 0.3781]	[0.3117, 0.2228]
4	$3.10 \times 10^{-6}$	[0.1073, 0.1065]	[0.2449, 0.2519]
5	$1.95 \times 10^{-8}$	[0.0126, 0.0066]	[0.4450, 0.3539]
6	$4.46 \times 10^{-6}$	[0.0637, 0.1267]	[0.3116, 0.4229]
7	$7.75 \times 10^{-5}$	[0.2290, 0.0032]	[0.5521, 0.4906]
8	$6.91 \times 10^{-5}$	[0.0764, 0.4485]	[0.8881, 0.3545]
9	$4.85 \times 10^{-5}$	[0.0255, 0.2767]	[0.6261, 0.9934]
10	$6.10 \times 10^{-7}$	[0.0187, 0.0318]	[0.1003, 0.2432]
Total		[1.0406, 1.8559]	

**Table 12.** Percentage errors of ten test cases using NSGA-II<sub>WS</sub> via surrogate models over ten runs.

the database for the surrogate models and the optimization time using surrogates, or

$$\begin{aligned} \text{Total time} &= \text{time to build the database for surrogate modeling} + \text{average optimization time for 10 runs} \\ &= 7135 + 107 = 7242 \text{ s.} \end{aligned}$$

**4.5. Algorithm 5: GBL<sub>S</sub><sub>WoS</sub>.** The GBL<sub>S</sub><sub>WoS</sub> approach is basically the minimization of the objective function directly from the FE models. This approach is very time-consuming, requiring about three hours to execute an average of 603 function calls over 100 different start points. The function calls cannot be limited to the 400 function evaluations used in the other algorithms and it takes a bit more function calls to obtain the result shown in Table 13. Prediction results are unsatisfactory with the maximum total sum of errors for the ten test cases given as 29% in both location and size. The individual maximum errors in location and size predictions are 7% and 6%, respectively. This approach requires no database creation and takes an average total time of 10016 s.

**4.6. Algorithm 6: GBL<sub>S</sub><sub>WS</sub>.** The GBL<sub>S</sub><sub>WS</sub> approach is essentially the minimization of the objective function via the surrogate models instead of the FE models. This approach is very time saving and

S/N	Actual $[X, a]$	Predicted $[X, a]$	Minimum <sub>Obj</sub>	% $([(E - X), (E - a)])$
1	[50, 26]	[50.7110, 27.0946]	2.7400	[1.4220, 4.2100]
2	[68, 24]	[67.8292, 24.8455]	2.9079	[0.2512, 3.5229]
3	[50, 20]	[52.4490, 18.7351]	3.0489	[4.8980, 6.3245]
4	[30, 36]	[30.7860, 36.3687]	1.0421	[2.6200, 1.0242]
5	[34, 44]	[34.7297, 45.7929]	4.1068	[2.1462, 4.0748]
6	[58, 40]	[54.8648, 39.4625]	3.9815	[5.4055, 1.3438]
7	[48, 34]	[51.1505, 34.1076]	1.5220	[6.5635, 0.3165]
8	[58, 30]	[59.9530, 31.3064]	4.4239	[3.3672, 4.3547]
9	[48, 28]	[48.2997, 29.0019]	2.4912	[0.6244, 3.5782]
10	[62, 36]	[61.0460, 35.9119]	1.8632	[1.5387, 0.2447]
Total				[28.8367, 28.9942]

**Table 13.** Percentage errors of ten test cases using  $GBLS_{WoS}$  over 100 start points.

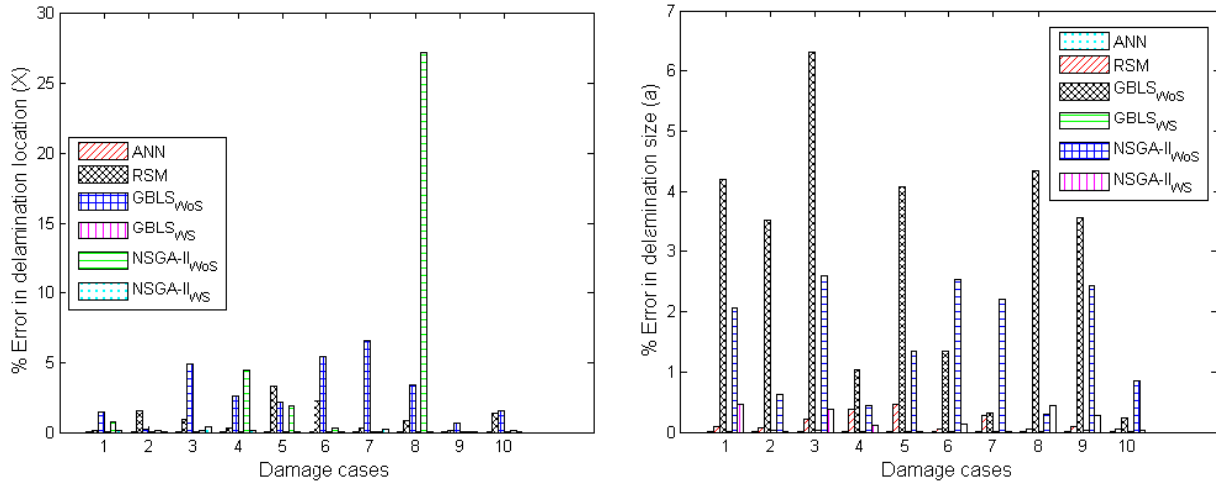
increases optimization performance and results. At an average 524 function calls over ten different start points for the ten test cases, the prediction results using this method are shown in Table 14. The results are also highly satisfactory, with the total sums of errors for the ten test cases being 0.04% and 0.12% in location and size, respectively. The individual maximum errors in location and size predictions are 0.06% and 0.04%, respectively. The excellent results of  $GBLS_{WS}$  can be attributed to a very low objective function, with the lowest at  $3.63 \times 10^{-9}$ . The total time taken for this approach is calculated as the time taken for building the database for the surrogate models and the optimization time using surrogates:

$$\begin{aligned} \text{Total time} &= \text{time to build database for surrogate modeling} + \text{average optimization time for 10 runs} \\ &= 7135 + 124 = 7259 \text{ s.} \end{aligned}$$

**4.7. Summary of comparative results.** Tables 15 and 16 give a comprehensive comparative analysis in terms of completion time and least minimum objective function value and prediction errors, respectively,

S/N	Actual $[X, a]$	Predicted $[X, a]$	# of calls	Minimum <sub>ObjS</sub>	% $([(E - X), (E - a)])$
1	[50, 26]	[49.9983, 26.0013]	657	$4.40 \times 10^{-7}$	[0.0034, 0.0050]
2	[68, 24]	[68.0023, 24.0063]	477	$1.26 \times 10^{-6}$	[0.0034, 0.0262]
3	[50, 20]	[50.0011, 19.9926]	292	$1.08 \times 10^{-5}$	[0.0022, 0.0370]
4	[30, 36]	[30.0017, 36.0010]	530	$4.45 \times 10^{-9}$	[0.0057, 0.0028]
5	[34, 44]	[33.9993, 44.0010]	545	$3.63 \times 10^{-9}$	[0.0021, 0.0023]
6	[58, 40]	[57.9984, 40.0036]	635	$5.50 \times 10^{-8}$	[0.0028, 0.0090]
7	[48, 34]	[48.0014, 34.0022]	661	$5.22 \times 10^{-8}$	[0.0029, 0.0065]
8	[58, 30]	[58.0036, 29.9994]	552	$7.76 \times 10^{-8}$	[0.0062, 0.0020]
9	[48, 28]	[48.0024, 28.0058]	670	$2.83 \times 10^{-7}$	[0.0050, 0.0207]
10	[62, 36]	[61.9991, 36.0038]	523	$4.48 \times 10^{-8}$	[0.0015, 0.0106]
Total					[0.0350, 0.1220]

**Table 14.** Percentage errors of ten test cases using  $GBLS_{WS}$  over ten start points.



**Figure 7.** Performance comparison of different methods based on error of prediction in delamination location  $X$  (left) and size  $a$  (right).

for the different algorithms ANN, RSM, NSGA-II<sub>WoS</sub>, NSGA-II<sub>WS</sub>, GBLS<sub>WoS</sub>, and GBLS<sub>WS</sub>. The prediction percentage errors in delamination location and size for the ten test cases are plotted for better comparison in Figure 7. From Tables 15 and 16 and Figure 7, the following deductions can be made.

- Giving a general conclusion on the algorithm with the best performance in terms of negligible prediction error irrespective of the time, we see from Table 16 that ANN outperforms all other methods, with a maximum prediction error of 0.012% in predicting delamination location and size. This is followed by GBLS<sub>WS</sub>, with a maximum prediction error of 0.06% in terms of delamination location and size, and then NSGA<sub>WS</sub>, with a maximum prediction error of 0.5%. Fourth in the performance order is the highly flexible RSM, with maximum prediction errors of 3.3% and 0.4% in predicting delamination location and size, respectively. The optimization methods without surrogates (GBLS<sub>WoS</sub> and NSGA-II<sub>WoS</sub>) both perform badly, with high amounts of error, which justifies the approach and objective of this study since optimization without surrogates is not only computationally demanding but yields poor results in terms of prediction accuracy because of the huge amount of computational effort required to explore the design space.

S/N	Algorithm	ACT = DT + RT (s)	Minimum objfun
1	ANN	7135 + 133 = 7268	N/A
2	RSM	7135 + 5 = 7140	N/A
3	GBLS <sub>WoS</sub>	10016	4.42
4	GBLS <sub>WS</sub>	7135 + 124 = 7259	$3.63 \times 10^{-9}$
5	NSGA-II <sub>WoS</sub>	7135	2.82
6	NSGA-II <sub>WS</sub>	7135 + 107 = 7242	$1.95 \times 10^{-8}$

**Table 15.** Average completion time (ACT) and lowest minimum objective function values (minimum objfun) for different algorithms, where DT is the time to create database and RT is the run time for the algorithm.

S/N	ANN	RSM	GBLS <sub>WoS</sub>
1	[0.0109, 0.0001]	[0.1448, 0.0850]	[1.4220, 4.2100]
2	[0.0139, 0.0042]	[1.5731, 0.0808]	[0.2512, 3.5229]
3	[0.0136, 0.0003]	[0.9572, 0.2250]	[4.8980, 6.3245]
4	[0.0031, 0.0001]	[0.3057, 0.3786]	[2.6200, 1.0242]
5	[0.0019, 0.0001]	[3.2618, 0.4691]	[2.1462, 4.0748]
6	[0.0006, 0.0011]	[2.2603, 0.0452]	[5.4055, 1.3438]
7	[0.0116, 0.0002]	[0.3396, 0.2771]	[6.5635, 0.3165]
8	[0.0139, 0.0011]	[0.7902, 0.0587]	[3.3672, 4.3547]
9	[0.0043, 0.0004]	[0.1131, 0.0914]	[0.6244, 3.5782]
10	[0.0008, 0.0013]	[1.3452, 0.0589]	[1.5387, 0.2447]
Total	[0.0745, 0.0088]	[11.0909, 1.7698]	[28.8367, 28.9942]
S/N	GBLS <sub>WS</sub>	NSGA-II <sub>WoS</sub>	NSGA-II <sub>WS</sub>
1	[0.0034, 0.0050]	[0.7108, 2.0482]	[0.1419, 0.4668]
2	[0.0034, 0.0262]	[0.1039, 0.6156]	[0.0016, 0.0111]
3	[0.0022, 0.0370]	[0.1268, 2.5863]	[0.3640, 0.3781]
4	[0.0057, 0.0028]	[4.4363, 0.4396]	[0.1073, 0.1065]
5	[0.0021, 0.0023]	[1.8721, 1.3376]	[0.0126, 0.0066]
6	[0.0028, 0.0090]	[0.3376, 2.5282]	[0.0637, 0.1267]
7	[0.0029, 0.0065]	[0.0606, 2.2037]	[0.2290, 0.0032]
8	[0.0062, 0.0020]	[27.1428, 0.2983]	[0.0764, 0.4485]
9	[0.0050, 0.0207]	[0.0610, 2.4284]	[0.0255, 0.2767]
10	[0.0015, 0.0106]	[0.1622, 0.8456]	[0.0187, 0.0318]
Total	[0.0350, 0.1220]	[35.0140, 15.3315]	[1.0406, 1.8559]

**Table 16.** Summary of comparative prediction percentage error results — that is, pairs  $[(E - X), (E - a)]$  — for different proposed algorithms.

- ANN and GBLS<sub>WS</sub> give the best delamination predictions. Results show not more than 0.06% error using both methods for predicting both delamination location and size at a known interface. However, ANN results are achieved by solving only the inverse problem whereas the optimization method requires the solution of the forward problems.
- When ANN and RSM are compared, it is evident that ANN's results beat those of RSM in terms of prediction results; however, the ease and flexibility of the RSM methods when compared to ANN and even other methods offer vital advantages. This is basically because, unlike other methods, RSM gives mathematical expressions (models) that can be used to predict delamination location and size of a modeled structure at any given time.
- GBLS<sub>WS</sub>, in comparison with NSGA-II<sub>WS</sub>, offers better results in terms of accuracy, as is evident in the lowest minimum objective function values for the former and latter,  $3.63 \times 10^{-9}$  and  $1.95 \times 10^{-8}$ , respectively, as shown in Table 15. However, subsequent comparisons between the two approaches when the number of variables considered increases from two to five, as in a more complicated



composite plate, revealed that NSGA-II<sub>WS</sub> performs consistently better than GBLS<sub>WS</sub> due to its ability to accommodate both discrete and continuous variables, which are limited in other methods.

- When the minimum objective function values are compared in Table 15 for the optimization techniques, it is shown that GBLS<sub>WS</sub> gives the lowest minimum value of objective function, followed by NSGA-II<sub>WS</sub>. The objective functions for optimization without surrogates are very high leading to their inefficiency in delamination predictions.
- The number of iterations or function calls for GBLS cannot be controlled, unlike NSGA-II, and hence it yields better results even with function evaluations of less than 300 calls. This justifies the need to use surrogate models.
- It can also be deduced from Table 15 that it takes about the same completion time for all the algorithms to successfully predict delaminations.

## 5. Conclusions

This paper presents potential candidate methods for the detection of delamination damage in composite laminates using variations of natural frequencies before and after damage. Key algorithms (artificial neural networks (ANN), response surface methodology (RSM), and optimization techniques) for real-world implementation of structural health monitoring (SHM) systems for delamination detection in composite structures are presented. The structural health monitoring systems proposed are incorporated with vibration-based damage detection techniques, which has been regarded as an efficient way to assess delamination damage in a structure and foresee probable costly failures. This research should help to identify starting points for damage detection via vibration-based monitoring and also guide practitioners in the field in choosing and implementing the most effective available damage-identification algorithms. The conclusions of our findings can be summarized as thus:

- Different inverse algorithms for delamination damage detection for SHM have been successfully developed and tested with numerical results. The notable excellent delamination prediction results obtained reiterate the robustness and accuracy of the algorithms as well as the approach. The results also underline the advantages of using *K*-means clustering for the choice of database while integrating surrogates in the optimization loop.
- An interesting finding in this research is that related to the superiority of the incorporation of surrogate models in the optimization loop. Surrogates with adequate approximation accuracy can be used to replace computationally expensive analysis. Management of such surrogates, that is, training regimes, selection of training data, and validation schemes, play an important role in any surrogate-assisted optimization exercise. Hence, surrogates enhance the optimization search performance by exploring the entire design space in a relatively short time.
- The use of optimization techniques directly in the optimization loop for delamination detection in composite beams and plates requires evaluations of large numbers of candidate solutions, thereby making the evaluations computationally expensive. The research reported in this thesis is focused on improving the efficiency of delamination detection results by allowing minimum computational cost. The approach adopted was to use surrogate models in lieu of expensive simulations to evaluate the

natural frequencies of the laminates. Further, an optimization strategy has been developed by integration with surrogate models. The optimization methodology was successfully applied to delamination prediction with the objective of improving structural integrity and minimizing computational costs.

- ANN is a superior and more accurate modeling technique as compared with RSM, as it represents nonlinearities in a much better way. However, a major disadvantage of ANN is that the resulting weights of the trained network are difficult to interpret, unlike RSM, which provides physical mathematical models that are easy to compute and interpret. Another problem peculiar to ANN, in contrast to RSM, is the difficulty of finding an appropriate network architecture. The RSM algorithm is straightforward in that it gives a physical mathematical expression for a delamination prediction and its model can easily be validated by statistical means.
- Both RSM and ANN require a large number of numerical experiments for obtaining a trained ANN or regressed response surfaces for predicting the delamination location and size from the measured data. By means of  $K$ -means clustering, the total number of numerical experiments required to build an ANN and RSM model is reduced to a substantial size. This represents a significant computational cost reduction associated with the developed approach. It can be seen from the previous work by the authors that with small training datasets, ANN and RSM fail to give accurate prediction results. But when solving a forward problem by creating surrogates with small datasets (say of size 40) and doing surrogate-assisted optimization with gradient-based local search (GBLS) and a nondominated sorting genetic algorithm (NSGA-II), the prediction results are very satisfactory.
- When one is interested in solving only the two-variable problem (delamination prediction at a known interface), the RSM and ANN algorithms deliver a good job. However, for three variables (delamination prediction at an unknown interface), surrogate-assisted optimization using GBLS and NSGA-II is very efficient. Finally, for the five-variable problem, as in composite plates, NSGA-II is preferred.

## References

- [Addin et al. 2006] A. O. Addin, S. M. Sapuan, E. Mahdi, and M. Osman, "Review on prediction and detection of failures in laminated composite materials using neural networks", *Polym. Polym. Compos.* **14**:4 (2006), 433–441.
- [Ayala and dos Santos Coelho 2008] H. V. H. Ayala and L. dos Santos Coelho, "A multiobjective genetic algorithm applied to multivariable control optimization", pp. 736–745 in *19th COBEM: International Congress of Mechanical Engineering (Brasilia, 2007)*, edited by P. E. Miyagi et al., ABCM Symposium Series in Mechatronics **3**, Associação Brasileira de Engenharia e Ciências Mecânicas, Rio de Janeiro, 2008.
- [Bas and Boyaci 2005] D. Bas and I. H. Boyaci, "Modeling and optimization, I: Usability of response surface methodology", *J. Food Eng.* **78**:3 (2005), 836–845.
- [Broughton et al. 2000] W. R. Broughton, M. J. Lodeiro, G. D. Sims, B. Zeqiri, M. Hodnett, R. A. Smith, and L. D. Jones, "Standardised procedures for ultrasonic inspection of polymer matrix composites", in *ICCM-12 Europe: 12th International Conference on Composite Materials (Paris, 1999)*, edited by M. Thierry and A. Vautrin, ICCM, Tours, 2000. Paper #1030.
- [Buynak et al. 1989] C. F. Buynak, T. J. Moran, and R. W. Martin, "Delamination and crack imaging in graphite-epoxy composites", *Mater. Eval.* **47** (1989), 438–441.
- [Cantwell and Morton 1992] W. J. Cantwell and J. Morton, "The significance of damage and defects and their detection in composite materials: a review", *J. Strain Anal. Eng. Des.* **27**:1 (1992), 29–42.
- [Cawley and Adams 1979] P. Cawley and R. D. Adams, "The location of defects in structures from measurements of natural frequencies", *J. Strain Anal. Eng. Des.* **14**:2 (1979), 49–57.

- [Chakraborty 2005] D. Chakraborty, "Artificial neural network based delamination prediction in laminated composites", *Mater. Des.* **26**:1 (2005), 1–7.
- [Chen et al. 2004] H.-P. Chen, H. Le, J.-H. Kin, and A. Chattopadhyay, "Delamination detection problems using a combined genetic algorithm and neural network technique", in *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference* (Albany, NY, 2004), AIAA, Reston, VA, 2004. AIAA #2004-4397.
- [Dackermann 2010] U. Dackermann, *Vibration-based damage identification methods for civil engineering structures using artificial neural networks*, Ph.D. thesis, University of Technology, Sydney, 2010, Available at <http://hdl.handle.net/2100/1127>.
- [Deb et al. 2002] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II", *IEEE Trans. Evol. Comput.* **6**:2 (2002), 182–197.
- [Delashmit and Manry 2005] W. H. Delashmit and M. T. Manry, "Recent developments in multilayer perceptron neural networks", in *Proceedings of the 7th Annual Memphis Area Engineering and Science Conference* (Memphis, TN, 2005), MAESC, Memphis, TN, 2005.
- [Della and Shu 2005] C. N. Della and D. Shu, "Vibration of beams with double delaminations", *J. Sound Vib.* **282** (2005), 919–935.
- [Doebbling et al. 1998] S. W. Doebbling, C. R. Farrar, M. B. Prime, and D. W. Shevitz, "A review of damage identification methods that examine changes in dynamic properties", *Shock Vib. Digest* **30**:2 (1998), 91–105.
- [Fang and Tang 2005] X. Fang and J. Tang, "Frequency response based damage detection using principal component analysis", pp. 407–412 in *Proceedings of the IEEE International Conference on Information Acquisition* (Hong Kong and Macau, 2005), edited by M. Meng et al., IEEE, Piscataway, NJ, 2005.
- [Harrison and Butler 2001] C. Harrison and R. Butler, "Locating delaminations in composite beams using gradient techniques and a genetic algorithm", *AIAA J.* **39**:7 (2001), 1383–1389.
- [Herath et al. 2010] M. T. Herath, K. Bandyopadhyay, and J. D. Logan, "Modelling of delamination damage in composite beams", pp. 1189–1198 in *6th Australasian Congress on Applied Mechanics* (Perth, 2010), edited by K. Teh et al., Engineers Australia, Perth, 2010.
- [Ihesiulor et al. 2012a] O. K. Ihesiulor, K. Shankar, Z. Zhang, and T. Ray, "Delamination detection using methods of computational intelligence", pp. 303–310 in *Proceedings of the Sixth Global Conference on Power Control and Optimization*, edited by N. Barsoum et al., American Institute of Physics Conference Series **1499**, August 2012.
- [Ihesiulor et al. 2012b] O. K. Ihesiulor, K. Shankar, Z. Zhang, and T. Ray, "Effectiveness of artificial neural networks and surrogate-assisted optimization techniques in delamination detection for structural health monitoring", pp. 198–203 in *Proceedings of the IASTED International Conference on Modelling and Simulation*, edited by S. Al Zahir et al., July 2012. 783-042.
- [Ishak et al. 2001] S. I. Ishak, G. R. Liu, H. M. Shang, and S. P. Lim, "Locating and sizing of delamination in composite laminates using computational and experimental methods", *Compos. B Eng.* **32**:4 (2001), 287–298.
- [Islam and Craig 1994] A. S. Islam and K. C. Craig, "Damage detection in composite structures using piezoelectric materials (and neural net)", *Smart Mater. Struct.* **3**:3 (1994), 318–328.
- [Kanungo et al. 2002] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation", *IEEE Trans. Pattern Anal. Mach. Intell.* **24**:7 (2002), 881–892.
- [Kessler et al. 2005] S. Kessler, K. Amaratunga, and B. Wardle, "An assessment of durability requirements for aircraft structural health monitoring sensors", in *Proceedings of the 5th International Workshop on Structural Health Monitoring* (Stanford, CA, 2005), edited by F.-K. Chang, DEStech Publications, Lancaster, PA, 2005.
- [Kim and Yiu 2004] H. L. Kim and W. M. Yiu, "Delamination detection in smart composite beams using Lamb waves", *Smart Mater. Struct.* **13** (2004), 544–551.
- [MacQueen 1967] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations", pp. 281–297 in *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, I: Statistics* (Berkeley, CA, 1965–1966), edited by L. M. Le Cam and J. Neyman, University of California Press, Berkeley, CA, 1967.
- [Mahdi and El Kadi 2007] E. S. Mahdi and H. El Kadi, "Crushing behavior of laterally compressed composite elliptical tubes: experiments and predictions using artificial neural networks", *Compos. Struct.* **83**:4 (2007), 399–412.

- [Montgomery 2005] D. C. Montgomery, *Design and analysis of experiments: response surface method and designs*, 6th ed., Wiley, Hoboken, NJ, 2005. 8th ed. published in 2012.
- [Mufti 2001] A. Mufti, *Guidelines for structural health monitoring*, ISIS Canada, Winnipeg, 2001.
- [Mujumdar and Suryanarayan 1988] P. J. Mujumdar and S. Suryanarayan, "Flexural vibrations of beams with delaminations", *J. Sound Vib.* **125**:3 (1988), 441–461.
- [Nag et al. 2002] A. Nag, D. R. Mahapatra, and S. Gopalakrishnan, "Identification of delamination in composite beams using spectral estimation and a genetic algorithm", *Smart Mater. Struct.* **11**:6 (2002), 899–908.
- [Okafor et al. 1996] A. C. Okafor, K. Chandrashekhara, and Y. P. Jiang, "Delamination prediction in composite beams with built-in piezoelectric devices using modal analysis and neural network", *Smart Mater. Struct.* **5**:3 (1996), 338–347.
- [Pardoen 1989] G. C. Pardoen, "Effect of delamination on the natural frequency of composite laminates", *J. Compos. Mater.* **23**:12 (1989), 1200–1215.
- [Ramanamurthy and Chandrasekaran 2011] E. V. V. Ramanamurthy and K. Chandrasekaran, "Vibration analysis on a composite beam to identify damage and damage severity using finite element method", *IJEST* **3**:7 (2011), 5865–5888.
- [Ramkumar and Kulkarni 1979] R. Ramkumar and S. Kulkarni, "Free vibration frequencies of a delaminated beam", in *Reinforcing the future: proceedings of the 34th Annual Conference, Reinforced Plastics/Composites Institute* (New Orleans, LA, 1979), Society of the Plastics Industry, New York, 1979.
- [Rao 2007] M. R. P. D. Rao, "Review of nondestructive evaluation techniques for FRP composite structural components", Master's thesis, West Virginia University, Morgantown, WV, 2007, Available at <http://tinyurl.com/Rao-2007-masters-thesis>.
- [Ray et al. 2001] T. Ray, T. Kang, and K. C. Seow, "Multi-objective design optimization by an evolutionary algorithm", *Eng. Optim.* **33**:4 (2001), 399–424.
- [Sastry et al. 2006] K. Sastry, D. D. Johnson, A. L. Thompson, D. E. Goldberg, T. J. Martinez, J. Leiding, and J. Owens, "Multi-objective genetic algorithms for multiscaling excited state direct dynamics in photochemistry", pp. 1745–1752 in *GECCO '06: Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation* (Seattle, WA, 2006), edited by M. Cattolico, ACM, New York, 2006.
- [Schittkowski 1986] K. Schittkowski, "NLPQL: a FORTRAN subroutine solving constrained nonlinear programming problems", *Ann. Oper. Res.* **5**:4 (1986), 485–500.
- [Shang et al. 2001] Y. Shang, Y. Wan, M. P. J. Fromherz, and L. S. Crawford, "Toward adaptive cooperation between global and local solvers for continuous constraint problems", in *International Conference on Constraint Programming, CP 01: Proceedings of the CoSolv01 Workshop on Cooperative Solvers in Constraint Programming* (Paphos, 2001), 2001.
- [Su et al. 2005] Z. Su, H.-Y. Ling, L.-M. Zhou, K.-T. Lau, and L. Ye, "Efficiency of genetic algorithms and artificial neural networks for evaluating delamination in composite structures using fibre Bragg grating sensors", *Smart Mater. Struct.* **14**:6 (2005), 1541–1553.
- [Sul et al. 2011] J. H. Sul, B. G. Prusty, and T. Ray, "Prediction of low cycle fatigue life of short fibre composites at elevated temperatures using surrogate modelling", *Compos. B Eng.* **42**:6 (2011), 1453–1460.
- [Todoroki 2001] A. Todoroki, "The effect of number of electrodes and diagnostic tool for monitoring the delamination of CFRP laminates by changes in electrical resistance", *Compos. Sci. Technol.* **61**:13 (2001), 1871–1880.
- [Tsoukalas and Uhrig 1997] L. H. Tsoukalas and R. E. Uhrig, *Fuzzy and neural approaches in engineering*, Wiley, New York, 1997.
- [Valoor and Chandrashekhara 2000] M. T. Valoor and K. Chandrashekhara, "A thick composite-beam model for delamination prediction by the use of neural networks", *Compos. Sci. Technol.* **60**:9 (2000), 1773–1779.
- [Wang and Liu 1982] J. Wang and Y. Liu, "Vibration of split beams", *J. Sound Vib.* **84**:4 (1982), 491–502.
- [Watkins et al. 2002] S. E. Watkins, G. W. Sanders, F. Akhavan, and K. Chandrashekhara, "Modal analysis using fiber optic sensors and neural networks for prediction of composite beam delamination", *Smart Mater. Struct.* **11**:4 (2002), 489–495.
- [Zhang et al. 2010] Z. Zhang, K. Shankar, M. Tahtali, and E. V. Morozov, "Vibration modelling of composite laminates with delamination damage", in *ICA 2010: Proceedings of 20th International Congress on Acoustics* (Sydney, 2010), edited by M. Burgess et al., Australian Acoustical Society, Kensington, 2010.

[Zheng et al. 2011] S.-J. Zheng, Z.-Q. Li, and H.-T. Wang, "A genetic fuzzy radial basis function neural network for structural health monitoring of composite laminated beams", *Expert Syst. Appl.* **38**:9 (2011), 11837–11842.

Received 1 Nov 2012. Accepted 13 Jan 2013.

OBINNA K. IHESIULOR: [obinna.ihesiulor@student.adfa.edu.au](mailto:obinna.ihesiulor@student.adfa.edu.au)

*School of Engineering and Information Technology, University of New South Wales at the Australian Defense Force Academy, Northcott Drive, Canberra ACT 2600, Australia*

KRISHNA SHANKAR: [k.shankar@adfa.edu.au](mailto:k.shankar@adfa.edu.au)

*School of Engineering and Information Technology, University of New South Wales at the Australian Defense Force Academy, Northcott Drive, Canberra ACT 2600, Australia*

ZHIFANG ZHANG: [zhifang.zhang@student.adfa.edu.au](mailto:zhifang.zhang@student.adfa.edu.au)

*School of Engineering and Information Technology, University of New South Wales at the Australian Defense Force Academy, Northcott Drive, Canberra ACT 2600, Australia*

TAPABRATA RAY: [t.ray@adfa.edu.au](mailto:t.ray@adfa.edu.au)

*School of Engineering and Information Technology, University of New South Wales at the Australian Defense Force Academy, Northcott Drive, Canberra ACT 2600, Australia*



## EVALUATION OF THE EFFECTIVE ELASTIC MODULI OF PARTICULATE COMPOSITES BASED ON MAXWELL'S CONCEPT OF EQUIVALENT INHOMOGENEITY: MICROSTRUCTURE-INDUCED ANISOTROPY

VOLODYMYR I. KUSHCH, SOFIA G. MOGILEVSKAYA,  
HENRYK K. STOLARSKI AND STEVEN L. CROUCH

Maxwell's concept of equivalent inhomogeneity is employed for evaluating the effective elastic properties of macroscopically anisotropic particulate composites with isotropic phases. The effective anisotropic elastic properties of the material are obtained by comparing the far-field solutions for the problem of a finite cluster of isotropic particles embedded in an infinite isotropic matrix with those for the problem of a single anisotropic equivalent inhomogeneity embedded in the same matrix. The former solutions precisely account for the interactions between all particles in the cluster and for their geometrical arrangement. Illustrative examples involving periodic (simple cubic) and random composites suggest that the approach provides accurate estimates of their effective elastic moduli.

### 1. Introduction

This paper examines Maxwell's concept of equivalent inhomogeneity in the context of the effective elastic properties of macroscopically anisotropic particulate composites with spherical particles. The matrix and the particles are assumed to be isotropic, so the overall anisotropy is entirely due to the geometrical arrangement of particles. Maxwell [1873] originally proposed the concept for evaluating the effective electrical conductivity of isotropic particulate composites. He obtained an approximation formula by equating "the potential at a great distance from the sphere" for two problems: a finite cluster of conducting spherical particles embedded in an infinite conducting matrix, and a single equivalent sphere embedded in the same matrix. The formula did not account for the interaction between the particles and, therefore, for their geometrical arrangement. According to Maxwell, the formula was only valid for materials with low volume fractions of particles. Nonetheless, the formula and analogous estimates (for example, for dielectric, magnetic, optical, and elastic properties) remain extremely popular [Milton 2002; Torquato 2002; McCartney and Kelly 2008; McCartney 2010; Levin et al. 2012] due to their simple, ready-to-use analytical nature. The accuracy of these analytical formulas has been discussed in several publications (for example, [McCartney and Kelly 2008; McCartney 2010; Mogilevskaya et al. 2012]).

It was recently brought to our attention that a concept somewhat similar to that of Maxwell is widely used in the geophysics community. Kuster and Toksöz [1974] (who apparently were not aware of Maxwell's approach) suggested equating the displacement fields for waves scattered by the equivalent

---

*Keywords:* Maxwell's methodology, anisotropic elastic moduli, particulate composites, spherical particles, spherical harmonics.

spherical inhomogeneity and those by a cluster of spherical or spheroidal reinforcements in order to evaluate the effective elastic moduli of two-phase composites. These authors made the assumption that “multiple scattering effects are negligible”, which allowed them to neglect interactions between the reinforcements in the cluster and its geometry. The method is discussed in detail in [Berriman and Berge 1996], where it was compared with the Mori–Tanaka approach. Recently, Weng [2010] and Levin et al. [2012] have shown that the approach of [Kuster and Toksöz 1974] is a dynamical analog of the Maxwell scheme.

In a series of recent papers Maxwell’s concept was modified to evaluate the effective elastic properties of transversely isotropic composites [Mogilevskaya et al. 2010a; 2010b; 2012; Mogilevskaya and Crouch 2013], the thermal properties of isotropic particulate composites [Koroteeva et al. 2010; Mogilevskaya et al. 2011], and the viscoelastic properties of transversely isotropic composites [Pyatigorets and Mogilevskaya 2011]. The modified concept allowed for a precise account of both the interactions among the constituencies in the cluster and their geometrical arrangement. The comparisons of the estimates obtained using Maxwell’s modified approach with benchmark results for periodic and random composites and with the exact solutions demonstrated the estimates’ accuracy even for materials with high volume fractions. It has been suggested [Mogilevskaya et al. 2010b; Mogilevskaya and Crouch 2013] that the general methodology presented in those papers would be formally applicable to composite materials exhibiting anisotropic behavior if the solution of a single inhomogeneity with a corresponding degree of anisotropy was used as the reference solution.

The objective of the present paper is to demonstrate that Maxwell’s concept is applicable to macroscopically anisotropic particulate composites with isotropic phases (matrix and spherical particles). The formulation involves the solutions for two problems: an infinite isotropic matrix containing a spherical inhomogeneity with an arbitrary degree of anisotropy, and an infinite matrix containing a cluster of nonoverlapping isotropic elastic spherical particles. The effective stiffness tensor of the composite is evaluated by comparing the far-field asymptotic behavior of the displacements for both solutions.

The method of solving the problem of a single anisotropic ellipsoidal inhomogeneity was outlined in Eshelby [1961]; see also [Mura 1987] and the references therein. Closed-form solutions have been reported for the particular cases of material symmetries, for example, by Huang [1968], who considered a problem of a single anisotropic spherical inhomogeneity which possessed cubic symmetry.

The problem of a finite cluster of elastic spherical particles embedded in an infinite elastic matrix has been studied in several publications, but mostly under various simplified assumptions, for example, the assumption that the strains inside the inhomogeneities are uniform [Molinari and El Mouden 1996] and the equivalent transformational strains are polynomial [Moschovidis and Mura 1975]; or the assumption that the interactions are governed by the average equivalent transformational strains [Rodin and Hwang 1991; Shen and Yi 2001]. In several papers the interactions between the inhomogeneities were accounted for in different approximate manners (for example, pairwise interactions in [Ju and Yanase 2010] and eight-particle interactions in [Yin and Sun 2005]). The problem was also solved numerically, for example, in [Fu et al. 1998] with the boundary element method. Complete multipole-type analytical solutions for the problem of a finite cluster of isotropic inhomogeneities were obtained in [Golovchan et al. 1993] for spherical particles, [Kushch 1996] for aligned spheroidal inhomogeneities, [Kushch 1998] for arbitrary oriented spheroidal inhomogeneities, and [Kushch et al. 2011] for spherical inhomogeneities with imperfect, Gurtin–Murdoch type, interfaces.



The solutions for spatially periodic media have been reported for cubic arrays of rigid spheres [Nunan and Keller 1984], for cubic arrays of elastic spheres [Sangani and Lu 1987], for a medium with arbitrary periodic arrays of elastic spheres [Kushch 1985; 1987; Sangani and Mo 1997], for a medium with periodic arrays of elastic spheroids [Kushch 1997], and for a transversely isotropic medium with finite or infinite periodic arrays of transversely isotropic spheres [Kushch 2003; Kushch and Sevostianov 2004]. These solutions were used to calculate the effective elasticity tensors of periodic and quasirandom composites.

The three-dimensional effective elastic properties (isotropic and anisotropic) of particulate composite and porous materials were also calculated in [Nemat-Nasser and Taya 1981; Nemat-Nasser et al. 1982; Iwakuma and Nemat-Nasser 1983; Luciano and Barbero 1994; Cohen and Bergman 2003; Cohen 2004]. Those estimates were obtained using the simplifying assumptions of constant equivalent strains within each inhomogeneity. Torquato [1997] obtained exact series expansions for the effective stiffness tensor of macroscopically anisotropic, two-phase composite media in terms of the powers of the “elastic polarizabilities”. Numerical estimates are also available, for example, with the finite element method in [Segurado and Llorca 2002; 2006; Zohdi and Wriggers 2005] and with boundary element method in [Grzhibovskis et al. 2010]. In addition, various effective medium theories and variational bounds have been generalized to estimate the overall three-dimensional anisotropic elastic properties (for example, [Willis 1977; Benveniste 1987; Ponte Castañeda and Willis 1995; Milton 2002; Torquato 2002]).

In the present paper the cluster problem is solved semianalytically using the multipole expansion method of [Kushch et al. 2011]. The reference solution for a single anisotropic spherical inhomogeneity is rederived in a form more suitable for comparison with the cluster problem. A numerical procedure for calculating the effective stiffness tensor is described for materials with an arbitrary degree of overall anisotropy. Closed-form expressions are given for the particular case of cubic symmetry. The effective moduli obtained with the generalized Maxwell method are compared with those obtained by periodic homogenization [Kushch 1987; Sangani and Lu 1987] and with the various approximate estimates and bounding methods.

The paper is structured as follows. Sections 2–4 summarize the statement of the problem, governing equations, and numerical solution (with details provided in the Appendix). These are followed by illustrative examples involving periodic (simple cubic) and random composites in Section 5 and conclusions in Section 6.

## 2. The equivalent anisotropic inhomogeneity problem

Consider an infinite isotropic elastic matrix with shear modulus  $\mu_0$  and Poisson’s ratio  $\nu_0$  containing an anisotropic spherical elastic inhomogeneity of radius  $R_{\text{eff}}$  perfectly bonded to the matrix material. The entire system is subjected to the uniform far-field strain  $\mathbf{E} = E_{ij}\mathbf{i}_i\mathbf{i}_j$ . This is the standard Eshelby problem whose analytical solution is outlined elsewhere (for example, in [Eshelby 1961; Mura 1987]). Below it is derived in a somewhat different form, more suitable for our purposes. The derivation procedure is essentially that of [Kushch et al. 2011], with minor modifications.

The displacement vector  $\mathbf{u}^{(0)}$  in the matrix domain is sought as a sum of the far field and the disturbance field caused by the inhomogeneity

$$\mathbf{u}^{(0)}(\mathbf{r}) = \mathbf{E} \cdot \mathbf{r} + \mathbf{u}_{\text{dis}}(\mathbf{r}), \quad (1)$$

where  $\mathbf{r}$  is a position vector relative to the Cartesian coordinate system with origin at the center of the inhomogeneity, and  $\mathbf{u}_{\text{dis}}$  is the displacement disturbance field that should obey the condition  $\mathbf{u}_{\text{dis}}(\mathbf{r}) \rightarrow 0$  as  $\|\mathbf{r}\| \rightarrow \infty$ . To assure that this condition is satisfied,  $\mathbf{u}_{\text{dis}}$  is taken in the form

$$\mathbf{u}_{\text{dis}}(\mathbf{r}) = \sum_{i,t,s} B_{ts}^{(i)} \mathbf{U}_{ts}^{(i)}(\mathbf{r}) \quad \left( \sum_{i,t,s} = \sum_{i=1}^3 \sum_{t=0}^{\infty} \sum_{s=-t}^t \right), \quad (2)$$

where  $B_{ts}^{(i)}$  are the unknown complex coefficients, and the complex-value irregular vector functions  $\mathbf{U}_{ts}^{(i)}$  ( $i = 1, 2, 3$ ) are defined by (A.7). Specifically,  $\mathbf{U}_{ts}^{(1)}$  are the potential vectors (gradients of scalar potential),  $\mathbf{U}_{ts}^{(2)}$  are harmonic vectors with nonzero curl, and  $\mathbf{U}_{ts}^{(3)}$  are the biharmonic vectors with harmonic divergence. In particular,  $\mathbf{U}_{1s}^{(2)}$  and  $\mathbf{U}_{1s}^{(3)}$  represent the displacements due to the concentrated moment and force, respectively (see the Appendix). For a single inhomogeneity problem the series of (2) involves only the functions with  $t \leq 2$ , see [Kushch et al. 2011]. In addition, equilibrium conditions for the inhomogeneity, with the resultant force  $T = 0$  and the resultant torque  $M = 0$ , require that some coefficients with  $t \leq 2$  also be excluded (see the Appendix). Thus  $\mathbf{u}_{\text{dis}}$  involves only the functions possessing nonzero vector dipole moment:

$$\mathbf{u}_{\text{dis}}(\mathbf{r}) = B_{00}^{(1)} \mathbf{U}_{00}^{(1)}(\mathbf{r}) + \sum_{|s| \leq 2} B_{2s}^{(3)} \mathbf{U}_{3s}^{(3)}(\mathbf{r}). \quad (3)$$

The Cartesian projections of  $\mathbf{u}_{\text{dis}}$  are real numbers, which implies that  $B_{2,-s}^{(3)} = (-1)^s \overline{B_{2s}^{(3)}}$ . Thus, the total number of unknown coefficients includes two real ( $B_{00}^{(1)}$  and  $B_{20}^{(3)}$ ) and two complex ( $B_{21}^{(3)}$  and  $B_{22}^{(3)}$ ) coefficients in (3).

The linear far-field displacement field  $\mathbf{u}_{\text{far}} = \mathbf{E} \cdot \mathbf{r}$  is expressed as follows [Kushch et al. 2011]:

$$\mathbf{E} \cdot \mathbf{r} = c_{00}^{(3)} \mathbf{u}_{00}^{(3)}(\mathbf{r}) + \sum_{|s| \leq 2} c_{2s}^{(1)} \mathbf{u}_{2s}^{(1)}(\mathbf{r}), \quad (4)$$

where  $\mathbf{u}_{ts}^{(i)}$  are the regular vector functions defined by (A.5), and the coefficients  $c_{ts}^{(i)}$  are defined as follows

$$\begin{aligned} c_{00}^{(3)} &= \frac{(E_{11} + E_{22} + E_{33})}{3\gamma_0(v_0)}, & c_{20}^{(1)} &= \frac{(2E_{33} - E_{11} - E_{22})}{3}, \\ c_{21}^{(1)} &= -\overline{c_{2,-1}^{(1)}} = E_{13} - iE_{23}, & c_{22}^{(1)} &= \overline{c_{2,-2}^{(1)}} = E_{11} - E_{22} - 2iE_{12}, \end{aligned} \quad (5)$$

where the coefficients  $\gamma_t = \gamma_t(v_0)$  are given by (A.6) and  $i^2 = -1$ .

It is well-known [Eshelby 1961] that the strains  $D_{ij}$  in the inhomogeneity are uniform, thus the displacement  $\mathbf{u}^{(1)}(\mathbf{r})$  inside the inhomogeneity can be presented as the following linear function of Cartesian coordinates ( $\mathbf{r} = x_j \mathbf{i}_j$ ):

$$\mathbf{u}^{(1)}(\mathbf{r}) = \mathbf{D} \cdot \mathbf{r} = D_{ij} x_j \mathbf{i}_i. \quad (6)$$

By analogy with (4),  $\mathbf{u}^{(1)}(\mathbf{r})$  can be written as

$$\mathbf{u}^{(1)}(\mathbf{r}) = d_{00}^{(3)} \mathbf{u}_{00}^{(3)}(\mathbf{r}) + \sum_{|s| \leq 2} d_{2s}^{(1)} \mathbf{u}_{2s}^{(1)}(\mathbf{r}), \quad (7)$$

where

$$\begin{aligned}\gamma_0 d_{00}^{(3)} &= \frac{(D_{11} + D_{22} + D_{33})}{3}, & d_{20}^{(1)} &= \frac{(2D_{33} - D_{11} - D_{22})}{3}, \\ d_{21}^{(1)} &= D_{13} - iD_{23}, & d_{22}^{(1)} &= D_{11} - D_{22} - 2iD_{12},\end{aligned}\quad (8)$$

and  $d_{2,-s}^{(1)} = (-1)^s \overline{d_{2s}^{(1)}}$ . The corresponding uniform stress tensor inside the inhomogeneity is

$$\boldsymbol{\sigma}^{(1)}(\mathbf{r}) = \mathbf{S} = S_{ij} \mathbf{i}_i \mathbf{i}_j = \mathbf{C}^* : \mathbf{D}, \quad (9)$$

where  $\mathbf{C}^*$  is the stiffness tensor (anisotropic, in the general case).

The twelve unknowns that govern the problem of a single inhomogeneity include six real coefficients  $D_{ij}$  of (8), and two real ( $B_{00}^{(1)}$  and  $B_{20}^{(3)}$ ) and two complex ( $B_{21}^{(3)}$  and  $B_{22}^{(3)}$ ) coefficients of (3). They can be obtained from the following conditions of perfect bonding at the matrix/inhomogeneity interface  $S_{\text{eff}}(r = R_{\text{eff}})$ :

$$[[\mathbf{u}]]_{S_{\text{eff}}} = 0, \quad [[\mathbf{T}_r(\mathbf{u})]]_{S_{\text{eff}}} = 0, \quad (10)$$

where  $[[f]]_{S_{\text{eff}}} = (f^{(0)} - f^{(1)})_{S_{\text{eff}}}$  is a jump of the quantity  $f$  across the interface  $S_{\text{eff}}$  and  $\mathbf{T}_r(\mathbf{u})$  is the traction vector at  $S_{\text{eff}}$ . The latter has the following form ( $\mathbf{n} = \mathbf{e}_r = n_j \mathbf{i}_j$ ):

$$\mathbf{T}_r(\mathbf{u}^{(1)}) = \mathbf{S} \cdot \mathbf{e}_r = S_{ij} \mathbf{i}_i \mathbf{i}_j \cdot \mathbf{e}_r = S_{ij} n_j \mathbf{i}_i. \quad (11)$$

In view of  $\mathbf{r} = r\mathbf{e}_r$ , the tractions  $\mathbf{T}_r(\mathbf{u}^{(1)})$  can be expanded in series analogous to that of (7):

$$r\mathbf{T}_r(\mathbf{u}^{(1)}) = s_{00}^{(3)} \mathbf{u}_{00}^{(3)}(\mathbf{r}) + \sum_{|s| \leq 2} s_{2s}^{(1)} \mathbf{u}_{2s}^{(1)}(\mathbf{r}), \quad (12)$$

where

$$\begin{aligned}\gamma_0 s_{00}^{(3)} &= \frac{(S_{11} + S_{22} + S_{33})}{3}, & s_{20}^{(1)} &= \frac{(2S_{33} - S_{11} - S_{22})}{3}, \\ s_{21}^{(1)} &= S_{13} - iS_{23}, & s_{22}^{(1)} &= S_{11} - S_{22} - 2iS_{12}.\end{aligned}\quad (13)$$

In order to fulfill the interface conditions of (10), we express  $\mathbf{u}^{(0)}$  and  $\mathbf{u}^{(1)}$  in terms of vector spherical harmonics  $\mathbf{S}_{ts}^{(i)}$  (see (A.1)) and use their orthogonality at the spherical surface  $S_{\text{eff}}$ . With the aid of (A.5) and (A.7) we find that

$$\mathbf{u}_{\text{far}}(\mathbf{r}) = \gamma_0 c_{00}^{(3)} r \mathbf{S}_{00}^{(3)} + \sum_{|s| \leq 2} c_{2s}^{(1)} \frac{r}{(2+s)!} (\mathbf{S}_{2s}^{(1)} + 2\mathbf{S}_{2s}^{(3)}), \quad (14)$$

$$\mathbf{u}_{\text{dis}}(\mathbf{r}) = -B_{00}^{(1)} \frac{1}{r^2} \mathbf{S}_{00}^{(3)} + \sum_{|s| \leq 2} \frac{(2-s)!}{r^2} \left[ B_{2s}^{(1)} \frac{1}{r^2} (\mathbf{S}_{2s}^{(1)} - 3\mathbf{S}_{2s}^{(3)}) + B_{2s}^{(3)} (\beta_{-3} \mathbf{S}_{2s}^{(1)} + \gamma_{-3} \mathbf{S}_{ts}^{(3)}) \right], \quad (15)$$

and

$$\mathbf{u}^{(1)}(\mathbf{r}) = \mathbf{D} \cdot \mathbf{r} = \gamma_0 d_{00}^{(3)} r \mathbf{S}_{00}^{(3)} + \sum_{|s| \leq 2} d_{2s}^{(1)} \frac{r}{(2+s)!} (\mathbf{S}_{2s}^{(1)} + 2\mathbf{S}_{2s}^{(3)}), \quad (16)$$

where the coefficients  $\gamma_t = \gamma_t(v_0)$  and  $\beta_t = \beta_t(v_0)$  involved in (15) are defined by (A.6).

By equating the coefficients of  $\mathbf{S}_{00}^{(3)}$  on both sides of the equality  $\mathbf{u}^{(0)} = \mathbf{u}_{\text{far}} + \mathbf{u}_{\text{dis}} = \mathbf{u}^{(1)}$ , we obtain the algebraic equation

$$\gamma_0 c_{00}^{(3)} - \frac{1}{R_{\text{eff}}^3} B_{00}^{(1)} = \gamma_0 d_{00}^{(3)}. \quad (17)$$

Similarly, by equating the coefficients of  $\mathbf{S}_{2s}^{(1)}$  and  $\mathbf{S}_{2s}^{(3)}$  ( $s = 0, 1, 2$ ), one obtains

$$\begin{aligned} 2c_{2s}^{(1)} + \frac{(2+s)!(2-s)!}{R_{\text{eff}}^3} \left( \frac{2}{R_{\text{eff}}^2} B_{2s}^{(1)} + 2\beta_{-3} B_{2s}^{(3)} \right) &= 2d_{2s}^{(1)}, \\ 2c_{2s}^{(1)} + \frac{(2+s)!(2-s)!}{R_{\text{eff}}^3} \left( -\frac{3}{R_{\text{eff}}^2} B_{2s}^{(1)} + \gamma_{-3} B_{2s}^{(3)} \right) &= 2d_{2s}^{(1)}, \end{aligned} \quad (18)$$

In view of  $\gamma_{-3} - 2\beta_{-3} = 1$ , we find that

$$\frac{B_{2s}^{(1)}}{R_{\text{eff}}^2} = \frac{1}{5} B_{2s}^{(3)}. \quad (19)$$

By eliminating  $B_{2s}^{(1)}$  using (19) and taking into account that  $\beta_{-3}(\nu) = (1 - 2\nu)/3$ , (18) reduces to

$$c_{2s}^{(1)} + \frac{(2+s)!(2-s)!}{R_{\text{eff}}^3} \frac{(8 - 10\nu_0)}{15} B_{2s}^{(3)} = d_{2s}^{(1)}. \quad (20)$$

The same procedure is employed to fulfill the second condition in (10), namely  $\mathbf{T}_r(\mathbf{u}^{(0)}) = \mathbf{T}_r(\mathbf{u}^{(1)})$ . With the aid of (A.9) and (A.11) we write

$$\mathbf{T}_r(\mathbf{u}_{\text{far}}) = c_{00}^{(3)} 2\mu_0 g_0(\nu_0) \mathbf{S}_{00}^{(3)} + 2\mu_0 \sum_{|s| \leq 2} c_{2s}^{(1)} \frac{1}{(2+s)!} (\mathbf{S}_{2s}^{(1)} + 2\mathbf{S}_{2s}^{(3)}), \quad (21)$$

$$\mathbf{T}_r(\mathbf{u}_{\text{dis}}) = B_{00}^{(1)} \frac{4\mu_0}{R_{\text{eff}}^3} \mathbf{S}_{00}^{(3)} + 2\mu_0 \sum_{|s| \leq 2} \frac{(2-s)!}{r^3} \left[ -B_{2s}^{(1)} 4 \frac{1}{R_{\text{eff}}^2} (\mathbf{S}_{2s}^{(1)} - 3\mathbf{S}_{2s}^{(3)}) + B_{2s}^{(3)} (b_{-3} \mathbf{S}_{2s}^{(1)} + g_{-3} \mathbf{S}_{2s}^{(3)}) \right], \quad (22)$$

and

$$\mathbf{T}_r(\mathbf{u}^{(1)}) = \gamma_0 s_{00}^{(3)} \mathbf{S}_{00}^{(3)} + \sum_{|s| \leq 2} s_{2s}^{(1)} \frac{1}{(2+s)!} (\mathbf{S}_{2s}^{(1)} + 2\mathbf{S}_{2s}^{(3)}). \quad (23)$$

The coefficients  $b_t = b_t(\nu_0)$  and  $g_t = g_t(\nu_0)$  involved in (22) are defined by (A.10). By using (18), as well as the equalities  $2b_{-3} - g_{-3} = 4$  and  $b_{-3}(\nu) = (1 + \nu)/3$ , we obtain the following linear equations:

$$2\mu_0 g_0 c_{00}^{(3)} + \frac{4\mu_0}{R_{\text{eff}}^3} B_{00}^{(1)} = \gamma_0 s_{00}^{(3)}, \quad (24)$$

$$c_{2s}^{(1)} - \frac{(2+s)!(2-s)!}{R_{\text{eff}}^3} \frac{(7 - 5\nu_0)}{15} B_{2s}^{(3)} = \frac{s_{2s}^{(1)}}{2\mu_0} \quad (s = 0, 1, 2). \quad (25)$$

The system of (17), (18), (24), and (25) represents a complete set of four real and four complex algebraic equations (twelve real equations in total) needed to find all the unknown coefficients (six real

coefficients  $D_{ij}$  of (8) and two real ( $B_{00}^{(1)}$  and  $B_{20}^{(3)}$ ) and two complex ( $B_{21}^{(3)}$  and  $B_{22}^{(3)}$ ) coefficients of (3)). This system is uniquely resolved to get all the series expansion coefficients in (2) and (7).

### 3. The cluster problem

Consider an infinite elastic isotropic matrix with shear modulus  $\mu_0$  and Poisson's ratio  $\nu_0$  containing a cluster of  $N$  nonoverlapping isotropic elastic spherical particles, of the same radii  $R$  and elastic properties  $\mu_1$  and  $\nu_1$ , perfectly bonded to the matrix. In the global Cartesian coordinate frame  $Ox_1x_2x_3$ , the center of  $p$ -th particle is specified by vector  $\mathbf{R}_p$ . The entire system is subjected to the uniform far-field strain  $\mathbf{E} = E_{ij}\mathbf{i}_i\mathbf{i}_j$ . In the following this problem is referred to as a finite cluster model (FCM) of the composite.

The solution of the problem is described in detail in [Kushch et al. 2011], where it was obtained for the case of more general interface conditions. The solution procedure employs the superposition principle. Specifically,  $\mathbf{u}_{\text{dis}}(\mathbf{r})$  of (1) is sought as a superposition of the disturbance fields of (2) (vanishing at infinity) caused by each particle separately:

$$\mathbf{u}_{\text{dis}}(\mathbf{r}) = \sum_{p=1}^N \sum_{i,t,s} A_{ts}^{(i)(p)} \mathbf{U}_{ts}^{(i)}(\mathbf{r} - \mathbf{R}_p), \quad (26)$$

where  $\mathbf{U}_{ts}^{(i)}$  are the irregular vector functions (vanishing at infinity) defined by (A.7) and  $A_{ts}^{(i)(p)}$  are the multipole expansion coefficients related to the  $p$ -th particle. The displacement vector within the particle is bounded but not linear (due to the interactions between the particles in the cluster) and, therefore, is represented by the *infinite* series of the regular functions  $\mathbf{u}_{ts}^{(i)}$  defined by (A.5):

$$\mathbf{u}^{(p)}(\mathbf{r}) = \sum_{i,t,s} d_{ts}^{(i)(p)} \mathbf{u}_{ts}^{(i)}(\mathbf{r}). \quad (27)$$

An infinite system of linear equations for the unknown coefficients  $A_{ts}^{(i)(p)}$  and  $d_{ts}^{(i)(p)}$  is obtained from the interface conditions of (10), written for each particle, using the orthogonality properties of vector spherical harmonics. For this purpose, the matrix displacement  $\mathbf{u}^{(0)}$  given by (1) and (26) should be expanded in the local spherical coordinates of each specific particle with the aid of the reexpansion formulas for  $\mathbf{U}_{ts}^{(i)}$  due to shift of the coordinate frame [Kushch 1985; 2013] The system is solved numerically after the series of (26) and (27) are truncated (see [Kushch et al. 2011] for more details).

### 4. Effective stiffness of the composite using the generalized Maxwell approach

The generalized Maxwell concept of equivalent inhomogeneity implies that the effective stiffness tensor of the composite can be obtained by comparing the far-field asymptotic behavior of the displacements for the solutions obtained in Sections 2 and 3. Specifically, we equate the dipole moments of an entire cluster of particles to those of an equivalent spherical inhomogeneity with the effective elastic moduli to be found. The radius  $R_{\text{eff}}$  of the equivalent inhomogeneity is defined by Maxwell [1873] so as to preserve the volume fraction  $c$  of the inhomogeneities in the cluster, which results in

$$R_{\text{eff}}^3 = NR^3/c. \quad (28)$$

A comparison of the displacements given by (2) and (26) yields the following relation:

$$B_{i-1,s}^{(i)} = \sum_{p=1}^N A_{i-1,s}^{(i)(p)}, \quad i = 1, 2, 3. \quad (29)$$

In fact, (29) for  $i = 2$  is a trivial identity because  $B_{1s}^{(2)} = A_{1s}^{(2)} = 0$ , due to the equilibrium conditions (see the Appendix and the text preceding (3)). Equation (29) is a formal expression of the generalized Maxwell concept which consists in equating the dipole moments of an entire cluster to those of an equivalent inhomogeneity whose the effective elastic moduli are to be found.

The complete numerical procedure that utilizes the generalized Maxwell concept includes the following steps:

- (a) identification of the cluster of  $N$  inhomogeneities that adequately represent the composite material in question,
- (b) solution of the cluster problem for any given  $\mathbf{E}$  to get a whole set of the series expansion coefficients  $A_{1s}^{(i)(p)}$ ,
- (c) evaluation of the dipole coefficients  $B_{i-1,s}^{(i)}$  of the equivalent inhomogeneity from the relation (29),
- (d) substitution of the coefficients  $B_{i-1,s}^{(i)}$  into (17) and (18) to obtain the coefficients  $d_{00}^{(3)}$  and  $d_{2s}^{(1)}$  (these are later used to recover the coefficients  $D_{ij}$  from (8)),
- (e) substitution of the coefficients  $B_{i-1,s}^{(i)}$  into (24) and (25) to obtain the coefficients  $s_{00}^{(3)}$  and  $s_{2s}^{(1)}$  (these are later used to recover the coefficients  $S_{ij}$  from (13)), and
- (f) determination of the effective stiffness tensor  $\mathbf{C}^*$  from the constitutive relation  $\mathbf{S} = \mathbf{C}^* : \mathbf{D}$ . In order to determine all the components of this tensor, steps (b)–(e) need to be performed for six linearly independent realizations of the tensor  $\mathbf{E}$ , for example, for  $E_{11}\mathbf{i}_1\mathbf{i}_1$ ,  $E_{22}\mathbf{i}_2\mathbf{i}_2$ ,  $E_{33}\mathbf{i}_3\mathbf{i}_3$ ,  $E_{13}(\mathbf{i}_1\mathbf{i}_3 + \mathbf{i}_3\mathbf{i}_1)$ ,  $E_{23}(\mathbf{i}_2\mathbf{i}_3 + \mathbf{i}_3\mathbf{i}_2)$ , and  $E_{12}(\mathbf{i}_1\mathbf{i}_2 + \mathbf{i}_2\mathbf{i}_1)$ .

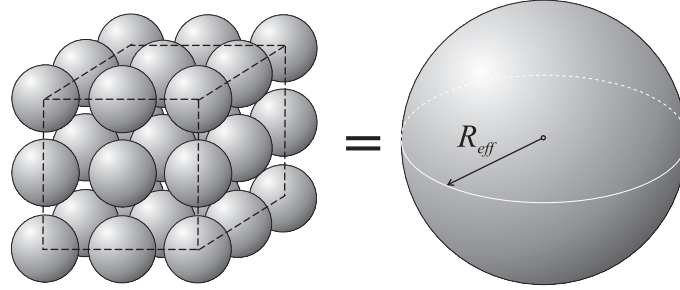
This procedure is illustrated below for the case of periodic composite material with simple cubic (SC) packing of isotropic spherical particles embedded into isotropic matrix.

**4.1. Cubic symmetry.** Consider the periodic composite with SC packing of spherical elastic particles. The cluster of  $N = n^3$  particles representing the material is shown on the left in Figure 1 ( $n$  is the number of particles in each of the coordinate directions). This composite is known to be macroscopically anisotropic and is characterized by three independent elastic moduli  $C_{1111}^*$ ,  $C_{1122}^*$ , and  $C_{1212}^*$  ( $C_{2222}^* = C_{3333}^* = C_{1111}^*$ ,  $C_{2233}^* = C_{3311}^* = C_{1122}^*$ , and  $C_{2323}^* = C_{3131}^* = C_{1212}^*$ ). Alternatively, the overall behavior can be characterized by the effective bulk modulus  $k^*$  and two shear moduli  $\mu_1^*$  and  $\mu_2^*$  as follows:

$$k^* = \frac{(C_{1111}^* + 2C_{1122}^*)}{3}, \quad \mu_1^* = \frac{(C_{1111}^* - C_{1122}^*)}{2}, \quad \mu_2^* = C_{1212}^*. \quad (30)$$

In these notations, the generalized Hooke's law  $\mathbf{S} = \mathbf{C}^* : \mathbf{D}$  is written as

$$\begin{aligned} S_{11} + S_{22} + S_{33} &= 3k^*(D_{11} + D_{22} + D_{33}), \\ 2S_{33} - S_{11} - S_{22} &= 2\mu_1^*(2D_{33} - D_{11} - D_{22}), \\ S_{11} - S_{22} - 2iS_{12} &= 2\mu_1^*(D_{11} - D_{22}) - 4i\mu_2^*D_{12}. \end{aligned} \quad (31)$$



**Figure 1.** Maxwell's equivalence principle: (left) the finite cluster model (FCM) and (right) the equivalent inhomogeneity.

**4.1.1. Effective bulk modulus.** Consider the far-field strain field characterized by nonzero components  $E_{11} = E_{22} = E_{33} = 1$ . The algebraic manipulations with (5), (8), (17), (24), and (31) and the use of the identity

$$\frac{g_0(\nu_0)}{\gamma_0(\nu_0)} = \frac{1 + \nu_0}{1 - 2\nu_0} = \frac{3k_0}{2\mu_0} \quad (32)$$

yield the following relations:

$$1 - B_{00}^{(1)} \frac{1}{R_{\text{eff}}^3} = \frac{(D_{11} + D_{22} + D_{33})}{3}, \quad 3k_0 + \frac{4\mu_0}{R_{\text{eff}}^3} B_{00}^{(1)} = 3k^* \frac{(D_{11} + D_{22} + D_{33})}{3}, \quad (33)$$

which reduce to

$$3k_0 + 4\mu_0 \frac{B_{00}^{(1)}}{R_{\text{eff}}^3} = 3k^* \left( 1 - \frac{B_{00}^{(1)}}{R_{\text{eff}}^3} \right). \quad (34)$$

Using (29) and the Maxwell definition of volume fraction given by (28), one gets

$$\frac{B_{00}^{(1)}}{R_{\text{eff}}^3} = \frac{c}{NR^3} \sum_{p=1}^N A_{00}^{(1)(p)} = c \langle A_{00}^{(1)} \rangle, \quad (35)$$

where  $\langle A_{00}^{(1)} \rangle$  is the mean dipole moment. Combination of the last two equations leads to the following expression for the effective bulk modulus  $k^*$ :

$$\frac{k^*}{k_0} = \frac{1 + (4\mu_0)/(3k_0)c \langle A_{00}^{(1)} \rangle}{1 - c \langle A_{00}^{(1)} \rangle}. \quad (36)$$

**4.1.2. Effective shear modulus  $\mu_1^*$ .** Consider the far-field strain field characterized by nonzero components  $E_{33} = 1$  and  $E_{11} = E_{22} = -E_{33}/2$ . The only nonzero coefficient of (5) is  $c_{20}^{(1)} = 1$ . Using (20) and (25) with  $s = 0$  as well as (31), one gets the following system of equations:

$$\begin{aligned} 1 + \frac{4}{R_{\text{eff}}^3} \frac{(8 - 10\nu_0)}{15} B_{20}^{(3)} &= \frac{(2D_{33} - D_{11} - D_{22})}{3}, \\ 1 - \frac{4}{R_{\text{eff}}^3} \frac{(7 - 5\nu_0)}{15} B_{20}^{(3)} &= \frac{\mu_1^* (2D_{33} - D_{11} - D_{22})}{\mu_0}. \end{aligned} \quad (37)$$

The system of (37), with the use of expression (29), yields the following solution for the effective shear modulus  $\mu_1^*$ :

$$\frac{\mu_1^*}{\mu_0} = \frac{1 - (7 - 5\nu_0)\frac{4}{15}c\langle A_{20}^{(1)} \rangle}{1 + (8 - 10\nu_0)\frac{4}{15}c\langle A_{20}^{(1)} \rangle}, \quad (38)$$

where the mean dipole  $\langle A_{2s}^{(1)} \rangle$  is defined as

$$\langle A_{2s}^{(1)} \rangle = \frac{1}{NR^3} \sum_{p=1}^N A_{2s}^{(1)(p)}. \quad (39)$$

**4.1.3. Effective shear modulus  $\mu_2^*$ .** Now consider the far-field strain field characterized by nonzero component  $E_{12} = 1$ . The only nonzero coefficient of (5) is  $c_{22}^{(1)} = -2i$ . Using (20) and (25) with  $s = 2$  as well as (31), one gets the following system of equations:

$$\begin{aligned} -2i + 24 \frac{(8 - 10\nu_0)}{15} \frac{B_{22}^{(3)}}{R_{\text{eff}}^3} &= D_{11} - D_{22} - 2iD_{12}, \\ -2i + 24 \frac{(-7 + 5\nu_0)}{15} \frac{B_{22}^{(3)}}{R_{\text{eff}}^3} &= \frac{\mu_1^*}{\mu_0} (D_{11} - D_{22}) - 2i \frac{\mu_2^*}{\mu_0} D_{12}. \end{aligned} \quad (40)$$

Using (29) and separating the imaginary parts in the equations of system (40), the following solution for the effective shear modulus  $\mu_2^*$  is obtained:

$$\frac{\mu_2^*}{\mu_0} = \frac{1 + (7 - 5\nu_0)\frac{4}{5}c\langle \text{Im} A_{22}^{(1)} \rangle}{1 - (8 - 10\nu_0)\frac{4}{5}c\langle \text{Im} A_{22}^{(1)} \rangle}. \quad (41)$$

**4.1.4. Noninteracting estimates.** In the case when the interactions between the particles are neglected, the cluster problem reduces to a set of  $N$  uncoupled single-particle problems, as in the original Maxwell approach. Hence,

$$\langle A_{00}^{(1)} \rangle = \frac{k_1 - k_0}{k_1 + \frac{4}{3}\mu_0}, \quad (42)$$

$$\langle A_{20}^{(1)} \rangle = \frac{15}{4} \frac{\mu_0 - \mu_1}{(8 - 10\nu_0)\mu_1 + (7 - 5\nu_0)\mu_0}, \quad (43)$$

$$\langle \text{Im} A_{22}^{(1)} \rangle = \frac{5}{4} \frac{\mu_0 - \mu_1}{(8 - 10\nu_0)\mu_1 + (7 - 5\nu_0)\mu_0}. \quad (44)$$

It could be shown that, in this case, the estimates of (36) for the effective bulk modulus reduce to those of [Kerner 1956; McCartney and Kelly 2008; McCartney 2010], and to one of the Hashin and Shtrikman [1963] bounds. They also coincide with the estimates of the composite sphere assemblage, the Mori–Tanaka method, and those of the generalized self-consistent method (for example, [Milton 2002; McCartney and Kelly 2008]). The estimates of (38) for the effective shear modulus  $\mu_1^*$  and the estimates of (41) for the effective shear modulus  $\mu_2^*$  coincide and reduce to those of [Kerner 1956; McCartney and Kelly 2008; McCartney 2010] and to the one of the Hashin and Shtrikman [1963] bounds. They



also coincide with the estimates of the composite sphere assemblage and the Mori–Tanaka method (for example, [Milton 2002; McCartney and Kelly 2008]).

It can be concluded that noninteracting estimates *cannot* capture the overall, microstructure induced, anisotropy of the composite material.

## 5. Numerical study

**5.1. Periodic (SC) composite.** In order to test the developed approach, we evaluate three effective elastic moduli of the periodic composite with simple cubic (SC) arrangement of particles considered in Section 4.1 (the representative cluster is depicted in Figure 1, left). Accurate values of the effective moduli of such composites are reported by Kushch [1987] and Sangani and Lu [1987], who provided complete solutions of the triple periodic homogenization problem. Their results are practically identical except for in the value of  $\mu_2^*(c)/\mu_0$  for porous material. It was suggested by Cohen [2004] that the estimates of [Sangani and Lu 1987] might be inaccurate. Therefore, in the subsequent analysis, all the periodic (SC) solutions have been recalculated with  $t_{\max} = 20$  and used as reference solutions. In addition, various bounds and approximate estimates are available (for example, [Milton 2002; Cohen and Bergman 2003; Cohen 2004]) and are used for comparison.

In order to provide comparison with the numerical data reported in the above cited works, in the following studies, as in [Kushch 1987; Sangani and Lu 1987], we assume  $\nu_0 = \nu_1 = 0.3$ . The equivalent inhomogeneity (Figure 1, right) is assumed to be anisotropic and to possess cubic symmetry of elastic properties with three independent elastic moduli defined by (30).

First simulations are conducted to analyze the convergence rate of the generalized Maxwell solution in terms of the cluster's size. The three effective elastic moduli for high-contrast composite ( $\mu_1/\mu_0 = 100$ ) are presented in Table 1 for the values  $n = 1, 2, 3, 4$  and two volume fractions  $c$  ( $c = 0.1$  and  $c = 0.5$ ). The packing limit for the SC composite is  $c_{\max} = \pi/6 \approx 0.5236$ . The extreme cases we study in this part are deliberately designed to test the developed method.

Here and below, the series of (26) and (27) are truncated at  $t_{\max} = 13$ . This number was sufficient to provide numerical solutions for dipole coefficients  $A_{0s}^{(1)(p)}$  and  $A_{2s}^{(3)(p)}$  that are accurate up to four significant digits for  $c \leq 0.45$  and up to at least three significant digits for  $c = 0.5$ . The last row of Table 1 contains values of SC results from [Kushch 1987; Sangani and Lu 1987] that are accurate up to four significant digits. It is seen from Table 1 that for  $c = 0.5$  the value of  $n = 4$  is sufficient to estimate all three moduli with accuracy of about 5%. This seems to be a good result, in view of the fact that  $c = 0.5$  is near the packing limit ( $c_{\max} = 0.5236$ ) and the contrast between the particles and the matrix is very high (almost rigid particles). For  $c = 0.1$  the convergence rate is higher, especially for the relative bulk modulus  $k^*/k_0$ . Therefore, in the subsequent numerical studies the value  $n = 4$  has been adopted.

The generalized Maxwell approach is capable of capturing the microstructure-induced overall anisotropy of the material (quite pronounced for  $c = 0.5$ ,  $\mu_1^*/\mu_2^* \approx 2$ ), whereas the noninteracting approach predicts macroscopic isotropy ( $\mu_2^* = \mu_1^*$ ).

The simulations to follow are conducted for porous ( $\mu_1 = 0$ ), moderate contrast ( $\mu_1/\mu_0 = 10$ ), and high-contrast ( $\mu_1/\mu_0 = 100$ ) composite materials and for a wide range of volume fractions. The normalized effective moduli  $k^*/k_0$ ,  $\mu_1^*(c)/\mu_0$ , and  $\mu_2^*(c)/\mu_0$  obtained using the generalized Maxwell approach, labeled FCM, are presented in Tables 2, 3, and 4, respectively, where they are compared with the SC

$n$	$c = 0.1$			$c = 0.5$		
	$k^*/k_0$	$\mu_1^*/\mu_0$	$\mu_2^*/\mu_0$	$k^*/k_0$	$\mu_1^*/\mu_0$	$\mu_2^*/\mu_0$
1	1.176	1.228	1.228	2.564	3.015	3.015
2	1.176	1.247	1.217	2.928	4.881	3.242
3	1.176	1.254	1.214	3.098	6.160	3.301
4	1.176	1.258	1.212	3.184	7.024	3.330
SC	1.176	1.263	1.209	3.287	6.697	3.399

**Table 1.** Convergence of the FCM solution to the exact solution in terms of the cluster size,  $n$ .

$c$	$k^*/k_0$			$\mu_1^*/\mu_0$			$\mu_2^*/\mu_0$		
	SC	FCM	C&B	SC	FCM	C&B	SC	FCM	C&B
0.10	0.774	0.774	0.774	0.841	0.839	0.841	0.812	0.814	0.812
0.20	0.602	0.601	0.604	0.718	0.712	0.719	0.640	0.644	0.641
0.30	0.465	0.462	0.471	0.609	0.600	0.612	0.490	0.494	0.496
0.40	0.348	0.343	0.364	0.504	0.495	0.512	0.360	0.360	0.379
0.45	0.295	0.289	0.318	0.450	0.443	0.463	0.301	0.298	0.330
0.50	0.242	0.238	0.276	0.393	0.389	0.413	0.243	0.237	0.288

**Table 2.** Comparison of normalized effective moduli of porous solid.

solutions. In addition, these tables contain the results predicted by the approximate formulas proposed for the cubic arrays by Cohen and Bergman [2003] and Cohen [2004], labeled C&B.

These formulas have the following forms [Cohen and Bergman 2003; Cohen 2004]:

$$\frac{k^*}{k_0} = 1 - \frac{c}{k_0/(k_0 - k_1) - 3k_0(1 - c)/(3k_0 + 4\mu_0)}, \quad (45)$$

$$\frac{\mu_1^*}{\mu_0} = 1 - \frac{c(1 - \mu_1/\mu_0)}{1 - (1 - c + G_1)s_2}, \quad (46)$$

$$\frac{\mu_2^*}{\mu_0} = 1 - \frac{c(1 - \mu_1/\mu_0)}{1 - (1 - c + G_2)s_2}, \quad (47)$$

where

$$G_1 = (3k_0 + \mu_0)(-0.929c + 1.1422c^{5/3})/(k_0 + 2\mu_0), \quad G_2 = -2G_1/3, \quad (48)$$

$$s_2 = 1.2(k_0 + 2\mu_0)(1 - \mu_1/\mu_0)/(3k_0 + 4\mu_0).$$

Expression (45) coincides with that of (36) and (42), which means that it does not include the effects of particles' interactions. In the case of porous materials Cohen [2004, Tables 2 and 3] showed that estimates (45)–(47) are in good agreement with those of [Iwakuma and Nemat-Nasser 1983], while the results of [Kushch 1987; Sangani and Lu 1987] for the bulk modulus are in good agreement with those of [Torquato 1998].

<i>c</i>	$k^*/k_0$			$\mu_1^*/\mu_0$			$\mu_2^*/\mu_0$		
	SC	FCM	C&B	SC	FCM	C&B	SC	FCM	C&B
0.10	1.150	1.150	1.150	1.208	1.204	1.208	1.173	1.175	1.173
0.20	1.331	1.331	1.330	1.504	1.490	1.500	1.363	1.369	1.361
0.30	1.558	1.561	1.551	1.915	1.890	1.890	1.593	1.604	1.581
0.40	1.869	1.872	1.829	2.492	2.467	2.375	1.907	1.919	1.856
0.45	2.087	2.083	1.996	2.873	2.865	2.646	2.130	2.137	2.024
0.50	2.391	2.378	2.189	3.348	3.381	2.927	2.463	2.446	2.222

**Table 3.** Comparison of normalized effective moduli of moderate contrast ( $\mu_1/\mu_0 = 10$ ) composite.

<i>c</i>	$k^*/k_0$			$\mu_1^*/\mu_0$			$\mu_2^*/\mu_0$		
	SC	FCM	C&B	SC	FCM	C&B	SC	FCM	C&B
0.10	1.176	1.176	1.176	1.263	1.258	1.263	1.209	1.212	1.209
0.20	1.397	1.399	1.396	1.676	1.653	1.669	1.445	1.455	1.442
0.30	1.690	1.694	1.677	2.336	2.287	2.270	1.746	1.765	1.723
0.40	2.137	2.139	2.048	3.505	3.449	3.106	2.208	2.230	2.092
0.45	2.514	2.508	2.284	4.556	4.570	3.610	2.599	2.609	2.330
0.50	3.287	3.184	2.564	6.697	7.024	4.156	3.399	3.330	2.623

**Table 4.** Comparison of normalized effective moduli of high-contrast ( $\mu_1/\mu_0 = 100$ ) composite.

It can be seen from Tables 2–4 that the generalized Maxwell approach provides estimates of the effective stiffnesses of composites consistent with SC results for the considered range of *c* and  $\mu_1/\mu_0$ . On the other hand, the approximate expressions of [Cohen and Bergman 2003; Cohen 2004], in general, provide accurate estimates only for low volume fractions of particles for high-contrast composites, as suggested by the authors themselves.

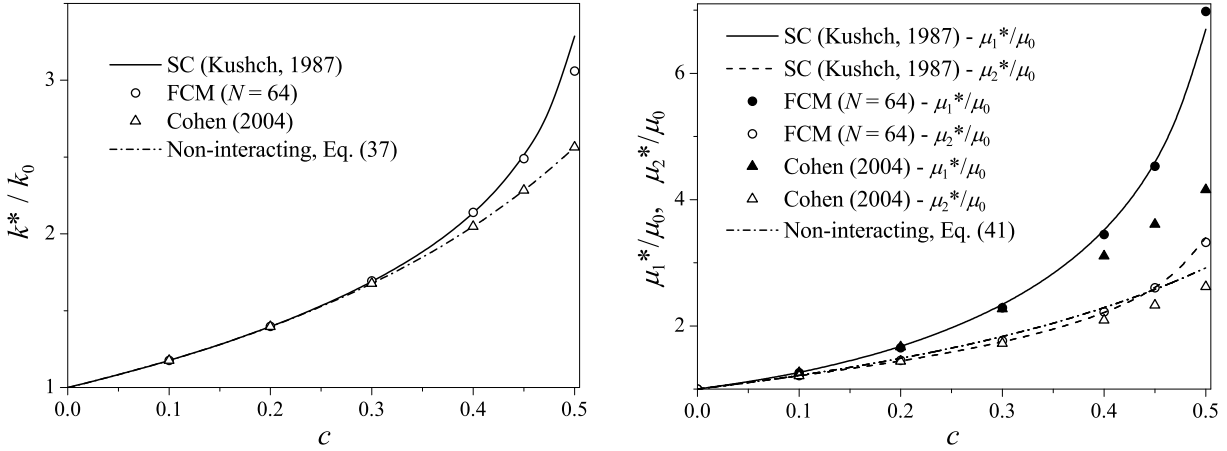
The normalized effective moduli  $k^*(c)/k_0$ ,  $\mu_1^*(c)/\mu_0$ , and  $\mu_2^*(c)/\mu_0$  of the high-contrast composite ( $\mu_1/\mu_0 = 100$ ) are also plotted in Figure 2, in which the FCM-based results ( $n = 4$ ) are marked by open and solid circles, the dash-dotted lines represent the noninteracting approach, the solid and dashed lines represent the SC solution of [Kushch 1987], and the open and solid triangles represent the results of [Cohen 2004].

In addition, in Table 5, we verified that the effective shear moduli  $\mu_1^*(c)$  and  $\mu_2^*(c)$  of high-contrast composite ( $\mu_1/\mu_0 = 100$ ,  $\nu_0 = \nu_1 = 0.3$ ) obtained with both SC and FCM models satisfy the following inequalities developed for a composite with cubic symmetry [Milton 2002]:

$$\frac{3c}{2[\mu_1^*(c) - \mu_0]} + \frac{c}{\mu_2^*(c) - \mu_0} \leq \frac{5}{2(\mu_1 - \mu_0)} + \frac{3(1 - c)(k_0 + 2\mu_0)}{\mu_0(3k_0 + 4\mu_0)}, \tag{49}$$

$$\frac{3(1 - c)}{2[\mu_1 - \mu_1^*(c)]} + \frac{(1 - c)}{\mu_1 - \mu_2^*(c)} \leq \frac{5}{2(\mu_1 - \mu_0)} - \frac{3c(k_1 + 2\mu_1)}{\mu_1(3k_1 + 4\mu_1)}. \tag{50}$$

As can be seen from Table 5, the noninteracting estimates coincide with the right-hand side of inequality (49), while they also fulfill inequality (50). The estimates based on Cohen and Bergman’s formulas



**Figure 2.** Normalized effective bulk modulus (left) and shear moduli (right) of high-contrast ( $\mu_1/\mu_0 = 100$ ) composite with SC array of spherical particles.

$c$	LHS (49)				RHS (49)	LHS (50)				
	(43)	C&B	FCM	SC		(43)	C&B	FCM	SC	RHS (50)
0.10	1.0967	1.048	1.053	1.049	1.0967	0.023	0.023	0.023	0.023	0.0241
0.20	0.9776	0.901	0.899	0.893	0.9776	0.020	0.020	0.020	0.020	0.0229
0.30	0.8586	0.769	0.742	0.739	0.8586	0.018	0.018	0.018	0.018	0.0217
0.40	0.7395	0.651	0.570	0.571	0.7395	0.015	0.015	0.015	0.015	0.0205
0.45	0.6800	0.597	0.472	0.471	0.6800	0.014	0.014	0.014	0.014	0.0199
0.50	0.6205	0.546	0.341	0.340	0.6205	0.013	0.013	0.013	0.013	0.0193

**Table 5.** The values of the left and right-hand sides (LHS and RHS) of (49) and (50) for  $\mu_1/\mu_0 = 100$ .

(46) and (47) also satisfy both inequalities. Interestingly the combination involved in the left-hand side of (50) is the same (up to three significant digits) for all four methods considered.

**5.2. Random composite.** In this section the generalized Maxwell approach is used to evaluate the effective elastic stiffness of random particulate composites. In the case of statistically uniform random microstructure, the composite is known to be macroscopically isotropic and characterized by two elastic moduli,  $k^*$  and  $\mu^*$ .

The representative finite clusters of such material are constructed by random generations of particles in a cube by employing the molecular dynamics algorithm for growing particles used in [Sangani and Mo 1997]. To provide statistical validity to the results, the simulation data are averaged over 20 random configurations. The standard error of the mean (the standard deviation divided by the square root of the number of configurations) is indicated in Tables 6 and 7, with the error estimate for the last significant digit enclosed in parentheses. For example, 4.35(3) means  $4.35 \pm 0.03$ .

In Table 6, the normalized effective bulk ( $k^*/k_0$ ) and shear ( $\mu^*/\mu_0$ ) moduli of the rigid particle composite ( $\mu_1 = \infty$ ) are presented for a wide range of volume fractions:  $c = 0.1, 0.25, 0.45, 0.6$ . In order to check isotropy of the model, the normalized shear modulus was estimated using both (38) and

$c$	$k^*/k_0$			$\mu^*/\mu_0$			
	(36)	S&M	(42)	(38)	(41)	S&M	(43)
0.10	1.184(0)	1.183(1)	1.179	1.245(0)	1.244(0)	1.242(2)	1.233
0.25	1.569(2)	1.60(1)	1.538	1.794(3)	1.775(2)	1.85(2)	1.700
0.45	2.55(1)	2.53(1)	2.322	3.32(2)	3.32(3)	3.42(3)	2.718
0.60	4.22(3)	4.35(3)	3.423	6.5(1)	6.2(2)	6.6(1)	4.15

**Table 6.** Normalized effective moduli of random composite with rigid particles.

$c$	$k^*/k_0$			$\mu^*/\mu_0$			
	(36)	S&M	(42)	(38)	(41)	S&M	(43)
0.10	0.771(0)	0.771(1)	0.774	0.821(0)	0.821(0)	0.820(1)	0.825
0.25	0.522(0)	0.520(1)	0.533	0.599(0)	0.596(1)	0.588(3)	0.611
0.45	0.294(1)	0.299(1)	0.318	0.364(1)	0.351(3)	0.355(2)	0.390
0.60	0.168(1)	0.177(1)	0.202	0.219(1)	0.202(4)	0.205(1)	0.259

**Table 7.** Normalized effective moduli of random porous solid.

(41). Also, Table 6 contains the data obtained by the accurate approach of [Sangani and Mo 1997], labeled S&M, and that obtained by noninteracting method of [McCartney 2010], labeled with (42) and (43).

In order to provide macroscopic isotropy of the composite, the simulations in [Sangani and Mo 1997] have been conducted with  $N = 32$  for  $c = 0.6$  and with  $N = 16$  for other volume fractions. In our computations, as before,  $N = 64$  and  $t_{\max} = 13$ . As seen from Table 6, the modified Maxwell approach provides good estimates for a whole range of  $c$ : for  $c \leq 0.45$ , the difference between the estimates for  $k^*/k_0$  obtained with the generalized Maxwell approach and those reported in [Sangani and Mo 1997] does not exceed the statistical error margins. Note that the configurations with  $c > 0.49$  strongly depend on the method used in generating the random microstructure and on the value of  $N$  (since, at such high  $c$ , the hard-sphere system may be in a metastable fluid state, a semicrystalline state, or a disordered glassy state; for example, [Rintoul and Torquato 1996; Sangani and Mo 1997; Sierou and Brady 2001]). However, even for  $c = 0.6$  the relative error in the estimates for  $k^*/k_0$  obtained from (36) is about 4%. At the same time, underestimation of the effective stiffness by the noninteracting Maxwell method is of the order of 22% for the normalized bulk modulus and 38% for the normalized shear modulus. A minor anisotropy (within 5%) predicted by (38) and (41) can be due to the cubic shape of finite cluster we used and/or the specific  $N (= 4^3 = 64)$  where the semicrystalline arrangement of particles in the generated configuration is likely. This issue deserves separate consideration.

The analogous data for the porous solid are collected in Table 7. Here, we observe the same tendencies and patterns as in the previous case, with the only difference being that this time the standard (noninteracting) Maxwell procedure overestimates  $k^*/k_0$  and  $\mu^*/\mu_0$  for  $c = 0.6$  by 15% and 26%, respectively. The use of the generalized Maxwell procedure reduces the error for  $k^*/k_0$  to 5% and for  $\mu^*/\mu_0$  to 1–7%. For  $c < 0.6$ , an accuracy of the suggested method is even higher, see Tables 6 and 7.

## 6. Conclusions

In this paper the generalized Maxwell approach based on the concept of equivalent anisotropic inhomogeneity is applied for evaluating the anisotropic overall elastic properties of particulate composites with isotropic phases (matrix and spherical particles). A numerical procedure using the approach is outlined for materials with an arbitrary degree of overall anisotropy. This procedure accurately accounts for the geometrical arrangement of particles and their interactions. In a special case of the periodic composite material with simple cubic packing of spherical particles, the closed-form expressions for the three elastic moduli that characterize the overall behavior are provided in terms of dipole coefficients for individual particles. It is demonstrated that noninteracting estimates cannot capture the overall, microstructure-induced anisotropy of the composite materials considered in this work. Illustrative examples involving simple cubic and random composites demonstrate that the approach provides estimates that are consistent with those predicted by the triple-periodic model for the whole range of volume fractions. Based on the results of this work, as well as on the two-dimensional results of [Mogilevskaya and Crouch 2013], it is clear that the approach can be used for the three-dimensional analysis of materials reinforced with particles of arbitrary shapes, if the cluster problem (Section 3) is solved with the boundary element method.

There are several interesting, problem-related issues which are left unresolved/unaddressed in this paper. The first problem is related to the choice of the spherical shape of the equivalent inhomogeneity, which might not adequately represent the shape of the cluster. While this is likely to affect the values of the effective moduli, the reasonable agreement of our results with exact solution for the periodic composite indicates that this effect is rather small. In addition, the recent studies by Sevostianov and Giraud [2012], who investigated the effect of the inhomogeneity's shape on compliance, found that a sphere and a cube of equal volumes possess quite similar compliance tensors. Therefore, it is expected that the dipole moments (apparently expressed in terms of the compliance contribution tensor) would also be similar for the sphere and cube of equal volume. From that point of view, our assumption of spherical shape of the equivalent inhomogeneity should be considered as an approximation which can be refined by taking a more appropriate shape of the cluster or equivalent inhomogeneity. The convergence study in terms of cluster size is a separate problem which also deserves much more attention. In the context of this paper, the most important finding is that macroscopic elastic anisotropy of a composite can be predicted with reasonable accuracy from a relatively small fragment/cluster of the composite structure. In the present work we assumed isotropic constituents of the composite, which means that the macroscopic anisotropy is entirely due to the geometric arrangement of particles. Similar methodology can also be used for composites with anisotropic particles (for example, cubic or transverse isotropic). All of these issues will be addressed in subsequent publications.

## Appendix

The vector surface spherical harmonics  $\mathbf{S}_{ts}^{(i)} = \mathbf{S}_{ts}^{(i)}(\mathbf{r})$  (for example, [Morse and Feshbach 1953]) are defined in terms of their scalar counterparts,  $\chi_t^s = P_t^s(\cos \theta) \exp(is\varphi)$ , as

$$\begin{aligned}\mathbf{S}_{ts}^{(1)} &= r \nabla(\chi_t^s) = \mathbf{e}_\theta \frac{\partial}{\partial \theta} \chi_t^s + \frac{\mathbf{e}_\varphi}{\sin \theta} \frac{\partial}{\partial \varphi} \chi_t^s, \\ \mathbf{S}_{ts}^{(2)} &= r \nabla \times (\mathbf{e}_r \chi_t^s) = \frac{\mathbf{e}_\theta}{\sin \theta} \frac{\partial}{\partial \varphi} \chi_t^s - \mathbf{e}_\varphi \frac{\partial}{\partial \theta} \chi_t^s,\end{aligned}\tag{A.1}$$

$$\mathbf{S}_{ts}^{(3)} = \mathbf{e}_r \chi_t^s \quad (t \geq 0, |s| \leq t).$$

These functions constitute a complete and orthogonal set on the sphere  $S$ . Specifically,

$$\frac{1}{S} \int_S \mathbf{S}_{ts}^{(i)} \cdot \overline{\mathbf{S}_{kl}^{(j)}} dS = \alpha_{ts}^{(i)} \delta_{tk} \delta_{sl} \delta_{ij}, \quad (\text{A.2})$$

where  $\alpha_{ts}^{(1)} = \alpha_{ts}^{(2)} = t(t+1)\alpha_{ts}$  and  $\alpha_{ts}^{(3)} = \alpha_{ts} = \frac{1}{2t+1} \frac{(t+s)!}{(t-s)!}$ . The vector surface spherical harmonics satisfy the following useful relations:

$$\mathbf{S}_{t,-s}^{(i)} = (-1)^{s+i-1} \overline{\mathbf{S}_{ts}^{(i)}}, \quad \mathbf{S}_{-(t+1),s}^{(i)} = \mathbf{S}_{ts}^{(i)}. \quad (\text{A.3})$$

The complete set of the partial solutions of Lamé's equation

$$\frac{2(1-\nu)}{(1-2\nu)} \nabla(\nabla \cdot \mathbf{u}) - \nabla \times \nabla \times \mathbf{u} = 0, \quad (\text{A.4})$$

where  $\mathbf{u}$  is the displacement vector and  $\nu$  is the Poisson ratio, have been introduced in [Kushch 1985]. The regular (bounded everywhere but  $r \rightarrow \infty$ ) complex-value vector functions  $\mathbf{u}_{ts}^{(i)}$  are written in terms of the vector spherical harmonics  $\mathbf{S}_{ts}^{(i)}$  of (A.1) as

$$\begin{aligned} \mathbf{u}_{ts}^{(1)} &= \frac{r^{t-1}}{(t+s)!} (\mathbf{S}_{ts}^{(1)} + t\mathbf{S}_{ts}^{(3)}), & \mathbf{u}_{ts}^{(2)} &= -\frac{1}{(t+1)} \frac{r^t}{(t+s)!} \mathbf{S}_{ts}^{(2)}, \\ \mathbf{u}_{ts}^{(3)} &= \frac{r^{t+1}}{(t+s)!} [\beta_t(\nu)\mathbf{S}_{ts}^{(1)} + \gamma_t(\nu)\mathbf{S}_{ts}^{(3)}], \end{aligned} \quad (\text{A.5})$$

where the coefficients

$$\beta_t(\nu) = \frac{t+5-4\nu}{(t+1)(2t+3)} \quad \text{and} \quad \gamma_t(\nu) = \frac{t-2+4\nu}{(2t+3)} \quad (\text{A.6})$$

are related by  $\gamma_t + (t+1)\beta_t \equiv 1$ . The irregular (infinitely growing at  $r \rightarrow 0$  and vanishing at infinity) complex-value functions  $\mathbf{U}_{ts}^{(i)}$  are

$$\begin{aligned} \mathbf{U}_{ts}^{(1)} &= \frac{(t-s)!}{r^{t+2}} [\mathbf{S}_{ts}^{(1)} - (t+1)\mathbf{S}_{ts}^{(3)}], & \mathbf{U}_{ts}^{(2)} &= \frac{1}{t} \frac{(t-s)!}{r^{t+1}} \mathbf{S}_{ts}^{(2)}, \\ \mathbf{U}_{ts}^{(3)} &= \frac{(t-s)!}{r^t} [\beta_{-(t+1)}(\nu)\mathbf{S}_{ts}^{(1)} + \gamma_{-(t+1)}(\nu)\mathbf{S}_{ts}^{(3)}]. \end{aligned} \quad (\text{A.7})$$

The traction vector  $\mathbf{T}_n = \boldsymbol{\sigma} \cdot \mathbf{n}$  at the surface  $S : r = \text{constant}$  is

$$\frac{1}{2\mu} \mathbf{T}_r(\mathbf{u}) = \frac{\nu}{1-2\nu} \mathbf{e}_r (\nabla \cdot \mathbf{u}) + \frac{\partial}{\partial r} \mathbf{u} + \frac{1}{2} \mathbf{e}_r \times (\nabla \times \mathbf{u}). \quad (\text{A.8})$$

For the regular vector functions of (A.5), this results in [Kushch 1985]

$$\begin{aligned} \frac{1}{2\mu} \mathbf{T}_r(\mathbf{u}_{ts}^{(1)}) &= \frac{(t-1)}{r} \mathbf{u}_{ts}^{(1)}, & \frac{1}{2\mu} \mathbf{T}_r(\mathbf{u}_{ts}^{(2)}) &= \frac{(t-1)}{2r} \mathbf{u}_{ts}^{(2)}, \\ \frac{1}{2\mu} \mathbf{T}_r(\mathbf{u}_{ts}^{(3)}) &= \frac{r^t}{(t+s)!} [b_t(\nu)\mathbf{S}_{ts}^{(1)} + g_t(\nu)\mathbf{S}_{ts}^{(3)}], \end{aligned} \quad (\text{A.9})$$

where

$$b_t(v) = (t+1)\beta_t - 2(1-v)/(t+1), \quad g_t(v) = (t+1)\gamma_t - 2v. \quad (\text{A.10})$$

For the irregular solutions  $\mathbf{U}_{ts}^{(i)}$  of (A.7), the vector takes the following form:

$$\begin{aligned} \frac{1}{2\mu} \mathbf{T}_r(\mathbf{U}_{ts}^{(1)}) &= -\frac{(t+2)}{r} \mathbf{U}_{ts}^{(1)}, & \frac{1}{2\mu} \mathbf{T}_r(\mathbf{u}_{ts}^{(2)}) &= -\frac{(t+2)}{2r} \mathbf{U}_{ts}^{(2)}, \\ \frac{1}{2\mu} \mathbf{T}_r(\mathbf{U}_{ts}^{(3)}) &= \frac{(t-s)!}{r^{t+1}} [b_{-(t+1)} \mathbf{S}_{ts}^{(1)} + g_{-(t+1)} \mathbf{S}_{ts}^{(3)}] \end{aligned} \quad (\text{A.11})$$

In view of (A.5),  $\mathbf{T}_r(\mathbf{u}_{ts}^{(i)})$  can be represented in terms of vector spherical harmonics of (A.1).

The resultant force  $\mathbf{T}$  and resultant torque (moment)  $\mathbf{M}$  acting on the spherical surface  $S : r = R$  enclosing the point  $\mathbf{r} = 0$  are

$$\mathbf{T} = \int_S \mathbf{T}_r dS, \quad \mathbf{M} = \int_S \mathbf{r} \times \mathbf{T}_r dS. \quad (\text{A.12})$$

It is readily found that  $\mathbf{T} = \mathbf{M} = 0$  for all the regular functions  $\mathbf{u}_{ts}^{(i)}$ . Among the irregular functions  $\mathbf{U}_{ts}^{(i)}$ , only three functions have nonzero resultant force, namely

$$\mathbf{T}(\mathbf{U}_{10}^{(3)}) = 16\mu\pi(v-1)\mathbf{i}_3, \quad \mathbf{T}(\mathbf{U}_{11}^{(3)}) = -\overline{\mathbf{T}(\mathbf{U}_{1,-1}^{(3)})} = 32\mu\pi(v-1)(\mathbf{i}_1 + \mathbf{i}_2). \quad (\text{A.13})$$

Hence,  $\mathbf{U}_{1s}^{(3)}$  can be regarded as vector monopoles. The net resultant torque is zero for all the Lamé solutions except for  $\mathbf{U}_{1s}^{(2)}$ , for which

$$\mathbf{M}(\mathbf{U}_{10}^{(2)}) = -8\mu\pi\mathbf{i}_3, \quad \mathbf{M}(\mathbf{U}_{11}^{(2)}) = \overline{\mathbf{M}(\mathbf{U}_{1,-1}^{(2)})} = -16\mu\pi(\mathbf{i}_1 + \mathbf{i}_2). \quad (\text{A.14})$$

Formulas (A.13) and (A.14) provide an insight into the physical meaning of the irregular vector functions  $\mathbf{U}_{1s}^{(2)}$  and  $\mathbf{U}_{1s}^{(3)}$ , these being the displacements due to concentrated moment and force, respectively, applied at the point  $\mathbf{r} = 0$ .

## References

- [Benveniste 1987] Y. Benveniste, "A new approach to the application of Mori–Tanaka's theory in composite materials", *Mech. Mater.* **6** (1987), 147–157.
- [Berriman and Berge 1996] J. Berriman and P. Berge, "Critique of two explicit schemes for estimating elastic properties of multiphase composites", *Mech. Mater.* **22** (1996), 149–164.
- [Cohen 2004] I. Cohen, "Simple algebraic approximations for the effective elastic moduli of cubic arrays of spheres", *J. Mech. Phys. Solids* **52**:9 (2004), 2167–2183.
- [Cohen and Bergman 2003] I. Cohen and D. J. Bergman, "Clausius–Mossotti-type approximations for elastic moduli of a cubic array of spheres", *Phys. Rev. B* **68**:2 (2003), Article ID #24104.
- [Eshelby 1961] J. D. Eshelby, "Elastic inclusions and inhomogeneities", pp. 87–140 in *Progress in solid mechanics*, vol. 2, edited by I. N. Sneddon and R. Hill, North-Holland, Amsterdam, 1961.
- [Fu et al. 1998] Y. Fu, K. J. Klimkowski, G. J. Rodin, E. Berger, J. C. Browne, J. K. Singer, R. A. Van De Geijn, and K. S. Vemaganti, "A fast solution method for three-dimensional many-particle problems of linear elasticity", *Int. J. Numer. Methods Eng.* **42**:7 (1998), 1215–1229.
- [Golovchan et al. 1993] V. T. Golovchan, A. N. Guz, Y. V. Kohanenko, and V. I. Kushch, *Mechanics of composites*, vol. VI, Naukova Dumka, Kiev, 1993.



- [Grzhibovskis et al. 2010] R. Grzhibovskis, S. Rjasanow, H. Andrä, and A. Zemitis, “Boundary element method for calculation of effective elastic moduli in 3D linear elasticity”, *Math. Methods Appl. Sci.* **33**:8 (2010), 1021–1034.
- [Hashin and Shtrikman 1963] Z. Hashin and S. Shtrikman, “A variational approach to the theory of the elastic behaviour of multiphase materials”, *J. Mech. Phys. Solids* **11** (1963), 127–140.
- [Huang 1968] Y. C. Huang, “The effect of a spherical inclusion in anisotropic solid”, *Appl. Sci. Res.* **18** (1968), 436–445.
- [Iwakuma and Nemat-Nasser 1983] T. Iwakuma and S. Nemat-Nasser, “Composites with periodic microstructure”, *Comput. Struct.* **16** (1983), 13–19.
- [Ju and Yanase 2010] J. W. Ju and K. Yanase, “Micromechanics and effective elastic moduli of particle-reinforced composites with near-field particle interactions”, *Acta Mech.* **215** (2010), 135–153.
- [Kerner 1956] E. H. Kerner, “The elastic and thermo-elastic properties of composite media”, *Proc. Phys. Soc. B* **69** (1956), 808–813.
- [Koroteeva et al. 2010] O. Koroteeva, S. G. Mogilevskaya, E. Gordeliy, and S. L. Crouch, “A computational technique for evaluating the effective thermal conductivity of isotropic porous materials”, *Eng. Anal. Bound. Elem.* **34** (2010), 793–801.
- [Kushch 1985] V. I. Kushch, “Elastic equilibrium of a medium containing periodic spherical inclusions”, *Prikl. Mekh.* **21**:5 (1985), 18–27. In Russian; translated in *Sov. Appl. Mech.* **21**:5 (1985), 435–442.
- [Kushch 1987] V. I. Kushch, “Computation of the effective elastic moduli of a granular composite material of regular structure”, *Prikl. Mekh.* **23**:4 (1987), 57–61. In Russian; translated in *Sov. Appl. Mech.* **23**:4 (1987), 362–365.
- [Kushch 1996] V. I. Kushch, “Elastic equilibrium of a medium containing finite number of aligned spheroidal inclusions”, *Int. J. Solids Struct.* **33** (1996), 1175–1189.
- [Kushch 1997] V. I. Kushch, “Microstresses and effective elastic moduli of a solid reinforced by periodically distributed spheroidal inclusions”, *Int. J. Solids Struct.* **34** (1997), 1353–1366.
- [Kushch 1998] V. I. Kushch, “Elastic equilibrium of a solid containing a finite number of arbitrarily oriented spheroidal inclusions”, *Int. J. Solids Struct.* **35** (1998), 1751–1762.
- [Kushch 2003] V. I. Kushch, “Stress concentrations in the particulate composite with transversely isotropic phases”, *Int. J. Solids Struct.* **40** (2003), 6369–6388.
- [Kushch 2013] V. I. Kushch, *Micromechanics of composites: multipole expansion approach*, Elsevier, Amsterdam, 2013.
- [Kushch and Sevostianov 2004] V. I. Kushch and I. Sevostianov, “Effective elastic properties of the particulate composite with transversely isotropic phases”, *Int. J. Solids Struct.* **41** (2004), 885–906.
- [Kushch et al. 2011] V. I. Kushch, S. G. Mogilevskaya, H. K. Stolarski, and S. L. Crouch, “Elastic interaction of spherical nanoinhomogeneities with Gurtin–Murdoch type interfaces”, *J. Mech. Phys. Solids* **59**:9 (2011), 1702–1716.
- [Kuster and Toksöz 1974] G. T. Kuster and M. N. Toksöz, “Velocity and attenuation of seismic waves in two-phase media, I. Theoretical formulations”, *Geophys.* **39** (1974), 587–606.
- [Levin et al. 2012] V. Levin, S. Kanaun, and M. Markov, “Generalized Maxwell’s scheme for homogenization of poroelastic composites”, *Int. J. Eng. Sci.* **61** (2012), 75–86.
- [Luciano and Barbero 1994] R. Luciano and E. J. Barbero, “Formulas for the stiffness of composites with periodic microstructure”, *Int. J. Solids Struct.* **31** (1994), 2933–2944.
- [Maxwell 1873] J. C. Maxwell, *A treatise on electricity and magnetism*, vol. 1, Clarendon, Oxford, 1873. 3rd ed. in 1892.
- [McCartney 2010] L. N. McCartney, “Maxwell’s far-field methodology predicting elastic properties of multi-phase composites reinforced with aligned transversely isotropic spheroids”, *Philos. Mag.* **90** (2010), 4175–4207.
- [McCartney and Kelly 2008] L. N. McCartney and A. Kelly, “Maxwell’s far-field methodology applied to the prediction of properties of multi-phase isotropic particulate composites”, *Proc. R. Soc. Lond. A* **464**:2090 (2008), 423–446.
- [Milton 2002] G. W. Milton, *The theory of composites*, Cambridge Monographs on Applied and Computational Mathematics **6**, Cambridge University Press, 2002.
- [Mogilevskaya and Crouch 2013] S. G. Mogilevskaya and S. L. Crouch, “Combining Maxwell’s methodology with the BEM for evaluating the two-dimensional effective properties of composite and micro-cracked materials”, *Comput. Mech.* **51** (2013), 377–389.

- [Mogilevskaya et al. 2010a] S. G. Mogilevskaya, S. L. Crouch, A. La Grotta, and H. K. Stolarski, “The effects of surface elasticity and surface tension on the trasverse overall elastic behavior of unidirectional nano-composites”, *Compos. Sci. Technol.* **70** (2010), 427–434.
- [Mogilevskaya et al. 2010b] S. G. Mogilevskaya, S. L. Crouch, H. K. Stolarski, and A. Benusiglio, “Equivalent inhomogeneity method for evaluating the effective elastic properties of unidirectional multi-phase composites with surface/interface effects”, *Int. J. Solids Struct.* **47** (2010), 407–418.
- [Mogilevskaya et al. 2011] S. G. Mogilevskaya, V. I. Kushch, O. Koroteeva, and S. L. Crouch, “Equivalent inhomogeneity method for evaluating the effective conductivities of isotropic particulate composites”, *J. Mech. Mater. Struct.* **7** (2011), 103–117.
- [Mogilevskaya et al. 2012] S. G. Mogilevskaya, H. K. Stolarski, and S. L. Crouch, “On Maxwell’s concept of equivalent inhomogeneity: when do the interactions matter?”, *J. Mech. Phys. Solids* **60**:3 (2012), 391–417.
- [Molinari and El Mouden 1996] A. Molinari and M. El Mouden, “The problem of elastic inclusions at finite concentration”, *Int. J. Solids Struct.* **33** (1996), 3131–3150.
- [Morse and Feshbach 1953] P. M. Morse and H. Feshbach, *Methods of theoretical physics*, McGraw-Hill, New York, 1953.
- [Moschovidis and Mura 1975] Z. A. Moschovidis and T. Mura, “Two ellipsoidal inhomogeneities by the equivalent inclusion method”, *J. Appl. Mech. (ASME)* **42** (1975), 847–852.
- [Mura 1987] T. Mura, *Micromechanics of defects in solids*, Mechanics of Elastic and Inelastic Solids **3**, Martinus Nijhoff, Dordrecht, 1987.
- [Nemat-Nasser and Taya 1981] S. Nemat-Nasser and M. Taya, “On effective moduli of an elastic body containing periodically distributed voids”, *Quart. Appl. Math.* **39** (1981), 49–59.
- [Nemat-Nasser et al. 1982] S. Nemat-Nasser, T. Iwakuma, and M. Hejazi, “On composites with periodic structure”, *Mech. Mater.* **1** (1982), 239–267.
- [Nunan and Keller 1984] K. C. Nunan and J. B. Keller, “Effective elasticity tensor of a periodic composite”, *J. Mech. Phys. Solids* **32** (1984), 259–280.
- [Ponte Castañeda and Willis 1995] P. Ponte Castañeda and J. R. Willis, “The effect of spatial distribution on the effective behavior of composite materials and cracked media”, *J. Mech. Phys. Solids* **43**:12 (1995), 1919–1951.
- [Pyatigorets and Mogilevskaya 2011] A. V. Pyatigorets and S. G. Mogilevskaya, “Evaluation of effective transverse mechanical properties of transversely isotropic viscoelastic composite materials”, *J. Compos. Mater.* **45** (2011), 2641–2658.
- [Rintoul and Torquato 1996] M. D. Rintoul and S. Torquato, “Computer simulations of dense hard-sphere systems”, *J. Chem. Phys.* **105** (1996), 9258–9265.
- [Rodin and Hwang 1991] G. J. Rodin and Y. L. Hwang, “On the problem of linear elasticity for an infinite region containing a finite number of non-intersecting spherical inhomogeneities”, *Int. J. Solids Struct.* **27** (1991), 145–159.
- [Sangani and Lu 1987] A. S. Sangani and W. Lu, “Elastic coefficients of composites containing spherical inclusions in a periodic array”, *J. Mech. Phys. Solids* **35** (1987), 1–21.
- [Sangani and Mo 1997] A. S. Sangani and G. Mo, “Elastic interactions in particulate composites with perfect as well as imperfect interfaces”, *J. Mech. Phys. Solids* **45**:11-12 (1997), 2001–2031.
- [Segurado and Llorca 2002] J. Segurado and J. Llorca, “A numerical approximation to the elastic properties of sphere-reinforced composites”, *J. Mech. Phys. Solids* **50** (2002), 2107–2121.
- [Segurado and Llorca 2006] J. Segurado and J. Llorca, “Computational micromechanics of composites: the effect of particle spatial distribution”, *Mech. Mater.* **38** (2006), 873–883.
- [Sevostianov and Giraud 2012] I. Sevostianov and A. Giraud, “On the compliance contribution tensor for a concave superspherical pore”, *Int. J. Fract.* **177** (2012), 199–206.
- [Shen and Yi 2001] L. Shen and S. Yi, “An effective inclusion model for effective moduli of heterogeneous materials with ellipsoidal inhomogeneities”, *Int. J. Solids Struct.* **38** (2001), 5789–5805.
- [Sierou and Brady 2001] A. Sierou and J. F. Brady, “Accelerated Stokesian dynamics simulations”, *J. Fluid Mech.* **448** (2001), 115–146.

- [Torquato 1997] S. Torquato, “Effective stiffness tensor of composite media, I: Exact series expansions”, *J. Mech. Phys. Solids* **45**:9 (1997), 1421–1448.
- [Torquato 1998] S. Torquato, “Effective stiffness tensor of composite media, II: Applications to isotropic dispersions”, *J. Mech. Phys. Solids* **46**:8 (1998), 1411–1440.
- [Torquato 2002] S. Torquato, *Random heterogeneous materials: microstructure and macroscopic properties*, Interdisciplinary Applied Mathematics **16**, Springer, New York, 2002.
- [Weng 2010] G. J. Weng, “A dynamical theory for the Mori–Tanaka and Ponte Castañeda–Willis estimates”, *Mech. Mater.* **42**:9 (2010), 886–893.
- [Willis 1977] J. R. Willis, “Bounds and self-consistent estimates for the overall properties of anisotropic composites”, *J. Mech. Phys. Solids* **25** (1977), 185–202.
- [Yin and Sun 2005] H. M. Yin and L. Z. Sun, “Elastic modelling of periodic composites with particle interactions”, *Philos. Mag. Lett.* **85** (2005), 163–173.
- [Zohdi and Wriggers 2005] T. I. Zohdi and P. Wriggers, *An introduction to computational micromechanics*, Lecture Notes in Applied and Computational Mechanics **20**, Springer, Berlin, 2005.

Received 21 Nov 2012. Revised 26 Mar 2013. Accepted 4 Apr 2013.

VOLODYMYR I. KUSHCH: vkushch@bigmir.net

*Institute for Superhard Materials, National Academy of Sciences of Ukraine, 2 Avtozavodskaya St., Kiev 04074, Ukraine*

SOFIA G. MOGILEVSKAYA: mogil003@umn.edu

*Department of Civil Engineering, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, MN 55455, United States*  
<http://www.ce.umn.edu/people/faculty/mogilevs/>

HENRYK K. STOLARSKI: stola001@umn.edu

*Department of Civil Engineering, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, MN 55455, United States*

STEVEN L. CROUCH: crouch@umn.edu

*Department of Civil Engineering, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, MN 55455, United States*  
<http://www.ce.umn.edu/directory/faculty/crouch>



## ON SUCCESSIVE DIFFERENTIATIONS OF THE ROTATION TENSOR: AN APPLICATION TO NONLINEAR BEAM ELEMENTS

TEODORO MERLINI AND MARCO MORANDINI

Successive differentiations of the rotation tensor are characterized by successive differential rotation vectors. Useful expressions of the differential rotation vectors for differentiations up to third order are derived. In the context of the exponential parameterization, explicit expressions for the differential maps (the maps providing the differential rotation vectors from the differentials of the parameters chosen) are obtained by resorting to an original infinite family of recursive subexponential maps. Useful properties of the mapping tensors are discussed.

The formulation is appropriate for nonlinear problems of computational solid mechanics, when spatial, incremental, and virtual variations of particle orientations must be dealt with together. As an application, the classical problem of modeling space-curved slender beams by finite elements is considered. The variational formulation and the nonlinear interpolation of the orientations, together with the relevant linearizations, consistently exploit the proposed differentiations and lead to an objective beam element. Two test cases are discussed.

### 1. Introduction

The motivation for a circumstantial study of the differentiations of the rotation tensor comes from specific demands in computational continuum mechanics by the finite element method and in the relevant variational formulations. As far as three-dimensional solids are concerned, the rotation field is manifestly an unknown variable in nonlinear mechanics of polar materials [Grekova and Zhilin 2001; Bauer et al. 2010], but is introduced as well as an unknown variable in some discrete representations with classical nonpolar materials [Simo et al. 1992; Atluri and Cazzani 1995; Merlini 1997]. Moreover, the rotation field is a primary unknown variable in Cosserat-type formulations of structured solid mechanics, namely beams, that behave as one-dimensional polar continua, and shells (refer to [Altenbach et al. 2010] and references therein), that feature a mixed polar/nonpolar constitutive behavior.

Multiple differentiations of the finite-rotation field are involved in nonlinear continuum mechanics problems. The particle orientations (that is, rotations from an absolute reference frame) within a body in any (deformed) configurations “differ” in general from each other, and the differential (hence the gradient) of the orientation field is used to define angular curvatures within the body (and hence angular strains by comparing curvatures in different configurations). Throughout the body deformation, the particle orientations undergo rotations; differential rotations are the unknowns of an incremental solution process in nonlinear boundary value problems. Following a variational approach with its discrete representation by the finite element method, one handles virtual functionals depending on virtual variable fields, so

---

*Keywords:* finite rotation differentiations, exponential map, rotation differential maps, space-curved slender beams, orientation interpolation, nonlinear beam elements.

virtual (differential) rotations of the particle orientations must be accounted for as well. This outline of the computational strategy lets us see that we have to deal with (at least) three independent variations of the orientation field: a spatial variation within the material body (by itself a complicated task in three dimensions), an incremental variation along with the configuration evolution, and a virtual variation that enables us to cast a discrete equation set.

A consistent setting that allows one to handle successive derivatives of the rotation tensor in a systematic way is lacking in the literature. A number of papers on the derivatives of wider classes of tensor-valued functions of a tensor were published during the past decade. Often, these investigations were motivated in the field of continuum mechanics [Rosati 1999; Itskov 2002; Jog 2008] and in the modeling of nonlinear elastoplastic constitutive laws [de Souza Neto 2001]. Most papers concern the representation and the derivatives of isotropic tensor functions of either a symmetrical or a generally unsymmetrical tensor [Ortiz et al. 2001; de Souza Neto 2001; 2004; Itskov and Aksel 2002; Itskov 2003; Fung 2004; Lu 2004; Dui et al. 2006; Wang and Dui 2007; Jog 2008]. Some of these works deal explicitly with exponential functions, and in this context reference is due to the accurate and worthy paper [Najfeld and Havel 1995], published in a field far from the scope of engineering. Though some cues may come from these works, it seems difficult to bring them in the bed of orthogonal tensors, as the exponential functions of skew-symmetric tensors are.

In the case of orthogonal tensors, explicit formulae for the derivatives of the tangent map of the rotation tensor are available, for example, in [Borri et al. 1990; Ritto-Corrêa and Camotim 2002; Mäkinen 2008], but again a systematic setting of successive differentiations is missing. However, the specific properties of the special orthogonal group should help in obtaining closed-form expressions for any differentiations. This task is undertaken in this paper. Our approach is based on an original decomposition of the exponential map into an infinite family of recursive subexponential maps, whose lowest differentiations are affordable. This enables us to manage successive differentiations of the rotation tensor with analytical expressions that are safe to implement and exact within machine precision. The proposed methodology was implemented in a nonlinear finite element code and successfully tested with solids [Merlini and Morandini 2005] and shells [Merlini and Morandini 2011a]; in [Merlini and Morandini 2004b, Appendices] a brief description of the methodology was given.

The outline of the paper is as follows. In Section 2 the essential structure of successive differentiations of an orthogonal tensor is presented and appropriate differential rotation vectors are proposed, however no parameterization of the rotation tensor is introduced yet. In Section 3 the exponential map is assumed for the rotation tensor and a useful family of subexponential maps is conveniently set. The lowest differentiations of the subexponential maps are dealt with in Section 4 and are used in Section 5 in order to provide expressions for the differential maps of the rotation, that is, the maps from the differentials of the rotation vector to the differential rotation vectors themselves.

The remaining part of the paper deals with an application of the proposed formulation to a classical problem in computational mechanics, the finite-element modeling of space-curved slender beams. The essentials of Reissner–Simo beam variational mechanics are discussed in Section 6 and a beam element based on a consistent nonlinear interpolation is discussed in Section 7. The formulation is contrasted with the analogous paper [Ritto-Corrêa and Camotim 2002] to highlight the significance of adopting a nonlinear element. In Section 8, two popular numerical tests demonstrate the performance of both

two-node and three-node beam elements. As further remarked in Section 9, the formulation is kept parameterization-free for as long as possible.

The style of the mathematical developments is kept as plain as possible. Abstract formalisms, typical of rotation math, are avoided as unnecessary to understanding matters that are meant for people with a mechanical background. When tensor-valued functions are differentiable — as rotations are, being exponential functions of the rotation vector — we by far prefer working with differentials than with directional derivatives. On the other hand, a heavy use of higher-order tensors is unavoidable in this context, and an index-free tensor notation is adopted throughout the paper to make it easier to follow complicated developments. Tensor notations and rules used in the paper are gathered in the Appendix.

## 2. Structure of successive differentiations of an orthogonal tensor

From a merely geometric standpoint, a rotation is a tensor  $\Phi$  that transforms a frame of three vectors (say  $\mathbf{a}$ ,  $\mathbf{b}$ , and  $\mathbf{c}$ ) into another frame ( $\Phi\mathbf{a}$ ,  $\Phi\mathbf{b}$ , and  $\Phi\mathbf{c}$ ), while preserving the vector lengths and their mutual orientations, and hence the frame volume  $\mathbf{a} \times \mathbf{b} \cdot \mathbf{c}$ . To work so, a rotation tensor must obey the symmetrical tensor equation

$$\Phi \Phi^T = I \quad (1)$$

and the scalar condition  $\det \Phi = +1$ . Equation (1) is called the *orthogonality condition* and represents six scalar constraints to fulfill in order to classify a tensor as a rotation; as a consequence, it makes a rotation tensor depend on just three scalar parameters.

**2.1. Differential rotation vectors.** It is well known that the differentiation of the rotation tensor is characterized by a skew-symmetric tensor, hence by the relevant axial vector [Pietraszkiewicz and Badur 1983; Cardona and Gérardin 1988; Ibrahimbegović et al. 1995]. It has been found that any successive independent differentiation of the rotation tensor is in turn characterized by a further vector that depends on the differentiations of the preceding characteristic vectors themselves [Merlini and Morandini 2004a]. In this section we delve into this topic.

Let us start by differentiating the orthogonality condition (1). We denote successive independent variations with  $d_1, d_2, d_3, \dots$  and keep evaluating  $d_1(\Phi \Phi^T), d_2 d_1(\Phi \Phi^T), d_3 d_2 d_1(\Phi \Phi^T), \dots$  in sequence. We use the symbolic notation  $d^n(\cdot)$  for a multiple differential  $d_n \dots d_3 d_2 d_1(\cdot)$ . Evaluation of the  $n$ -th differential  $d^n(\Phi \Phi^T)$  yields a sum of  $2^n$  terms, specifically the extreme tensors, namely  $d^n \Phi \Phi^T$  and its transpose  $\Phi d^n \Phi^T$ , and the bulk of the remaining tensors that contain lower-order differentials (up to  $d^{n-1} \Phi$ ) of the rotation. Let us denote by  $-2\Phi_{d^n}^S$  the sum of this bulk of tensors, which is of course a symmetric tensor (we denote a symmetric tensor by a superscript  $(\cdot)^S$  and a skew-symmetric tensor by means of its axial vector  $\mathbf{a}$  as  $\mathbf{a} \times$ , see (A.1)). Thus, a multiple differentiation of (1) ends up in a form like  $d^n(\Phi \Phi^T) = d^n \Phi \Phi^T - 2\Phi_{d^n}^S + \Phi d^n \Phi^T = d^n \Phi \Phi^T - \Phi_{d^n}^S + (d^n \Phi \Phi^T - \Phi_{d^n}^S)^T = \mathbf{0}$ . This means that tensor  $d^n \Phi \Phi^T - \Phi_{d^n}^S$  is skew-symmetric, say  $\varphi_{d^n} \times$ , and finally we obtain the decomposition of tensor  $d^n \Phi \Phi^T$  into its symmetric and skew-symmetric parts,

$$d^n \Phi \Phi^T = (\varphi_{d^n} \times) + \Phi_{d^n}^S.$$

As it is well known, tensor  $d_1 \Phi \Phi^T = \varphi_{d_1} \times$  is skew-symmetric ( $\Phi_{d_1}^S = \mathbf{0}$ ) and  $\varphi_{d_1}$  represents the characteristic vector of the first differential  $d_1 \Phi$  of the rotation tensor. The successive tensors  $d^n \Phi \Phi^T$

with  $n > 1$  have both symmetric and skew-symmetric parts, instead. The former ones ( $\Phi_{d^n}^S$ ) are algebraic functions of the characteristic vectors  $\varphi_{d^1}, \varphi_{d^2}, \dots, \varphi_{d^{n-1}}$  of the preceding lower-order differentials, whereas the latter ones ( $\varphi_{d^n} \times$ ) contain the differentials of those characteristic vectors as well. Therefore, the axial vector  $\varphi_{d^n}$  can be assumed as a characteristic vector of the  $n$ -th differential  $d^n \Phi$  of the rotation tensor. Note that vectors  $\varphi_{d^n}$  with  $n > 1$  also contain algebraic functions of the preceding characteristic vectors  $\varphi_{d^1}, \varphi_{d^2}, \dots, \varphi_{d^{n-1}}$ , thus there is some arbitrariness in the choice of the characteristic vector. Our choice is to assume the whole axial vector of  $d^n \Phi \Phi^T$  as *the* characteristic vector of the differential  $d^n \Phi$ .

The above characterization of successive differentials of the rotation tensor is now made explicit for differentiations up to third order. We denote three successive independent variations by  $\delta$ ,  $\partial$ , and  $d$ . In computational finite elasticity, such symbols are conveniently associated respectively with virtual variations, incremental variations, and spatial variations — the latter being derivatives along either a one-coordinate domain (beams), or surface gradients on a two-dimensional domain (shells), or even gradients on a three-dimensional solid domain. After evaluation of  $\delta(\Phi \Phi^T) = \mathbf{0}$ ,  $\partial \delta(\Phi \Phi^T) = \mathbf{0}$ , and  $d \partial \delta(\Phi \Phi^T) = \mathbf{0}$ , we easily obtain

$$\begin{aligned} \delta \Phi \Phi^T &= \varphi_\delta \times, \\ \partial \delta \Phi \Phi^T &= \varphi_{\partial \delta} \times + \frac{1}{2}(\varphi_\partial \times \varphi_\delta \times + \varphi_\delta \times \varphi_\partial \times), \\ d \partial \delta \Phi \Phi^T &= \varphi_{d \partial \delta} \times + \frac{1}{2}(\varphi_{d \partial} \times \varphi_\delta \times + \varphi_{\partial \delta} \times \varphi_d \times + \varphi_{\delta d} \times \varphi_\partial \times + \varphi_d \times \varphi_{\partial \delta} \times + \varphi_\partial \times \varphi_{\delta d} \times + \varphi_\delta \times \varphi_{d \partial} \times), \end{aligned} \quad (2)$$

where  $\varphi_\delta$ ,  $\varphi_{\partial \delta}$ , and  $\varphi_{d \partial \delta}$  are proposed here as the characteristic vectors introduced at each successive variation.

It is worth noting that no parameterization of the rotation tensor is implied in the foregoing characterization. In spite of that, the first characteristic vector  $\varphi_\delta$  has been properly referred to as the virtual rotation vector [Borri et al. 1990]. A simple reason for such terminology comes from considering the rotation built with the infinitesimal rotation vector  $\varphi_\delta$ , that is,  $\exp(\varphi_\delta \times) \cong \mathbf{I} + \varphi_\delta \times$ , and appending it to a rotation  $\Phi$ ; the composed rotation becomes  $\exp(\varphi_\delta \times) \Phi \cong (\mathbf{I} + \varphi_\delta \times) \Phi = \Phi + \delta \Phi$  and matches the expression of what is often understood — perhaps improperly — as a varied rotation. However, it is worth stressing that  $\varphi_\delta$  is an (independent) infinitesimal differential rotation vector and by no means has to coincide (in general) with the differential  $\delta \varphi$  of the rotation vector  $\varphi = \text{ax} \ln \Phi$  [Borri et al. 1990]. Note that vector  $\varphi_\delta$  is also related to the concept of spin and is often given a perhaps misleading notation using the differential symbol  $\delta$  applied at a different variable, for example,  $\delta \omega$  in [Ritto-Corrêa and Camotim 2002]. We by far prefer the notation with an appended subscript  $\delta$  to the rotation vector itself, a notation that stresses unequivocally that  $\varphi_\delta$  is not the variation of a vector-valued function.

We extend the same notation convention to vectors  $\varphi_{\partial \delta}$  and  $\varphi_{d \partial \delta}$  and adopt a general terminology for the characteristic vectors of any differentials of the rotation tensor by calling them *differential rotation vectors*. These vectors account for the variations of the differential vectors of lower order, as it is shown next.

**2.2. Evaluation of successive differential rotation vectors.** Expressions for evaluating the second and third differential rotation vectors follow from their definitions as axial vectors, that is,  $\varphi_{\partial \delta} = \text{ax}(\partial \delta \Phi \Phi^T)$  and  $\varphi_{d \partial \delta} = \text{ax}(d \partial \delta \Phi \Phi^T)$ . From

$$\varphi_{\partial \delta} \times = \frac{1}{2}(\partial \delta \Phi \Phi^T - (\partial \delta \Phi \Phi^T)^T) \quad \text{and} \quad \varphi_{d \partial \delta} \times = \frac{1}{2}(d \partial \delta \Phi \Phi^T - (d \partial \delta \Phi \Phi^T)^T)$$



we obtain, after some algebraic manipulations involving (2), the expressions

$$\begin{aligned}\boldsymbol{\varphi}_{\partial\delta} &= \partial\boldsymbol{\varphi}_\delta - \frac{1}{2}\boldsymbol{\varphi}_\partial \times \boldsymbol{\varphi}_\delta, \\ \boldsymbol{\varphi}_{d\partial\delta} &= d\boldsymbol{\varphi}_{\partial\delta} - \frac{1}{2}\boldsymbol{\varphi}_d \times \boldsymbol{\varphi}_{\partial\delta} - \frac{1}{2}(\boldsymbol{\varphi}_d \otimes \boldsymbol{\varphi}_\partial \cdot \boldsymbol{\varphi}_\delta + (\boldsymbol{\varphi}_\partial \otimes \boldsymbol{\varphi}_\delta)^S \cdot \boldsymbol{\varphi}_d) \\ &= d\partial\boldsymbol{\varphi}_\delta - \frac{1}{2}\boldsymbol{\varphi}_{d\partial} \times \boldsymbol{\varphi}_\delta - \frac{1}{2}(\boldsymbol{\varphi}_d \times \boldsymbol{\varphi}_{\partial\delta} + \boldsymbol{\varphi}_\partial \times \boldsymbol{\varphi}_{\delta d}) - (\boldsymbol{\varphi}_d \otimes \boldsymbol{\varphi}_\partial)^S \cdot \boldsymbol{\varphi}_\delta,\end{aligned}\quad (3)$$

where the symbols  $d$ ,  $\partial$ , and  $\delta$  may exchange cyclically. Other expressions, symmetrical with respect to the variations  $d$ ,  $\partial$ , and  $\delta$ , easily follow from (3):

$$\begin{aligned}\boldsymbol{\varphi}_{\partial\delta} &= \frac{1}{2}(\partial\boldsymbol{\varphi}_\delta + \delta\boldsymbol{\varphi}_\partial), \\ \boldsymbol{\varphi}_{d\partial\delta} &= \frac{1}{3}(d\boldsymbol{\varphi}_{\partial\delta} + \partial\boldsymbol{\varphi}_{\delta d} + \delta\boldsymbol{\varphi}_{d\partial}) - \frac{1}{6}(\boldsymbol{\varphi}_d \times \boldsymbol{\varphi}_{\partial\delta} + \boldsymbol{\varphi}_\partial \times \boldsymbol{\varphi}_{\delta d} + \boldsymbol{\varphi}_\delta \times \boldsymbol{\varphi}_{d\partial}) \\ &\quad - \frac{1}{3}(\boldsymbol{\varphi}_d \otimes \boldsymbol{\varphi}_\partial \cdot \boldsymbol{\varphi}_\delta + \boldsymbol{\varphi}_\partial \otimes \boldsymbol{\varphi}_\delta \cdot \boldsymbol{\varphi}_d + \boldsymbol{\varphi}_\delta \otimes \boldsymbol{\varphi}_d \cdot \boldsymbol{\varphi}_\partial) \quad (4) \\ &= \frac{1}{3}(d\partial\boldsymbol{\varphi}_\delta + \partial\delta\boldsymbol{\varphi}_d + \delta d\boldsymbol{\varphi}_\partial) - \frac{1}{6}(\boldsymbol{\varphi}_d \times \boldsymbol{\varphi}_{\partial\delta} + \boldsymbol{\varphi}_\partial \times \boldsymbol{\varphi}_{\delta d} + \boldsymbol{\varphi}_\delta \times \boldsymbol{\varphi}_{d\partial}) \\ &\quad - \frac{1}{3}(\boldsymbol{\varphi}_d \otimes \boldsymbol{\varphi}_\partial \cdot \boldsymbol{\varphi}_\delta + \boldsymbol{\varphi}_\partial \otimes \boldsymbol{\varphi}_\delta \cdot \boldsymbol{\varphi}_d + \boldsymbol{\varphi}_\delta \otimes \boldsymbol{\varphi}_d \cdot \boldsymbol{\varphi}_\partial),\end{aligned}$$

together with the identity

$$d\boldsymbol{\varphi}_{\partial\delta} + \partial\boldsymbol{\varphi}_{\delta d} + \delta\boldsymbol{\varphi}_{d\partial} = d\partial\boldsymbol{\varphi}_\delta + \partial\delta\boldsymbol{\varphi}_d + \delta d\boldsymbol{\varphi}_\partial.$$

In practical applications, a parameterization of the rotation tensor must be resorted to, and the differential rotation vectors are solved for the differentials of the parameters chosen. Evaluation of the first differential vector follows from its definition  $\boldsymbol{\varphi}_\delta = \text{ax}(\delta\boldsymbol{\Phi}\boldsymbol{\Phi}^T)$ ; then, the next differential vectors  $\boldsymbol{\varphi}_{\partial\delta}$  and  $\boldsymbol{\varphi}_{d\partial\delta}$  are evaluated, via (3) or (4), by differentiating the lower-order differential vectors themselves—which in turn is a parameterization-dependent operation. Such customized evaluations of the differential vectors are addressed in Section 5, with focus on our preferred natural parameterization.

Alternative expressions, based on corotational instead of direct differentiations of lower-order differential vectors, are available for  $\boldsymbol{\varphi}_{\partial\delta}$  and  $\boldsymbol{\varphi}_{d\partial\delta}$ . Introducing the corotational differentiations

$$\begin{aligned}\boldsymbol{\Phi} \partial(\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\delta) &= \partial\boldsymbol{\varphi}_\delta - \boldsymbol{\varphi}_\partial \times \boldsymbol{\varphi}_\delta, \\ \boldsymbol{\Phi} d(\boldsymbol{\Phi}^T \boldsymbol{\varphi}_{\partial\delta}) &= d\boldsymbol{\varphi}_{\partial\delta} - \boldsymbol{\varphi}_d \times \boldsymbol{\varphi}_{\partial\delta}, \\ \boldsymbol{\Phi} d\partial(\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\delta) &= d\partial\boldsymbol{\varphi}_\delta - \boldsymbol{\varphi}_{d\partial} \times \boldsymbol{\varphi}_\delta - (\boldsymbol{\varphi}_d \times \boldsymbol{\varphi}_{\partial\delta} + \boldsymbol{\varphi}_\partial \times \boldsymbol{\varphi}_{\delta d}),\end{aligned}\quad (5)$$

and the related formulae

$$\begin{aligned}\boldsymbol{\Phi} d(\boldsymbol{\Phi}^T \boldsymbol{\varphi}_{\partial\delta}) + \boldsymbol{\Phi} \partial(\boldsymbol{\Phi}^T \boldsymbol{\varphi}_{\delta d}) + \boldsymbol{\Phi} \delta(\boldsymbol{\Phi}^T \boldsymbol{\varphi}_{d\partial}) &= \boldsymbol{\Phi} d\partial(\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\delta) + \boldsymbol{\Phi} \partial\delta(\boldsymbol{\Phi}^T \boldsymbol{\varphi}_d) + \boldsymbol{\Phi} \delta d(\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\partial) \\ &= d\boldsymbol{\varphi}_{\partial\delta} + \partial\boldsymbol{\varphi}_{\delta d} + \delta\boldsymbol{\varphi}_{d\partial} - (\boldsymbol{\varphi}_d \times \boldsymbol{\varphi}_{\partial\delta} + \boldsymbol{\varphi}_\partial \times \boldsymbol{\varphi}_{\delta d} + \boldsymbol{\varphi}_\delta \times \boldsymbol{\varphi}_{d\partial}) \\ &= d\partial\boldsymbol{\varphi}_\delta + \partial\delta\boldsymbol{\varphi}_d + \delta d\boldsymbol{\varphi}_\partial - (\boldsymbol{\varphi}_d \times \boldsymbol{\varphi}_{\partial\delta} + \boldsymbol{\varphi}_\partial \times \boldsymbol{\varphi}_{\delta d} + \boldsymbol{\varphi}_\delta \times \boldsymbol{\varphi}_{d\partial}),\end{aligned}$$

equation (3) can be written

$$\begin{aligned}\boldsymbol{\varphi}_{\partial\delta} &= \boldsymbol{\Phi} \partial(\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\delta) + \frac{1}{2}\boldsymbol{\varphi}_\partial \times \boldsymbol{\varphi}_\delta, \\ \boldsymbol{\varphi}_{d\partial\delta} &= \boldsymbol{\Phi} d(\boldsymbol{\Phi}^T \boldsymbol{\varphi}_{\partial\delta}) + \frac{1}{2}\boldsymbol{\varphi}_d \times \boldsymbol{\varphi}_{\partial\delta} - \frac{1}{2}(\boldsymbol{\varphi}_d \otimes \boldsymbol{\varphi}_\partial \cdot \boldsymbol{\varphi}_\delta + (\boldsymbol{\varphi}_\partial \otimes \boldsymbol{\varphi}_\delta)^S \cdot \boldsymbol{\varphi}_d) \\ &= \boldsymbol{\Phi} d\partial(\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\delta) + \frac{1}{2}\boldsymbol{\varphi}_{d\partial} \times \boldsymbol{\varphi}_\delta + \frac{1}{2}(\boldsymbol{\varphi}_d \times \boldsymbol{\varphi}_{\partial\delta} + \boldsymbol{\varphi}_\partial \times \boldsymbol{\varphi}_{\delta d}) - (\boldsymbol{\varphi}_d \otimes \boldsymbol{\varphi}_\partial)^S \cdot \boldsymbol{\varphi}_\delta,\end{aligned}\quad (6)$$

and the symmetrical forms (4) become

$$\begin{aligned}
\varphi_{\partial\delta} &= \frac{1}{2} (\Phi \partial(\Phi^T \varphi_\delta) + \Phi \delta(\Phi^T \varphi_\partial)), \\
\varphi_{d\partial\delta} &= \frac{1}{3} (\Phi d(\Phi^T \varphi_{\partial\delta}) + \Phi \partial(\Phi^T \varphi_{\delta d}) + \Phi \delta(\Phi^T \varphi_{d\partial})) + \frac{1}{6} (\varphi_d \times \varphi_{\partial\delta} + \varphi_\partial \times \varphi_{\delta d} + \varphi_\delta \times \varphi_{d\partial}) \\
&\quad - \frac{1}{3} (\varphi_d \otimes \varphi_\partial \cdot \varphi_\delta + \varphi_\partial \otimes \varphi_\delta \cdot \varphi_d + \varphi_\delta \otimes \varphi_d \cdot \varphi_\partial) \\
&= \frac{1}{3} (\Phi d\partial(\Phi^T \varphi_\delta) + \Phi \partial\delta(\Phi^T \varphi_d) + \Phi \delta d(\Phi^T \varphi_\partial)) + \frac{1}{6} (\varphi_d \times \varphi_{\partial\delta} + \varphi_\partial \times \varphi_{\delta d} + \varphi_\delta \times \varphi_{d\partial}) \\
&\quad - \frac{1}{3} (\varphi_d \otimes \varphi_\partial \cdot \varphi_\delta + \varphi_\partial \otimes \varphi_\delta \cdot \varphi_d + \varphi_\delta \otimes \varphi_d \cdot \varphi_\partial).
\end{aligned}$$

Combining (3), (5), and (6), some remarkable relations follow:

$$\begin{aligned}
\Phi \partial(\Phi^T \varphi_\delta) &= \partial\varphi_\delta - \varphi_\partial \times \varphi_\delta \\
&= \delta\varphi_\partial,
\end{aligned} \tag{7}$$

$$\begin{aligned}
\Phi d(\Phi^T \varphi_{\partial\delta}) &= d\varphi_{\partial\delta} - \varphi_d \times \varphi_{\partial\delta} \\
&= \partial\delta\varphi_d - \frac{1}{2} (\varphi_\partial \times \delta\varphi_d + \varphi_\delta \times \partial\varphi_d),
\end{aligned} \tag{8}$$

$$\begin{aligned}
\Phi d\partial(\Phi^T \varphi_\delta) &= d\partial\varphi_\delta - \varphi_{d\partial} \times \varphi_\delta - (\varphi_d \times \varphi_{\partial\delta} + \varphi_\partial \times \varphi_{\delta d}) \\
&= \delta\varphi_{d\partial} - \frac{1}{2} (\varphi_d \times \delta\varphi_\partial + \varphi_\partial \times \delta\varphi_d).
\end{aligned} \tag{9}$$

They will be exploited in Section 5.

### 3. The exponential and subexponential maps

The structure of the differentiations of the rotation tensor discussed above has general validity, independently of any particular parameterization. From now on, a specific parameterization is assumed instead. We adopt the natural vectorial parameterization and resort to the so-called *exponential map*,

$$\Phi = \exp(\varphi \times) = \sum_{n=0}^{\infty} \frac{1}{n!} \varphi \times^n, \tag{10}$$

where  $\varphi \times$  is the skew-symmetric tensor built on the rotation vector  $\varphi$  (refer, for example, to [Argyris 1982; Ritto-Corrêa and Camotim 2002; Bauchau and Trainelli 2003; Mäkinen 2008]).

In this section we propose a helpful representation of the rotation tensor by means of a family of recursive subexponential maps; such representation will be profitably exploited in the parameterized differentiations in Sections 4 and 5. However, let us introduce this family in the realm of ordinary scalar functions first.

**3.1. The family of subexponential functions.** Consider the exponential function and the relevant series expansion  $X(x) = \exp(x) = \sum_{n=0}^{\infty} (1/n!)x^n$ . Collect all terms after the first (unity) and take the second term itself as the common factor of a subsequent series  $X_1(x)$ . Operate in the same way on  $X_1(x)$  to define a subsequent series  $X_2(x)$ , and so on. By proceeding recursively, an infinite family of functions

$X_0(x), X_1(x), X_2(x), \dots$  is defined:

$$\begin{aligned}
 X(x) &= 1 + x + \underbrace{\frac{x^2}{2} + \frac{x^3}{6} + \dots}_{X_0(x)} && = X_0(x) \\
 &= 1 + x \underbrace{\left(1 + \frac{x}{2} + \frac{x^2}{6} + \frac{x^3}{24} + \dots\right)}_{X_1(x)} && = 1 + x X_1(x) \\
 &= 1 + x \left(1 + \frac{x}{2} \underbrace{\left(1 + \frac{x}{3} + \frac{x^2}{12} + \frac{x^3}{60} + \dots\right)}_{X_2(x)}\right) && = 1 + x \left(1 + \frac{x}{2} X_2(x)\right) \\
 &= 1 + x \left(1 + \frac{x}{2} \left(1 + \frac{x}{3} \underbrace{\left(1 + \frac{x}{4} + \frac{x^2}{20} + \frac{x^3}{120} + \dots\right)}_{X_3(x)}\right)\right) && = 1 + x \left(1 + \frac{x}{2} \left(1 + \frac{x}{3} X_3(x)\right)\right) \\
 &= \dots && = \dots
 \end{aligned} \tag{11}$$

In (11), functions  $X_m(x)$  ( $m = 0, 1, 2, \dots$ ) are nested together inside each other as matryoshkas. Alternatively, the exponential function can be given one of the following forms of finite series:

$$\begin{aligned}
 X(x) &= X_0(x) \\
 &= 1 + x X_1(x) \\
 &= 1 + x + \frac{1}{2} x^2 X_2(x) \\
 &= \dots \\
 &= \sum_{n=0}^{m-1} \frac{1}{n!} x^n + \frac{1}{m!} x^m X_m(x).
 \end{aligned} \tag{12}$$

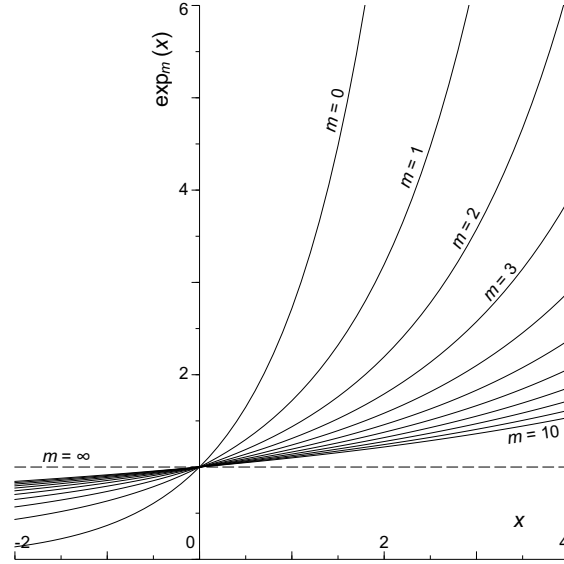
The meaning of (12) is clear: truncating the series expansion of  $X(x)$  at the  $m$ -th term is exact, provided the last retained term is multiplied by the  $m$ -th element of the family  $X_0(x), X_1(x), X_2(x), \dots$ . Thus, the elements of this family behave as plugs to properly truncate the series expansion of  $X(x)$ .

Functions  $X_m(x)$  are equipped with the following series expansions:

$$X_m(x) = \exp_m(x) = \sum_{n=0}^{\infty} \frac{m!}{(n+m)!} x^n \quad (\forall m \geq 0), \tag{13}$$

where we resort to the symbolic notation  $\exp_m(x)$  owing to the evident similarity with the exponential expansion. We refer to the set of functions  $\exp_m(x)$  as the family of *subexponential* functions (a plot of the first few is illustrated in Figure 1). The base function  $\exp_0(x)$  is the exponential function itself, whereas at very high integer  $m$  the subexponential function approaches unity,  $\exp_{m \rightarrow \infty}(x) \rightarrow 1$ .

It is clear from the nesting representation (11) that each subexponential function can in turn be expressed via some higher subexponential functions. This recursive character allows us to extend the representation (12) to the whole family, in the form of the relation



**Figure 1.** The family of subexponential functions.

$$X_m(x) = \sum_{n=0}^{l-1} \frac{m!}{(n+m)!} x^n + \frac{m!}{(l+m)!} x^l X_{l+m}(x) \quad (\forall m \geq 0, \forall l \geq 1), \quad (14)$$

which can also be obtained directly from (13).

**3.2. The family of subexponential maps.** The family of *subexponential maps*  $\Phi_0, \Phi_1, \Phi_2, \dots$  of the rotation tensor is introduced exactly the same way as above. Tensors  $\Phi_m$  ( $m = 0, 1, 2, \dots$ ) take the series expansion form

$$\Phi_m = \exp_m(\varphi \times) = \sum_{n=0}^{\infty} \frac{m!}{(n+m)!} \varphi \times^n \quad (\forall m \geq 0). \quad (15)$$

The first element  $\exp_0(\varphi \times)$  of the family coincides with the exponential map (10) and the asymptotic element is the unit tensor,  $\exp_{m \rightarrow \infty}(\varphi \times) \rightarrow \mathbf{I}$ . The subexponential maps are nested together just like the subexponential functions are in (11) and allow us to properly truncate the exponential map as in (12). This truncation property applies recursively to each element of the family, so, as in (14), we can state

$$\Phi_m = \sum_{n=0}^{l-1} \frac{m!}{(n+m)!} \varphi \times^n + \frac{m!}{(l+m)!} \varphi \times^l \Phi_{l+m} \quad (\forall m \geq 0, \forall l \geq 1).$$

In particular, the strongest truncation ( $l = 1$ ),

$$\Phi_m = \mathbf{I} + \frac{1}{m+1} \varphi \times \Phi_{m+1} \quad (\forall m \geq 0), \quad (16)$$

lets us envisage a nested representation of the rotation tensor employing all the subexponential maps in cascade.

It can be noted that all the subexponential maps  $\exp_m(\varphi \times)$  share the same eigenvector  $\varphi$ ; they are transparent to their argument,  $\varphi \times \Phi_m = \Phi_m \varphi \times$ , and commute with each other,  $\Phi_m \Phi_n = \Phi_n \Phi_m$ . Other useful identities involving subsequent tensors  $\Phi_m$  (with  $m \geq 1$ ) are easily proven:

$$\begin{aligned} \varphi \times \Phi_m &= m(\Phi_{m-1} - I), \\ \Phi_{m-1} - \Phi_m &= \frac{1}{m} \sum_{n=1}^{\infty} \frac{m!}{(n+m)!} \varphi \times^n \Phi_{n+m} = \left( \frac{1}{m} \Phi_m - \frac{1}{m+1} \Phi_{m+1} \right) \varphi \times, \\ (\Phi_m - \Phi_{m+1}) \varphi \times &= I + m \Phi_{m-1} - (m+1) \Phi_m, \\ \Phi_m \Phi_0^T &= \sum_{n=1}^m \frac{(-1)^{n-1} m!}{n!(m-n)!} \Phi_n^T. \end{aligned} \quad (17)$$

Finally, the following property is quoted:

$$(\Phi_m - \Phi_n) I^\times \varphi \times = \varphi \times I^\times (\Phi_m - \Phi_n) \quad (\forall m, n),$$

which can be proved using the expansion (15) and the tensor identities (A.7), (A.13), and (A.14). Tensor  $I^\times$  is the unitary third-order tensor of skew-symmetrical nature, namely Ricci's tensor (A.6).

**3.3. Compact form.** Series expansions  $\Phi_m$  can be brought to forms made of a finite number of terms and more suited for computations. This is accomplished by recursively using the formula  $\varphi \times^3 + \varphi^2 \varphi \times = \mathbf{0}$ , where  $\varphi$  is the magnitude of vector  $\varphi$ , namely the rotation angle. This formula follows from the Cayley–Hamilton theorem applied to skew-symmetric tensor  $\varphi \times$ , or it can be drawn directly from tensor identity (A.9); it can be generalized to any power  $\varphi \times^n$  ( $\forall n \geq 1$ ) in the forms

$$\begin{aligned} \varphi \times^{2n-1} &= (-1)^{n-1} \varphi^{2(n-1)} \varphi \times, \\ \varphi \times^{2n} &= (-1)^{n-1} \varphi^{2(n-1)} \varphi \times^2. \end{aligned}$$

Using these formulae and the trigonometric series expansions in (15) yields

$$\begin{aligned} \Phi_0 &= I + a \varphi \times + b_0 \varphi \times^2, \\ \Phi_m &= I + m b_{m-1} \varphi \times + b_m \varphi \times^2 \quad (\forall m \geq 1). \end{aligned} \quad (18)$$

Equation (18) is referred to as the *compact form* of the family of tensors  $\Phi_m$ ; they stand on three minimal tensorial bases, the identity  $I$  and tensors  $\varphi \times$  and  $\varphi \times^2$ . The coefficients  $a$  and  $b_m$  are recursive functions of the rotation angle,

$$\begin{aligned} a &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \varphi^{2n} = \frac{1}{\varphi} \sin \varphi, \\ b_0 &= 0! \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+2)!} \varphi^{2n} = \frac{1}{\varphi^2} (1 - \cos \varphi), \\ b_1 &= 1! \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+3)!} \varphi^{2n} = \frac{1}{\varphi^2} (1 - a), \\ b_m &= m! \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+2+m)!} \varphi^{2n} = \frac{1}{\varphi^2} (1 - m(m-1)b_{m-2}) \quad (\forall m \geq 2). \end{aligned} \quad (19)$$

Details about the computational implementation of these coefficient functions are discussed in [Merlini and Morandini 2004b, Appendix A].

It is seen that  $(18)_1$  is the well-known Euler–Rodrigues formula of the rotation tensor [Cheng and Gupta 1989]. The expression for  $\Phi_1$  in  $(18)_2$  corresponds to another well-known tensor in the finite-rotation literature, sometimes referred to as the associated differential tensor [Bauchau and Trainelli 2003]. This tensor represents the mapping of the differential  $d\boldsymbol{\varphi}$  of the rotation vector onto the differential rotation vector  $\boldsymbol{\varphi}_d$ , which characterizes the tangent space of the rotation [Ibrahimbegović et al. 1995; Borri et al. 2000; Ritto-Corrêa and Camotim 2002; Mäkinen 2008].

#### 4. Differentiation of subexponential maps

Subexponential maps  $\exp_m(\boldsymbol{\varphi} \times)$  are power series expansions of the argument itself and as such are differentiable tensor functions. In fact, each differential  $d(\boldsymbol{\varphi} \times^n)$  of a power term in (15) is easily brought to the form [de Souza Neto 2001]

$$d(\boldsymbol{\varphi} \times^n) = \sum_{k=1}^n \boldsymbol{\varphi} \times^{k-1} d\boldsymbol{\varphi} \times \boldsymbol{\varphi} \times^{n-k} \quad (\forall n > 0), \quad (20)$$

linear with respect to  $d\boldsymbol{\varphi} \times$ . So, the differential of a tensor  $\Phi_m$  can be brought to an expression like  $d\Phi_m = \Phi_{m/}^{T132} \cdot d\boldsymbol{\varphi}$ , linear with the variation of the rotation vector, where  $\Phi_{m/}$  is a third-order tensor, a function itself of powers of  $\boldsymbol{\varphi} \times$ . It follows that tensor  $\Phi_{m/}$  is a differentiable tensor itself, and as a consequence tensor  $\Phi_m$  is two times differentiable.

Recursively, it can be seen that the subexponential maps are continuously differentiable with respect to their vector argument. However, for present purposes, our interest is in the first two successive, and independent, differentiations, say  $\delta$  and  $\partial$ . The relevant differentials can be cast in the form

$$\begin{aligned} \delta\Phi_m &= \Phi_{m/} : \delta\boldsymbol{\varphi} \otimes \mathbf{I}, & \partial\Phi_{m/} &= \Phi_{m//}^{1234} : \partial\boldsymbol{\varphi} \otimes \mathbf{I}, \\ \partial\delta\Phi_m &= \Phi_{m//} : \partial\delta\boldsymbol{\varphi} \otimes \mathbf{I} + \Phi_{m//}^{1234} : \partial\boldsymbol{\varphi} \otimes \delta\boldsymbol{\varphi} \otimes \mathbf{I}, \end{aligned} \quad (21)$$

where  $\Phi_{m//}^{1234}$  is a fourth-order tensor, symmetric with respect to the two inner polyadic legs (see (24); refer to the Appendix for notation and rules), and the variations  $\partial$  and  $\delta$  are of course interchangeable. We may refer to tensors  $\Phi_{m/}$  and  $\Phi_{m//}^{1234}$  as the *first derivative* and the *second derivative*, respectively, of tensor  $\Phi_m$  with respect to the rotation vector.

Equation (21) defines the first two derivatives of the subexponential maps, however successive derivatives can analogously be defined on demands. The recursive forms developed in Section 4.2 pave the way for building successive derivatives virtually up to any order. In this section we derive useful expressions of the first two derivative tensors.

**4.1. Series expansion form of the derivative tensors.** Simple series expansion forms of the derivative tensors are easily obtained recalling (20). Differentiation of (15) yields

$$\delta\Phi_m = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{m!}{(1+j+k+m)!} \boldsymbol{\varphi} \times^j \delta\boldsymbol{\varphi} \times \boldsymbol{\varphi} \times^k = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{-m!}{(1+j+k+m)!} \boldsymbol{\varphi} \times^j \mathbf{I} \times \boldsymbol{\varphi} \times^k : \delta\boldsymbol{\varphi} \otimes \mathbf{I}, \quad (22)$$

and comparison with (21)<sub>1</sub> provides the first derivative tensor in the power series expansion form

$$\Phi_{m/} = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{-m!}{(1+j+k+m)!} \varphi \times^j I^\times \varphi \times^k. \quad (23)$$

Subsequently, differentiation of (23) yields

$$\begin{aligned} \partial \Phi_{m/} &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{-m!}{(2+j+k+l+m)!} \varphi \times^j (\partial \varphi \times \varphi \times^k I^\times + I^\times \varphi \times^k \partial \varphi \times) \varphi \times^l \\ &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{2m!}{(2+j+k+l+m)!} (\varphi \times^j I^\times \varphi \times^k I^\times \varphi \times^l)^{S_{1234}} : \partial \varphi \otimes I, \end{aligned} \quad (24)$$

and comparison with (21)<sub>2</sub> provides the second derivative tensor in the power series expansion form

$$\Phi_{m//}^{1234} = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{2m!}{(2+j+k+l+m)!} (\varphi \times^j I^\times \varphi \times^k I^\times \varphi \times^l)^{S_{1234}}. \quad (25)$$

Equations (23) and (25) are concise straightforward expressions of the derivative tensors, however they may not be fit for numerical computations. More suitable expressions are derived in next subsections.

**4.2. Recursive form of the derivative tensors.** Differential (22) can be worked out using (15), (16), (17)<sub>1</sub>, and (17)<sub>2</sub>. After some algebraic manipulations, the following expression is obtained:

$$\begin{aligned} \delta \Phi_m &= \frac{1}{m+1} \left( I^\times - (I^\times \Phi_{m+1} + \Phi_{m+1} I^\times) - \frac{1}{2} (\varphi \times I^\times (\Phi_{m+1} - \Phi_{m+2}) \right. \\ &\quad \left. + (\Phi_{m+1} - \Phi_{m+2}) I^\times \varphi \times) \right) : \delta \varphi \otimes I. \end{aligned} \quad (26)$$

Recalling (21), subsequent differentiation yields

$$\begin{aligned} \partial \delta \Phi_m &= \frac{1}{m+1} \left( I^\times - (I^\times \Phi_{m+1} + \Phi_{m+1} I^\times) - \frac{1}{2} (\varphi \times I^\times (\Phi_{m+1} - \Phi_{m+2}) + (\Phi_{m+1} - \Phi_{m+2}) I^\times \varphi \times) \right) : \partial \delta \varphi \otimes I \\ &\quad + \frac{1}{m+1} \left( \frac{1}{2} (I^\times I^\times (\Phi_{m+1} - \Phi_{m+2}) + (\Phi_{m+1} - \Phi_{m+2}) I^\times I^\times) - (I^\times \Phi_{m+1/} + \Phi_{m+1/} I^\times) \right. \\ &\quad \left. - \frac{1}{2} (\varphi \times I^\times (\Phi_{m+1/} - \Phi_{m+2/}) + (\Phi_{m+1/} - \Phi_{m+2/}) I^\times \varphi \times) \right)^{S_{1234}} : \partial \varphi \otimes \delta \varphi \otimes I. \end{aligned} \quad (27)$$

In (26) and (27), new expressions of the derivative tensors  $\Phi_{m/}$  and  $\Phi_{m//}^{1234}$ , defined in (21), are clearly recognized. They are referred to as the *recursive form* of the derivatives of tensor  $\Phi_m$  as they are functions of the next subexponential maps and of their first derivative tensors. Such recursive forms are better suited for numerical computations than the series expansion forms.

**4.3. Compact form of the derivative tensors.** Other expressions of the derivative tensors can be drawn from (26) and (27) using the compact forms (18) of the subexponential maps. These new expressions

stand on five tensorial bases and are called the *compact form* of the derivative tensors:

$$\begin{aligned} \Phi_{m/} = f_{m0} \mathbf{I}^\times + f_{m1} (\boldsymbol{\varphi} \times \mathbf{I}^\times + \mathbf{I}^\times \boldsymbol{\varphi} \times) + f_{m2} (\boldsymbol{\varphi} \times^2 \mathbf{I}^\times + \boldsymbol{\varphi} \times \mathbf{I}^\times \boldsymbol{\varphi} \times + \mathbf{I}^\times \boldsymbol{\varphi} \times^2) \\ + f_{m3} \frac{1}{2} (\boldsymbol{\varphi} \times^2 \mathbf{I}^\times \boldsymbol{\varphi} \times + \boldsymbol{\varphi} \times \mathbf{I}^\times \boldsymbol{\varphi} \times^2) + f_{m4} \boldsymbol{\varphi} \times^2 \mathbf{I}^\times \boldsymbol{\varphi} \times^2, \end{aligned} \quad (28)$$

$$\begin{aligned} \Phi_{m//}^{1234} = (g_{m0} \mathbf{I}^\times \mathbf{I}^\times + g_{m1} (\boldsymbol{\varphi} \times \mathbf{I}^\times \mathbf{I}^\times + \mathbf{I}^\times \boldsymbol{\varphi} \times \mathbf{I}^\times + \mathbf{I}^\times \mathbf{I}^\times \boldsymbol{\varphi} \times) + g_{m2} (\boldsymbol{\varphi} \times^2 \mathbf{I}^\times \mathbf{I}^\times + \mathbf{I}^\times \boldsymbol{\varphi} \times^2 \mathbf{I}^\times + \mathbf{I}^\times \mathbf{I}^\times \boldsymbol{\varphi} \times^2 \\ + \boldsymbol{\varphi} \times \mathbf{I}^\times \boldsymbol{\varphi} \times \mathbf{I}^\times + \boldsymbol{\varphi} \times \mathbf{I}^\times \mathbf{I}^\times \boldsymbol{\varphi} \times + \mathbf{I}^\times \boldsymbol{\varphi} \times \mathbf{I}^\times \boldsymbol{\varphi} \times + \boldsymbol{\varphi}^2 \mathbf{I}^\times \mathbf{I}^\times) \\ + g_{m3} \boldsymbol{\varphi} \times \mathbf{I}^\times \boldsymbol{\varphi} \times \mathbf{I}^\times \boldsymbol{\varphi} \times + g_{m4} \boldsymbol{\varphi} \times \mathbf{I}^\times \boldsymbol{\varphi} \times^2 \mathbf{I}^\times \boldsymbol{\varphi} \times) \stackrel{S1234}{.} \end{aligned} \quad (29)$$

The coefficient functions in (28) and (29) are defined as follows:

$$\begin{aligned} f_{m0} &= \frac{-1}{m+1}, & g_{m0} &= 2b_m, \\ f_{m1} &= -b_m, & g_{m1} &= b_m - b_{m+1}, \\ f_{m2} &= \frac{-1}{m+1} b_{m+1}, & g_{m2} &= \frac{1}{m+1} (b_{m+1} - b_{m+2}), \\ f_{m3} &= \frac{-1}{m+1} (b_{m+1} - b_{m+2}), & g_{m3} &= \frac{1}{m+1} (b_{m+1} - b_{m+2}) - \frac{1}{m+2} (b_{m+2} - b_{m+3}), \\ f_{m4} &= \frac{-1}{(m+1)(m+2)} (b_{m+2} - b_{m+3}), & g_{m4} &= \frac{1}{m+1} \left( \frac{1}{m+2} (b_{m+2} - b_{m+3}) - \frac{1}{m+3} (b_{m+3} - b_{m+4}) \right). \end{aligned}$$

## 5. Differential maps of the rotation

The differential maps of the rotation transform the multiple differentials of the rotation vector into the differential rotation vectors defined in Section 2. The differentiations of the lowest two subexponential maps, as derived above, are used in this section to develop the expected differential maps and to highlight some important properties.

**5.1. Explicit notation for the lowest two subexponential maps.** The first and second subexponential maps play important roles in computational mechanics. As mentioned in Section 3.3,  $\Phi_0$  is the exponential map itself (that is, the rotation tensor), whereas  $\Phi_1$  is known as its tangent map or the associated differential map [Borri et al. 2000]; in the present context, the latter will be referred to concisely as the *tangent tensor* (of the rotation). Both of these tensors are assigned hereafter specific notations:

$$\begin{aligned} \Phi &= \Phi_0 = \exp(\boldsymbol{\varphi} \times) = \sum_{n=0}^{\infty} \frac{1}{n!} \boldsymbol{\varphi} \times^n = \mathbf{I} + a \boldsymbol{\varphi} \times + b_0 \boldsymbol{\varphi} \times^2, \\ \Gamma &= \Phi_1 = \exp_1(\boldsymbol{\varphi} \times) = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} \boldsymbol{\varphi} \times^n = \mathbf{I} + b_0 \boldsymbol{\varphi} \times + b_1 \boldsymbol{\varphi} \times^2. \end{aligned} \quad (30)$$

Note that,  $\Gamma$  being the strongest truncation of  $\Phi$ , the relation  $\Phi = \mathbf{I} + \boldsymbol{\varphi} \times \Gamma$  holds as derived directly from (16).



The first and second differentiations of tensors  $\Phi$  and  $\Gamma$  are written, from (21), as

$$\begin{aligned}\delta\Phi &= \Phi_{/} : \delta\varphi \otimes I, & \partial\Phi_{/} &= \Phi_{//}^{1234} : \partial\varphi \otimes I, \\ \partial\delta\Phi &= \Phi_{/} : \partial\delta\varphi \otimes I + \Phi_{//}^{1234} : \partial\varphi \otimes \delta\varphi \otimes I,\end{aligned}\tag{31}$$

and

$$\begin{aligned}\delta\Gamma &= \Gamma_{/} : \delta\varphi \otimes I, & \partial\Gamma_{/} &= \Gamma_{//}^{1234} : \partial\varphi \otimes I, \\ \partial\delta\Gamma &= \Gamma_{/} : \partial\delta\varphi \otimes I + \Gamma_{//}^{1234} : \partial\varphi \otimes \delta\varphi \otimes I.\end{aligned}\tag{32}$$

Expressions of the relevant derivative tensors,

$$\begin{aligned}\Phi_{/} &= \Phi_{0/}, & \Phi_{//}^{1234} &= \Phi_{0//}^{1234}, \\ \Gamma_{/} &= \Phi_{1/}, & \Gamma_{//}^{1234} &= \Phi_{1//}^{1234},\end{aligned}\tag{33}$$

are easily written by setting  $m = 0$  or  $m = 1$  in (22)–(29). They are not repeated here.

The compact forms of the derivative tensors (33) are known in the literature. Let us use (28) and (29), together with the coefficient functions (19), to evaluate the following second-order tensors built with tensors (33) and arbitrary vectors  $\mathbf{u}$  and  $\mathbf{v}$ :

$$\begin{aligned}\Phi_{/} : \mathbf{u} \otimes I &= a\mathbf{u} \times + b_0(\mathbf{u} \otimes \varphi + \varphi \otimes \mathbf{u}) - a\varphi \cdot \mathbf{u} \otimes I \\ &\quad - (b_0 - b_1)\varphi \cdot \mathbf{u} \otimes \varphi \times - (b_1 - b_2)\varphi \cdot \mathbf{u} \otimes \varphi \otimes \varphi,\end{aligned}\tag{34}$$

$$\begin{aligned}\Gamma_{/} : \mathbf{u} \otimes I &= b_0\mathbf{u} \times + b_1(\mathbf{u} \otimes \varphi + \varphi \otimes \mathbf{u}) - (b_0 - b_1)\varphi \cdot \mathbf{u} \otimes I \\ &\quad - (b_1 - b_2)\varphi \cdot \mathbf{u} \otimes \varphi \times - \frac{1}{2}(b_2 - b_3)\varphi \cdot \mathbf{u} \otimes \varphi \otimes \varphi,\end{aligned}$$

and

$$\begin{aligned}\Phi_{//}^{1234} : \mathbf{u} \otimes \mathbf{v} \otimes I &= b_0(\mathbf{u} \otimes \mathbf{v} + \mathbf{v} \otimes \mathbf{u}) - a\mathbf{u} \cdot \mathbf{v} \otimes I - (b_0 - b_1)(\varphi \cdot \mathbf{u} \otimes \mathbf{v} \times + \varphi \cdot \mathbf{v} \otimes \mathbf{u} \times + \mathbf{u} \cdot \mathbf{v} \otimes \varphi \times) \\ &\quad - (b_1 - b_2)(\varphi \cdot \mathbf{u} \otimes (\mathbf{v} \otimes \varphi + \varphi \otimes \mathbf{v}) + \varphi \cdot \mathbf{v} \otimes (\mathbf{u} \otimes \varphi + \varphi \otimes \mathbf{u}) + \mathbf{u} \cdot \mathbf{v} \otimes \varphi \otimes \varphi) \\ &\quad + (b_0 - b_1)\varphi \cdot \mathbf{v} \otimes \varphi \cdot \mathbf{u} \otimes I + ((b_1 - b_2) - \frac{1}{2}(b_2 - b_3))\varphi \cdot \mathbf{v} \otimes \varphi \cdot \mathbf{u} \otimes \varphi \times \\ &\quad + (\frac{1}{2}(b_2 - b_3) - \frac{1}{3}(b_3 - b_4))\varphi \cdot \mathbf{v} \otimes \varphi \cdot \mathbf{u} \otimes \varphi \otimes \varphi,\end{aligned}\tag{35}$$

$$\begin{aligned}\Gamma_{//}^{1234} : \mathbf{u} \otimes \mathbf{v} \otimes I &= b_1(\mathbf{u} \otimes \mathbf{v} + \mathbf{v} \otimes \mathbf{u}) - (b_0 - b_1)\mathbf{u} \cdot \mathbf{v} \otimes I \\ &\quad - (b_1 - b_2)(\varphi \cdot \mathbf{u} \otimes \mathbf{v} \times + \varphi \cdot \mathbf{v} \otimes \mathbf{u} \times + \mathbf{u} \cdot \mathbf{v} \otimes \varphi \times) - \frac{1}{2}(b_2 - b_3)(\varphi \cdot \mathbf{u} \otimes (\mathbf{v} \otimes \varphi + \varphi \otimes \mathbf{v}) \\ &\quad + \varphi \cdot \mathbf{v} \otimes (\mathbf{u} \otimes \varphi + \varphi \otimes \mathbf{u}) + \mathbf{u} \cdot \mathbf{v} \otimes \varphi \otimes \varphi) + ((b_1 - b_2) - \frac{1}{2}(b_2 - b_3))\varphi \cdot \mathbf{v} \otimes \varphi \cdot \mathbf{u} \otimes I \\ &\quad + (\frac{1}{2}(b_2 - b_3) - \frac{1}{3}(b_3 - b_4))\varphi \cdot \mathbf{v} \otimes \varphi \cdot \mathbf{u} \otimes \varphi \times + \frac{1}{2}(\frac{1}{3}(b_3 - b_4) - \frac{1}{4}(b_4 - b_5))\varphi \cdot \mathbf{v} \otimes \varphi \cdot \mathbf{u} \otimes \varphi \otimes \varphi.\end{aligned}$$

It is seen that (34)<sub>1</sub> and (34)<sub>2</sub> coincide respectively with the directional derivatives  $D\Phi[\mathbf{u}]$  and  $D\Gamma[\mathbf{u}]$  found, respectively, in [Ritto-Corrêa and Camotim 2002, Equations (18) and (19)], once our coefficient functions (19) are converted to those used by them. It is also seen that (35)<sub>1</sub> and (35)<sub>2</sub> coincide, respectively, with the second directional derivatives  $D^2\Phi[\mathbf{u}, \mathbf{v}]$  and  $D^2\Gamma[\mathbf{u}, \mathbf{v}]$  [ibid., Equations (20) and (21)].

**5.2. Lowest three differential maps.** The differential (31)<sub>1</sub> of the rotation tensor is easily obtained from (26) with  $m = 0$ :

$$\delta\Phi = -\delta\varphi \times + (\delta\varphi \times \Phi_1 + \Phi_1 \delta\varphi \times) + \frac{1}{2}(\varphi \times \delta\varphi \times (\Phi_1 - \Phi_2) + (\Phi_1 - \Phi_2) \delta\varphi \times \varphi \times).$$

Using (16) with  $m = 1$ , namely  $\Phi_1 = I + \frac{1}{2}\varphi \times \Phi_2 = I + \frac{1}{2}\Phi_2\varphi \times$ , and the tensor identity (A.9),  $\delta\Phi$  is rewritten as

$$\delta\Phi = \delta\varphi \times + \frac{1}{2}(\varphi \otimes \delta\varphi \cdot \Phi_2 + \Phi_2 \cdot \delta\varphi \otimes \varphi) - \Phi_1 \otimes \varphi \cdot \delta\varphi$$

and is then multiplied by  $\Phi_0^T$  to draw an expression for tensor  $\delta\Phi\Phi^T$ . Using the property (17)<sub>4</sub> with  $m = 1$  (that is,  $\Phi_1\Phi_0^T = \Phi_1^T$ ) and  $m = 2$  (that is,  $\Phi_2\Phi_0^T = 2\Phi_1^T - \Phi_2^T$ ), the tensor identities (A.9) and (A.12), and (16) with  $m = 0$  and  $m = 1$ , the relation

$$\delta\Phi\Phi^T = (\Phi_1 \cdot \delta\varphi) \times$$

is finally obtained. Comparison with (2)<sub>1</sub> and (30)<sub>2</sub> provides the sought relation of  $\varphi_\delta$  as a linear function of  $\delta\varphi$ :

$$\varphi_\delta = \Gamma \cdot \delta\varphi, \quad (36)$$

a well-known result involving the tangent tensor  $\Gamma$ .

The foregoing derivation of the *first differential map* of the rotation is an alternative to other derivations found in the literature, for example, [Borri et al. 1990; Ibrahimbegović et al. 1995; Ritto-Corrêa and Camotim 2002].

Expressions for the next differential vectors are drawn directly from their definitions as elaborated in (4)<sub>1</sub> and (4)<sub>3</sub>, respectively. The *second differential map* is easily obtained using (36) and (32)<sub>1</sub>:

$$\varphi_{\partial\delta} = \Gamma \cdot \partial\delta\varphi + \Gamma_{/}^{S123} : \partial\varphi \otimes \delta\varphi, \quad (37)$$

where  $\Gamma_{/}^{S123} = \frac{1}{2}(\Gamma_{/} + \Gamma_{/}^{T132})$  is the right-symmetric part of  $\Gamma_{/}$  (that is, symmetric with respect to the two rightmost polyadic legs, see (A.2)<sub>1</sub>).

The *third differential map* is obtained, after a more involved derivation, in the form

$$\begin{aligned} \varphi_{d\partial\delta} = & \Gamma \cdot d\partial\delta\varphi + \Gamma_{/}^{S123} : (d\varphi \otimes \partial\delta\varphi + \partial\varphi \otimes d\delta\varphi + \delta\varphi \otimes d\partial\varphi) \\ & + (\Gamma_{//}^{1234} - \frac{1}{2}(I \times \Gamma)^{T132} \Gamma_{/}^{S123} - \Gamma \otimes \Gamma^T \Gamma)^{S1234} : d\varphi \otimes \partial\varphi \otimes \delta\varphi, \end{aligned} \quad (38)$$

where  $( )^{S1234}$  denotes the full-symmetric part, with respect to the three rightmost polyadic legs, of a fourth-order tensor, see (A.5). Derivation of (38) requires (36), (37), (32), and the property (39).

**5.3. Properties of the derivative tensors.** The derivative tensors  $\Gamma_{/}$  and  $\Gamma_{//}^{1234}$  are endowed with useful properties, which descend from the definition of the differential vectors themselves. Let us focus first on the property (7) and develop  $\Phi \partial(\Phi^T \varphi_\delta)$  from (7)<sub>1</sub> and (7)<sub>2</sub> separately, using the differential map (36) and the differentiation formula (32)<sub>1</sub>. Comparison of the results leads to

$$\Gamma_{/}^{T132} = \Gamma_{/} - (I \times \Gamma)^{T132} \Gamma. \quad (39)$$

Observing that  $(\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} \boldsymbol{\Gamma}$  is a right-skew-symmetric third-order tensor, see (A.17), (39) entails the right-symmetry of tensor  $\boldsymbol{\Gamma}_/ - \frac{1}{2} (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} \boldsymbol{\Gamma}$ , hence the property

$$\boldsymbol{\Gamma}_/ = \boldsymbol{\Gamma}_/^{S123} + \frac{1}{2} (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} \boldsymbol{\Gamma}. \quad (40)$$

Equation (40) represents the decomposition of the tangent-tensor first derivative into the symmetric and skew-symmetric parts of the rightmost polyadic legs. Note that property (7) enables a straightforward derivation of (40), while deriving the latter directly from the expressions of  $\boldsymbol{\Gamma}_/$  in (33) would be a much more involved task.

Next, let us focus on property (8) and develop  $\boldsymbol{\Phi} \text{d}(\boldsymbol{\Phi}^{\text{T}} \boldsymbol{\varphi}_{\partial\delta})$  from (8)<sub>1</sub> and (8)<sub>2</sub> separately, using the differential maps (36) and (37), the differentiation formulae (32), and the property (39). (Alternatively, one might start from property (9).) Comparison of the results leads to a useful property of the symmetric parts of another fourth-order tensor, specifically tensor  $\boldsymbol{\Gamma}_{//}^{1234} - (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} \boldsymbol{\Gamma}_/$ . This property can be written in the following equivalent forms:

$$\begin{aligned} (\boldsymbol{\Gamma}_{//}^{1234} - (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} \boldsymbol{\Gamma}_/)^{S1234} &= (\boldsymbol{\Gamma}_{//}^{1234} - (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} \boldsymbol{\Gamma}_/)^{\text{T}1342 S1234}, \\ (\boldsymbol{\Gamma}_{//}^{1234} - (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} \boldsymbol{\Gamma}_/)^{S1234} &= (\boldsymbol{\Gamma}_{//}^{1234} - (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} \boldsymbol{\Gamma}_/)^{\text{T}1423 S1234}, \\ (\boldsymbol{\Gamma}_{//}^{1234} - (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} \boldsymbol{\Gamma}_/)^{\text{T}1342 S1234} &= (\boldsymbol{\Gamma}_{//}^{1234} - (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} \boldsymbol{\Gamma}_/)^{\text{T}1423 S1234}. \end{aligned} \quad (41)$$

Other remarkable relations descend from the property

$$\boldsymbol{\Gamma} = \boldsymbol{\Gamma}^{\text{T}} \boldsymbol{\Phi} = \boldsymbol{\Phi} \boldsymbol{\Gamma}^{\text{T}}, \quad (42)$$

which is obtained by setting  $m = 1$  in (17)<sub>4</sub>. Equation (42) represents a factorization of the tangent tensor into itself, by means of the rotation tensor  $\boldsymbol{\Phi}$ .

Differentiating (42) and developing the identity  $\delta \boldsymbol{\Gamma} = \delta(\boldsymbol{\Phi} \boldsymbol{\Gamma}^{\text{T}})$  using (32)<sub>1</sub>, (2)<sub>1</sub>, (36), and (39), the relation  $\boldsymbol{\Gamma}_/^{\text{T}132} = \boldsymbol{\Phi} \boldsymbol{\Gamma}_/^{\text{T}321}$  is obtained; hence the factorization of the first derivative of the tangent tensor follows in the forms

$$\boldsymbol{\Gamma}_/ = \boldsymbol{\Gamma}_/^{\text{T}231} \boldsymbol{\Phi} = \boldsymbol{\Phi} \boldsymbol{\Gamma}_/^{\text{T}312}. \quad (43)$$

Differentiating (42) further and developing the identity  $\partial \delta \boldsymbol{\Gamma} = \partial \delta(\boldsymbol{\Phi} \boldsymbol{\Gamma}^{\text{T}})$  using (32), (2), (36), (37), and (39)–(43), a factorization of the second derivative of the tangent tensor is obtained in the form

$$\boldsymbol{\Gamma}_{//}^{1234} = (\boldsymbol{\Gamma}_{//}^{1234} - (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} \boldsymbol{\Gamma}_/)^{\text{T}3241 S1234} \boldsymbol{\Phi}. \quad (44)$$

Finally, factorization formulae of the derivative tensors  $\boldsymbol{\Phi}_/$  and  $\boldsymbol{\Phi}_{//}^{1234}$  are easily obtained from (31) using (2), (36) and (37), and (42) and (43):

$$\boldsymbol{\Phi}_/ = (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} \boldsymbol{\Phi} = \boldsymbol{\Phi} (\mathbf{I} \times \boldsymbol{\Gamma}^{\text{T}})^{\text{T}132}, \quad (45)$$

and

$$\boldsymbol{\Phi}_{//}^{1234} = ((\mathbf{I} \times \boldsymbol{\Gamma}_/)^{\text{T}1342} + (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132} (\mathbf{I} \times \boldsymbol{\Gamma})^{\text{T}132})^{S1234} \boldsymbol{\Phi}. \quad (46)$$

All the above properties will be exploited in the following sections.

## 6. Slender beam variational mechanics

As an application of the formulation discussed so far to computational elastostatics, we address the finite-element modeling of space-curved slender beams for geometrically nonlinear problems. In the context of a variational approach and a typical Newton–Raphson solution procedure, we are concerned with three independent variations, namely virtual variations, incremental variations, and field derivatives on a one-coordinate domain (the ordinary derivatives with respect to beam abscissa). The simple form taken by the spatial variations in this problem is the main motivation for the choice of this particular application, as it allows us to focus on the significance of correctly differentiating the rotation up to third order.

Within the domain of one-dimensional intrinsic continuum mechanics, we identify the beam cross-sections with material particles aligned along a curvilinear arc-length  $s$ . The particle position and orientation are independent configuration variables and allow for shear-deformable beam response. Customary hypotheses such as linear elastic material, small strains and small stresses with respect to the elastic moduli, constant cross-sections, and rigid-section deformations are understood. In these circumstances, we may refer to the beam theory proposed by Reissner [1973], reformulated by Simo [1985], and discussed in [Ritto-Corrêa and Camotim 2002] as Reissner–Simo beam theory (see also [Simo and Vu-Quoc 1986]). The essentials of the beam variational mechanics are addressed in this section, whereas a beam element is formulated in the next section. For our present purposes, we may focus on just the variational term arising from the internal virtual work.

### 6.1. Nonlinear mechanics setup.

*Kinematics.* Nonrigid beam deformation produces a one-dimensional strain that can be identified by the rotational and translational strain measures

$$\boldsymbol{\omega} = \mathbf{k}' - \boldsymbol{\Phi} \mathbf{k}, \quad \boldsymbol{\chi} = \mathbf{x}'_{,s} - \boldsymbol{\Phi} \mathbf{x}_{,s}, \quad (47)$$

also referred to as the angular and linear strain vectors, respectively. Here,  $(\cdot)_{,s}$  denotes a derivative with respect to the beam abscissa and the appended prime  $(\cdot)'$  distinguishes the current configuration from the reference one. The position vector  $\mathbf{x}$  and the orthogonal orientation tensor  $\boldsymbol{\alpha}$  of an orthonormal triad define the section configuration. Vector  $\mathbf{k}$  is the angular curvature defined by  $\boldsymbol{\alpha}^T \boldsymbol{\alpha}_{,s} = (\boldsymbol{\alpha}^T \mathbf{k}) \times$  as for (2)<sub>1</sub>, and  $\mathbf{x}_{,s}$  is the tangent vector to the beam axis. Tensor  $\boldsymbol{\Phi}$  is the section rotation from  $\boldsymbol{\alpha}$  to  $\boldsymbol{\alpha}' = \boldsymbol{\Phi} \boldsymbol{\alpha}$ . It can be seen (see (58)<sub>1</sub> and (62)<sub>1</sub>) that the expression of  $\boldsymbol{\omega}$  in (47) coincides with the angular strain induced by a variable rotation along the beam and defined by  $\boldsymbol{\Phi}^T \boldsymbol{\Phi}_{,s} = (\boldsymbol{\Phi}^T \boldsymbol{\omega}) \times$  as for (2)<sub>1</sub>. The *kinematical strain vectors*  $\boldsymbol{\omega}$  and  $\boldsymbol{\chi}$  in (47) represent differences of curvature and tangent vectors, that are made comparable thanks to a forward rotation of the reference values by  $\boldsymbol{\Phi}$ ; they vanish in the case of rigid deformation.

*Statics.* The differential equilibrium equations

$$\mathbf{M}'_{,s} + \mathbf{x}'_{,s} \times \mathbf{T}' + \mathbf{c}' = \mathbf{0}, \quad \mathbf{T}'_{,s} + \mathbf{f}' = \mathbf{0},$$

involve a one-dimensional stress state identified by the internal couple  $\mathbf{M}'$  and force  $\mathbf{T}'$  in the current configuration, also referred to as the angular and linear *stress resultants*. Current external loads  $\mathbf{c}'$  and  $\mathbf{f}'$  are couples and forces per unit length.

*Constitutive law.* Assuming a hyperelastic model, there exists a strain energy  $w(\boldsymbol{\beta}, \boldsymbol{\varepsilon})$ , a function of the

angular ( $\boldsymbol{\beta}$ ) and linear ( $\boldsymbol{\varepsilon}$ ) *strain parameters*, whose first derivatives with respect to the strain parameters define the conjugate angular and linear *stress parameters*,

$$\hat{\boldsymbol{M}} = w_{/\boldsymbol{\beta}}, \quad \hat{\boldsymbol{T}} = w_{/\boldsymbol{\varepsilon}}. \quad (48)$$

The second derivatives define the *elastic tensors*  $\hat{\boldsymbol{E}}_{\boldsymbol{\beta}\boldsymbol{\beta}} = w_{/\boldsymbol{\beta}\boldsymbol{\beta}}$ ,  $\hat{\boldsymbol{E}}_{\boldsymbol{\beta}\boldsymbol{\varepsilon}} = \hat{\boldsymbol{E}}_{\boldsymbol{\varepsilon}\boldsymbol{\beta}}^\top = w_{/\boldsymbol{\beta}\boldsymbol{\varepsilon}}$ , and  $\hat{\boldsymbol{E}}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = w_{/\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}$ , which allow us to write the tangent map that transforms strain-parameter variations into stress-parameter variations:

$$\begin{Bmatrix} \partial \hat{\boldsymbol{M}} \\ \partial \hat{\boldsymbol{T}} \end{Bmatrix} = \begin{bmatrix} \hat{\boldsymbol{E}}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \hat{\boldsymbol{E}}_{\boldsymbol{\beta}\boldsymbol{\varepsilon}} \\ \hat{\boldsymbol{E}}_{\boldsymbol{\varepsilon}\boldsymbol{\beta}} & \hat{\boldsymbol{E}}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} \end{bmatrix} \cdot \begin{Bmatrix} \partial \boldsymbol{\beta} \\ \partial \boldsymbol{\varepsilon} \end{Bmatrix}. \quad (49)$$

*Constitutive-to-mechanical variables connection.* The parameters governing the constitutive model are connected to the corresponding mechanical variables defined above by the relations (cf. [Ritto-Corrêa and Camotim 2002])

$$\begin{aligned} \boldsymbol{\beta} &= \boldsymbol{\Phi}^\top \boldsymbol{\omega}, & \hat{\boldsymbol{M}} &= \boldsymbol{\Phi}^\top \boldsymbol{M}', \\ \boldsymbol{\varepsilon} &= \boldsymbol{\Phi}^\top \boldsymbol{\chi}, & \hat{\boldsymbol{T}} &= \boldsymbol{\Phi}^\top \boldsymbol{T}'. \end{aligned} \quad (50)$$

Thus, the strain parameters are the back-rotated versions of the kinematical strains and the stress parameters are the back-rotated versions of the stress resultants. By analogy with three-dimensional elasticity, we may say that  $\boldsymbol{M}'$  and  $\boldsymbol{T}'$  are stress vectors of the first Piola–Kirchhoff kind, whereas  $\hat{\boldsymbol{M}}$  and  $\hat{\boldsymbol{T}}$  are stress vectors of the Biot kind.

The formulation summarized above matches the Reissner–Simo beam theory as described in [Ritto-Corrêa and Camotim 2002]. In that paper, the two representations of either strains or stresses are referred to as the spatial and the material representations, the latter being the back-rotated version of the former. Moreover, in that formulation (but not in the implementation, see [ibid., Section 5.2]) the cross-section in the material representation is oriented as the absolute reference frame, hence what they call section rotation corresponds to our section orientation ( $\boldsymbol{\alpha}$  in the reference configuration and  $\boldsymbol{\alpha}'$  in the current configuration), and the pull-back operation from the spatial to the material representation is performed by the orientation  $\boldsymbol{\alpha}'$ . This difference does not impair the equivalence of our formulation to theirs. However, it must be pointed out that tensors  $\bar{\boldsymbol{E}}_{\boldsymbol{\beta}\boldsymbol{\beta}}$ ,  $\bar{\boldsymbol{E}}_{\boldsymbol{\beta}\boldsymbol{\varepsilon}} = \bar{\boldsymbol{E}}_{\boldsymbol{\varepsilon}\boldsymbol{\beta}}^\top$ , and  $\bar{\boldsymbol{E}}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}$ , as built with the customary matrices of the section elastic properties, are meant to be defined in the absolute reference frame and must be rotated by  $\boldsymbol{\alpha}$  to build the elastic tensors in (49), that is,  $\hat{\boldsymbol{E}}_{\boldsymbol{\beta}\boldsymbol{\beta}} = \boldsymbol{\alpha} \bar{\boldsymbol{E}}_{\boldsymbol{\beta}\boldsymbol{\beta}} \boldsymbol{\alpha}^\top$ , etc.

The foregoing relations, together with the essential and natural boundary conditions, allow us to set up a variational functional  $\Pi_\delta = \int_s \pi_\delta ds$  and state the principle of virtual work for the beam as  $\Pi_\delta = 0$ . In nonlinear elasticity, we are usually concerned with the linearized form of the principle,  $\Pi_\delta + \partial \Pi_\delta = 0$ . In the forthcoming discussion, we focus on the internal work of stresses and address only the contributions  $\pi_{\text{int}\delta}$  and  $\partial \pi_{\text{int}\delta}$  to the virtual work per unit length; recalling the constitutive equations (48) and the relevant tangent map (49), they are given by

$$\begin{aligned} \pi_{\text{int}\delta} &= \begin{Bmatrix} \delta \boldsymbol{\beta} \\ \delta \boldsymbol{\varepsilon} \end{Bmatrix}^\top \cdot \begin{Bmatrix} \hat{\boldsymbol{M}} \\ \hat{\boldsymbol{T}} \end{Bmatrix}, \\ \partial \pi_{\text{int}\delta} &= \partial \pi_{\text{int}G\delta} + \partial \pi_{\text{int}E\delta} = \begin{Bmatrix} \partial \delta \boldsymbol{\beta} \\ \partial \delta \boldsymbol{\varepsilon} \end{Bmatrix}^\top \cdot \begin{Bmatrix} \hat{\boldsymbol{M}} \\ \hat{\boldsymbol{T}} \end{Bmatrix} + \begin{Bmatrix} \delta \boldsymbol{\beta} \\ \delta \boldsymbol{\varepsilon} \end{Bmatrix}^\top \cdot \begin{bmatrix} \hat{\boldsymbol{E}}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \hat{\boldsymbol{E}}_{\boldsymbol{\beta}\boldsymbol{\varepsilon}} \\ \hat{\boldsymbol{E}}_{\boldsymbol{\varepsilon}\boldsymbol{\beta}} & \hat{\boldsymbol{E}}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} \end{bmatrix} \cdot \begin{Bmatrix} \partial \boldsymbol{\beta} \\ \partial \boldsymbol{\varepsilon} \end{Bmatrix}. \end{aligned}$$

The incremental term  $\partial\pi_{\text{int}\delta}$  is made of a geometric contribution, that accounts for the current stress state, and an elastic contribution, that depends on the strain increments. Using (50), the contributions to the linearized internal virtual work are rewritten in terms of current mechanical variables as

$$\begin{aligned}\pi_{\text{int}\delta} &= \begin{Bmatrix} \Phi \delta(\Phi^T \omega) \\ \Phi \delta(\Phi^T \chi) \end{Bmatrix}^T \cdot \begin{Bmatrix} M' \\ T' \end{Bmatrix}, & \partial\pi_{\text{int}G\delta} &= \begin{Bmatrix} \Phi \partial\delta(\Phi^T \omega) \\ \Phi \partial\delta(\Phi^T \chi) \end{Bmatrix}^T \cdot \begin{Bmatrix} M' \\ T' \end{Bmatrix}, \\ \partial\pi_{\text{int}E\delta} &= \begin{Bmatrix} \Phi \delta(\Phi^T \omega) \\ \Phi \delta(\Phi^T \chi) \end{Bmatrix}^T \cdot \begin{bmatrix} \Phi \hat{E}_{\beta\beta} \Phi^T & \Phi \hat{E}_{\beta\epsilon} \Phi^T \\ \Phi \hat{E}_{\epsilon\beta} \Phi^T & \Phi \hat{E}_{\epsilon\epsilon} \Phi^T \end{bmatrix} \cdot \begin{Bmatrix} \Phi \partial(\Phi^T \omega) \\ \Phi \partial(\Phi^T \chi) \end{Bmatrix}.\end{aligned}\quad (51)$$

The development of the corotational variations of the kinematical strain that appear in (51) can be carried out in various ways. The proper choice mainly depends on the interpolating model one chooses to set up a beam element. Two different approaches are discussed in the next subsections.

**6.2. Vectorial parameterization of motion.** The first approach is more suitable for beam elements based on a linear interpolating model of the beam motion. According to (2), three variation variables ( $\varphi_\delta$ ,  $\varphi_\partial$ , and  $\varphi_{\partial\delta}$ ) characterize the virtual and incremental variations of the section rotation,

$$\begin{aligned}\Phi^T \delta\Phi &= (\Phi^T \varphi_\delta) \times, & \Phi^T \partial\Phi &= (\Phi^T \varphi_\partial) \times, \\ \Phi^T \partial\delta\Phi &= (\Phi^T \varphi_{\partial\delta}) \times + \frac{1}{2}((\Phi^T \varphi_\partial) \times (\Phi^T \varphi_\delta) \times + (\Phi^T \varphi_\delta) \times (\Phi^T \varphi_\partial) \times).\end{aligned}\quad (52)$$

Thus, recalling the properties (7) and (9), one immediately writes

$$\begin{aligned}\Phi \delta(\Phi^T \omega) &= \varphi_{\delta,s}, & \Phi \partial(\Phi^T \omega) &= \varphi_{\partial,s}, \\ \Phi \partial\delta(\Phi^T \omega) &= \varphi_{\partial\delta,s} - \frac{1}{2}(\varphi_\partial \times \varphi_{\delta,s} + \varphi_\delta \times \varphi_{\partial,s}).\end{aligned}\quad (53)$$

Moreover, recalling (47) and using (52), one easily obtains

$$\begin{aligned}\Phi \delta(\Phi^T \chi) &= \delta x'_{,s} - \varphi_\delta \times x'_{,s}, & \Phi \partial(\Phi^T \chi) &= \partial x'_{,s} - \varphi_\partial \times x'_{,s}, \\ \Phi \partial\delta(\Phi^T \chi) &= \partial\delta x'_{,s} - \varphi_{\partial\delta} \times x'_{,s} - (\varphi_\partial \times \delta x'_{,s} + \varphi_\delta \times \partial x'_{,s}) + \frac{1}{2}(\varphi_\partial \times \varphi_\delta \times + \varphi_\delta \times \varphi_\partial \times) x'_{,s}.\end{aligned}\quad (54)$$

The corotational variations of the kinematical strain vectors in (53) and (54) are expressed as functions of the differential rotation vectors  $\varphi_\delta$ ,  $\varphi_\partial$  (the spins, in the terminology of [Ritto-Corrêa and Camotim 2002]), and  $\varphi_{\partial\delta}$ ; their derivatives with respect to the beam abscissa; and the derivatives of the virtual and incremental variations  $\delta x'$ ,  $\partial x'$ , and  $\partial\delta x'$  of the current position vector—that is, the tangent vector virtual and incremental variations. The motivation for retaining, at this level, the mixed virtual-incremental variations  $\partial\delta x'$  will become clear later on.

Substituting (53) and (54) in (51), the virtual work contributions are rewritten in terms of the differential rotation and position vectors and their derivatives. In particular, contributions  $\pi_{\text{int}\delta}$  and  $\partial\pi_{\text{int}G\delta}$  take

the very neat expressions

$$\begin{aligned}\pi_{\text{int}\delta} &= \begin{Bmatrix} \boldsymbol{\varphi}_{\delta,s} \\ \delta \mathbf{x}'_{,s} \\ \boldsymbol{\varphi}_{\delta} \end{Bmatrix}^T \cdot \begin{Bmatrix} \mathbf{M}' \\ \mathbf{T}' \\ -\mathbf{x}'_{,s} \times \mathbf{T}' \end{Bmatrix}, \\ \partial \pi_{\text{intG}\delta} &= \begin{Bmatrix} \boldsymbol{\varphi}_{\partial\delta,s} \\ \partial \delta \mathbf{x}'_{,s} \\ \boldsymbol{\varphi}_{\partial\delta} \end{Bmatrix}^T \cdot \begin{Bmatrix} \mathbf{M}' \\ \mathbf{T}' \\ -\mathbf{x}'_{,s} \times \mathbf{T}' \end{Bmatrix} + \begin{Bmatrix} \boldsymbol{\varphi}_{\delta,s} \\ \delta \mathbf{x}'_{,s} \\ \boldsymbol{\varphi}_{\delta} \end{Bmatrix}^T \cdot \begin{bmatrix} \mathbf{0} & \mathbf{0} & \frac{1}{2} \mathbf{M}' \times^T \\ \mathbf{0} & \mathbf{0} & \mathbf{T}' \times^T \\ \frac{1}{2} \mathbf{M}' \times & \mathbf{T}' \times & (\mathbf{x}'_{,s} \times \mathbf{T}' \times)^S \end{bmatrix} \cdot \begin{Bmatrix} \boldsymbol{\varphi}_{\partial,s} \\ \partial \mathbf{x}'_{,s} \\ \boldsymbol{\varphi}_{\partial} \end{Bmatrix}.\end{aligned}\quad (55)$$

It is worth stressing that the virtual work contributions developed here through (55), are true and natural expressions in nonlinear beam variational mechanics, independently of any particular parameterization of the rotation one may choose to solve the elastic problem numerically. However, it is also worth noting that the multiplier  $\boldsymbol{\varphi}_{\partial\delta,s}$  is the derivative of a (second) differential rotation vector, not the characteristic differential rotation vector of the third differentiation of the rotation tensor.

Let us introduce now the vectorial parameterization of the rotation tensor defined in (10). Recalling the differential maps (36) and (37) and the derivative tensors defined in (32), the space-derivatives of the differential rotation vectors in (55) are expressed as functions of the variations of the rotation vector:

$$\begin{aligned}\boldsymbol{\varphi}_{\delta,s} &= \boldsymbol{\Gamma} \cdot \delta \boldsymbol{\varphi}_{,s} + (\boldsymbol{\varphi}_{,s} \cdot \boldsymbol{\Gamma}'^{\text{T}213}) \cdot \delta \boldsymbol{\varphi}, \\ \boldsymbol{\varphi}_{\partial,s} &= \boldsymbol{\Gamma} \cdot \partial \boldsymbol{\varphi}_{,s} + (\boldsymbol{\varphi}_{,s} \cdot \boldsymbol{\Gamma}'^{\text{T}213}) \cdot \partial \boldsymbol{\varphi}, \\ \boldsymbol{\varphi}_{\partial\delta,s} &= \boldsymbol{\Gamma} \cdot \partial \delta \boldsymbol{\varphi}_{,s} + (\boldsymbol{\varphi}_{,s} \cdot \boldsymbol{\Gamma}'^{\text{T}213}) \cdot \partial \delta \boldsymbol{\varphi} \\ &\quad + \boldsymbol{\Gamma}'^{\text{S}123} : (\partial \boldsymbol{\varphi}_{,s} \otimes \delta \boldsymbol{\varphi} + \partial \boldsymbol{\varphi} \otimes \delta \boldsymbol{\varphi}_{,s}) + (\boldsymbol{\varphi}_{,s} \cdot \boldsymbol{\Gamma}'^{\text{1234T}2134})^{\text{S}123} : \partial \boldsymbol{\varphi} \otimes \delta \boldsymbol{\varphi}.\end{aligned}\quad (56)$$

Then, the differential maps (36) and (37), the space-derivatives (56), and the relations (50) are introduced within (55). Note that the first term of the expression of  $\partial \pi_{\text{intG}\delta}$  splits into a term working for the mixed double variations ( $\partial\delta$ ) and a term working for the single virtual ( $\delta$ ) and incremental ( $\partial$ ) variations. The latter adds to the second term in the right-hand-side of (55)<sub>2</sub> and the final result follows in the form

$$\begin{aligned}\pi_{\text{int}\delta} &= \begin{Bmatrix} \delta \boldsymbol{\varphi}_{,s} \\ \delta \mathbf{x}'_{,s} \\ \delta \boldsymbol{\varphi} \end{Bmatrix}^T \cdot \begin{Bmatrix} \boldsymbol{\Gamma} \hat{\mathbf{M}} \\ \boldsymbol{\Phi} \hat{\mathbf{T}} \\ \boldsymbol{\varphi}_{,s} \cdot \boldsymbol{\Gamma}' \hat{\mathbf{M}} + \mathbf{x}'_{,s} \cdot \boldsymbol{\Phi}' \hat{\mathbf{T}} \end{Bmatrix}, \\ \partial \pi_{\text{intG}\delta} &= \begin{Bmatrix} \partial \delta \boldsymbol{\varphi}_{,s} \\ \partial \delta \mathbf{x}'_{,s} \\ \partial \delta \boldsymbol{\varphi} \end{Bmatrix}^T \cdot \begin{Bmatrix} \boldsymbol{\Gamma} \hat{\mathbf{M}} \\ \boldsymbol{\Phi} \hat{\mathbf{T}} \\ \boldsymbol{\varphi}_{,s} \cdot \boldsymbol{\Gamma}' \hat{\mathbf{M}} + \mathbf{x}'_{,s} \cdot \boldsymbol{\Phi}' \hat{\mathbf{T}} \end{Bmatrix} \\ &\quad + \begin{Bmatrix} \delta \boldsymbol{\varphi}_{,s} \\ \delta \mathbf{x}'_{,s} \\ \delta \boldsymbol{\varphi} \end{Bmatrix}^T \cdot \begin{bmatrix} \mathbf{0} & \mathbf{0} & \boldsymbol{\Gamma}' \hat{\mathbf{M}} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Phi}' \hat{\mathbf{T}} \\ (\boldsymbol{\Gamma}' \hat{\mathbf{M}})^T & (\boldsymbol{\Phi}' \hat{\mathbf{T}})^T & \boldsymbol{\varphi}_{,s} \cdot \boldsymbol{\Gamma}'^{\text{1234}} \hat{\mathbf{M}} + \mathbf{x}'_{,s} \cdot \boldsymbol{\Phi}'^{\text{1234}} \hat{\mathbf{T}} \end{bmatrix} \cdot \begin{Bmatrix} \partial \boldsymbol{\varphi}_{,s} \\ \partial \mathbf{x}'_{,s} \\ \partial \boldsymbol{\varphi} \end{Bmatrix}.\end{aligned}\quad (57)$$

Derivation of (57) is carried out with the help of the properties (40), (42)–(46), and (A.16)<sub>2</sub>.

It is seen that (57)<sub>1</sub> and the second term in (57)<sub>2</sub> coincide respectively with the integrands found in [Ritto-Corrêa and Camotim 2002, Equations (36) and (40)]. The first term of  $\partial \pi_{\text{intG}\delta}$  is lacking in that paper, as configuration-independent interpolations of both the displacement and the rotation are assumed

there, so the mixed double variations hold null in their formulation. This is not the case in our design of a nonlinear beam element, as it will be shown in the next section.

The foregoing development of the kinematical strain corotational variations is suitable for beam elements relying on a linear interpolation of the parameters of the section motion. This was a common approach in the nineties [Cardona and Géradin 1988; Ibrahimbegović 1995; Ibrahimbegović et al. 1995]. The beam element developed by Ritto-Corrêa and Camotim [2002] relies on the interpolation of the total position and rotation vectors as for  $\mathbf{x} = \sum_{J=1}^N W_J \mathbf{x}_J$  and  $\boldsymbol{\varphi} = \sum_{J=1}^N W_J \boldsymbol{\varphi}_J$  (whence  $\mathbf{x}_{,s} = \sum_{J=1}^N W_{J,s} \mathbf{x}_J$  and  $\boldsymbol{\varphi}_{,s} = \sum_{J=1}^N W_{J,s} \boldsymbol{\varphi}_J$ ). These very simple formulae are linear with the nodal unknowns  $\mathbf{x}_J$  and  $\boldsymbol{\varphi}_J$ , and substituting them in (57) leads to a linear beam element. However, as Ritto-Corrêa and Camotim themselves point out, that element was not frame invariant.

**6.3. Consistent account of the rotation variations.** The second approach is more suitable for a nonlinear interpolating model of the section orientation and allows us to build nonlinear beam elements that are frame invariant and path independent at the same time. Equation (2) is now written for all combinations of the virtual, incremental, and spatial independent variations of the section rotation; they provide seven equations, those in (52) and, in addition,

$$\begin{aligned}
\boldsymbol{\Phi}^T \boldsymbol{\Phi}_{,s} &= (\boldsymbol{\Phi}^T \boldsymbol{\omega}) \times, \\
\boldsymbol{\Phi}^T \delta \boldsymbol{\Phi}_{,s} &= (\boldsymbol{\Phi}^T \boldsymbol{\omega}_\delta) \times - \frac{1}{2} (\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\delta \times \boldsymbol{\omega}) \times + (\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\delta) \times (\boldsymbol{\Phi}^T \boldsymbol{\omega}) \times, \\
\boldsymbol{\Phi}^T \partial \boldsymbol{\Phi}_{,s} &= (\boldsymbol{\Phi}^T \boldsymbol{\omega}_\partial) \times - \frac{1}{2} (\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\partial \times \boldsymbol{\omega}) \times + (\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\partial) \times (\boldsymbol{\Phi}^T \boldsymbol{\omega}) \times, \\
\boldsymbol{\Phi}^T \partial \delta \boldsymbol{\Phi}_{,s} &= (\boldsymbol{\Phi}^T \boldsymbol{\omega}_{\partial\delta}) \times - \frac{1}{2} (\boldsymbol{\Phi}^T \boldsymbol{\varphi}_{\partial\delta} \times \boldsymbol{\omega}) \times + (\boldsymbol{\Phi}^T \boldsymbol{\varphi}_{\partial\delta}) \times (\boldsymbol{\Phi}^T \boldsymbol{\omega}) \times \\
&\quad - \frac{1}{2} (\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\partial \times \boldsymbol{\omega}_\delta + \boldsymbol{\Phi}^T \boldsymbol{\varphi}_\delta \times \boldsymbol{\omega}_\partial) \times + (\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\partial) \times (\boldsymbol{\Phi}^T \boldsymbol{\omega}_\delta) \times + (\boldsymbol{\Phi}^T \boldsymbol{\varphi}_\delta) \times (\boldsymbol{\Phi}^T \boldsymbol{\omega}_\partial) \times.
\end{aligned} \tag{58}$$

Equations (52) and (58) actually define seven variation variables: the virtual, incremental, and mixed virtual-incremental rotation vectors  $\boldsymbol{\varphi}_\delta$ ,  $\boldsymbol{\varphi}_\partial$ , and  $\boldsymbol{\varphi}_{\partial\delta}$ , the finite angular strain vector  $\boldsymbol{\omega}$ , and the virtual, incremental, and mixed virtual-incremental angular strain vectors  $\boldsymbol{\omega}_\delta$ ,  $\boldsymbol{\omega}_\partial$ , and  $\boldsymbol{\omega}_{\partial\delta}$ . Using (6), the angular strain corotational variations are given the expressions

$$\begin{aligned}
\boldsymbol{\Phi} \delta(\boldsymbol{\Phi}^T \boldsymbol{\omega}) &= \boldsymbol{\omega}_\delta - \frac{1}{2} \boldsymbol{\varphi}_\delta \times \boldsymbol{\omega}, & \boldsymbol{\Phi} \partial(\boldsymbol{\Phi}^T \boldsymbol{\omega}) &= \boldsymbol{\omega}_\partial - \frac{1}{2} \boldsymbol{\varphi}_\partial \times \boldsymbol{\omega}, \\
\boldsymbol{\Phi} \partial \delta(\boldsymbol{\Phi}^T \boldsymbol{\omega}) &= \boldsymbol{\omega}_{\partial\delta} - \frac{1}{2} \boldsymbol{\varphi}_{\partial\delta} \times \boldsymbol{\omega} - \frac{1}{2} (\boldsymbol{\varphi}_\partial \times \boldsymbol{\omega}_\delta + \boldsymbol{\varphi}_\delta \times \boldsymbol{\omega}_\partial) + (\boldsymbol{\varphi}_\partial \otimes \boldsymbol{\varphi}_\delta)^S \cdot \boldsymbol{\omega}.
\end{aligned} \tag{59}$$

Note that, in (58) and (59),  $\boldsymbol{\omega}_{\partial\delta}$  is a *third* characteristic differential rotation vector as defined in (2)<sub>3</sub>.

In view of the interpolating model that will be introduced in the next section, we also address the variations of the section orientation in the current configuration,  $\boldsymbol{\alpha}' = \boldsymbol{\Phi} \boldsymbol{\alpha}$ . Since the orientations are assumed to be orthogonal tensors, we may write, as for (52) and (58),

$$\begin{aligned}
\boldsymbol{\alpha}'^T \delta \boldsymbol{\alpha}' &= (\boldsymbol{\alpha}'^T \boldsymbol{\varphi}_\delta) \times, & \boldsymbol{\alpha}'^T \partial \boldsymbol{\alpha}' &= (\boldsymbol{\alpha}'^T \boldsymbol{\varphi}_\partial) \times, \\
\boldsymbol{\alpha}'^T \partial \delta \boldsymbol{\alpha}' &= (\boldsymbol{\alpha}'^T \boldsymbol{\varphi}_{\partial\delta}) \times + \frac{1}{2} ((\boldsymbol{\alpha}'^T \boldsymbol{\varphi}_\partial) \times (\boldsymbol{\alpha}'^T \boldsymbol{\varphi}_\delta) \times + (\boldsymbol{\alpha}'^T \boldsymbol{\varphi}_\delta) \times (\boldsymbol{\alpha}'^T \boldsymbol{\varphi}_\partial) \times),
\end{aligned} \tag{60}$$



and

$$\begin{aligned}
\boldsymbol{\alpha}^T \boldsymbol{\alpha}'_{,s} &= (\boldsymbol{\alpha}^T \mathbf{k}') \times, \\
\boldsymbol{\alpha}^T \delta \boldsymbol{\alpha}'_{,s} &= (\boldsymbol{\alpha}^T \mathbf{k}'_{\delta}) \times - \frac{1}{2} (\boldsymbol{\alpha}^T \boldsymbol{\varphi}_{\delta} \times \mathbf{k}') \times + (\boldsymbol{\alpha}^T \boldsymbol{\varphi}_{\delta}) \times (\boldsymbol{\alpha}^T \mathbf{k}') \times, \\
\boldsymbol{\alpha}^T \partial \boldsymbol{\alpha}'_{,s} &= (\boldsymbol{\alpha}^T \mathbf{k}'_{\partial}) \times - \frac{1}{2} (\boldsymbol{\alpha}^T \boldsymbol{\varphi}_{\partial} \times \mathbf{k}') \times + (\boldsymbol{\alpha}^T \boldsymbol{\varphi}_{\partial}) \times (\boldsymbol{\alpha}^T \mathbf{k}') \times, \\
\boldsymbol{\alpha}^T \partial \delta \boldsymbol{\alpha}'_{,s} &= (\boldsymbol{\alpha}^T \mathbf{k}'_{\partial\delta}) \times - \frac{1}{2} (\boldsymbol{\alpha}^T \boldsymbol{\varphi}_{\partial\delta} \times \mathbf{k}') \times + (\boldsymbol{\alpha}^T \boldsymbol{\varphi}_{\partial\delta}) \times (\boldsymbol{\alpha}^T \mathbf{k}') \times \\
&\quad - \frac{1}{2} (\boldsymbol{\alpha}^T \boldsymbol{\varphi}_{\partial} \times \mathbf{k}'_{\delta} + \boldsymbol{\alpha}^T \boldsymbol{\varphi}_{\delta} \times \mathbf{k}'_{\partial}) \times + (\boldsymbol{\alpha}^T \boldsymbol{\varphi}_{\partial}) \times (\boldsymbol{\alpha}^T \mathbf{k}'_{\delta}) \times + (\boldsymbol{\alpha}^T \boldsymbol{\varphi}_{\delta}) \times (\boldsymbol{\alpha}^T \mathbf{k}'_{\partial}) \times.
\end{aligned} \tag{61}$$

The differential vectors that characterize the virtual and incremental orientation variations in (60) are easily seen to coincide with vectors  $\boldsymbol{\varphi}_{\delta}$ ,  $\boldsymbol{\varphi}_{\partial}$ , and  $\boldsymbol{\varphi}_{\partial\delta}$ . Equation (61), instead, defines the current curvature vector  $\mathbf{k}'$  and three differential curvature vectors,  $\mathbf{k}'_{\delta}$ ,  $\mathbf{k}'_{\partial}$ , and  $\mathbf{k}'_{\partial\delta}$  (again, a *third* characteristic differential rotation vector).

The finite and differential strain vectors can be linked to the finite and differential current curvature vectors by differentiating the relation  $\boldsymbol{\alpha}' = \boldsymbol{\Phi} \boldsymbol{\alpha}$  and recalling the definition of the reference curvature,  $\boldsymbol{\alpha}^T \boldsymbol{\alpha}_{,s} = (\boldsymbol{\alpha}^T \mathbf{k}) \times$ . This yields

$$\begin{aligned}
\boldsymbol{\omega} &= \mathbf{k}' - \boldsymbol{\Phi} \mathbf{k}, \\
\boldsymbol{\omega}_{\delta} &= \mathbf{k}'_{\delta} - \frac{1}{2} \boldsymbol{\varphi}_{\delta} \times \boldsymbol{\Phi} \mathbf{k}, \\
\boldsymbol{\omega}_{\partial} &= \mathbf{k}'_{\partial} - \frac{1}{2} \boldsymbol{\varphi}_{\partial} \times \boldsymbol{\Phi} \mathbf{k}, \\
\boldsymbol{\omega}_{\partial\delta} &= \mathbf{k}'_{\partial\delta} - \frac{1}{2} \boldsymbol{\varphi}_{\partial\delta} \times \boldsymbol{\Phi} \mathbf{k} + \frac{1}{2} (\boldsymbol{\varphi}_{\partial} \otimes \boldsymbol{\varphi}_{\delta} + \boldsymbol{\varphi}_{\partial} \cdot \boldsymbol{\varphi}_{\delta} \otimes \mathbf{I})^S \cdot \boldsymbol{\Phi} \mathbf{k}.
\end{aligned} \tag{62}$$

Finally, using (62) within (59), the angular strain corotational variations are brought to the following expressions, which coincide with the curvature corotational variations:

$$\begin{aligned}
\boldsymbol{\Phi} \delta(\boldsymbol{\Phi}^T \boldsymbol{\omega}) &= \boldsymbol{\alpha}' \delta(\boldsymbol{\alpha}^T \mathbf{k}') = \mathbf{k}'_{\delta} - \frac{1}{2} \boldsymbol{\varphi}_{\delta} \times \mathbf{k}', \\
\boldsymbol{\Phi} \partial(\boldsymbol{\Phi}^T \boldsymbol{\omega}) &= \boldsymbol{\alpha}' \partial(\boldsymbol{\alpha}^T \mathbf{k}') = \mathbf{k}'_{\partial} - \frac{1}{2} \boldsymbol{\varphi}_{\partial} \times \mathbf{k}', \\
\boldsymbol{\Phi} \partial \delta(\boldsymbol{\Phi}^T \boldsymbol{\omega}) &= \boldsymbol{\alpha}' \partial \delta(\boldsymbol{\alpha}^T \mathbf{k}') = \mathbf{k}'_{\partial\delta} - \frac{1}{2} \boldsymbol{\varphi}_{\partial\delta} \times \mathbf{k}' - \frac{1}{2} (\boldsymbol{\varphi}_{\partial} \times \mathbf{k}'_{\delta} + \boldsymbol{\varphi}_{\delta} \times \mathbf{k}'_{\partial}) + (\boldsymbol{\varphi}_{\partial} \otimes \boldsymbol{\varphi}_{\delta})^S \cdot \mathbf{k}'.
\end{aligned} \tag{63}$$

The linear strain corotational variations are much easier to write and are seen to coincide with the tangent vector corotational variations. From (47)<sub>2</sub>, recalling (52), one obtains

$$\begin{aligned}
\boldsymbol{\Phi} \delta(\boldsymbol{\Phi}^T \boldsymbol{\chi}) &= \boldsymbol{\alpha}' \delta(\boldsymbol{\alpha}^T \mathbf{x}'_{,s}) = \delta \mathbf{x}'_{,s} - \boldsymbol{\varphi}_{\delta} \times \mathbf{x}'_{,s}, \\
\boldsymbol{\Phi} \partial(\boldsymbol{\Phi}^T \boldsymbol{\chi}) &= \boldsymbol{\alpha}' \partial(\boldsymbol{\alpha}^T \mathbf{x}'_{,s}) = \partial \mathbf{x}'_{,s} - \boldsymbol{\varphi}_{\partial} \times \mathbf{x}'_{,s}, \\
\boldsymbol{\Phi} \partial \delta(\boldsymbol{\Phi}^T \boldsymbol{\chi}) &= \boldsymbol{\alpha}' \partial \delta(\boldsymbol{\alpha}^T \mathbf{x}'_{,s}) = \partial \delta \mathbf{x}'_{,s} - \boldsymbol{\varphi}_{\partial\delta} \times \mathbf{x}'_{,s} - (\boldsymbol{\varphi}_{\partial} \times \delta \mathbf{x}'_{,s} + \boldsymbol{\varphi}_{\delta} \times \partial \mathbf{x}'_{,s}) + (\boldsymbol{\varphi}_{\partial} \times \boldsymbol{\varphi}_{\delta} \times)^S \mathbf{x}'_{,s},
\end{aligned} \tag{64}$$

which also coincides with (54).

The corotational variations of the kinematical strain vectors are expressed in (63) and (64) as functions of true characteristic differential vectors: the differential rotation vectors  $\boldsymbol{\varphi}_{\delta}$ ,  $\boldsymbol{\varphi}_{\partial}$ , and  $\boldsymbol{\varphi}_{\partial\delta}$ , the differential curvature vectors  $\mathbf{k}'_{\delta}$ ,  $\mathbf{k}'_{\partial}$ , and  $\mathbf{k}'_{\partial\delta}$ , and the differential tangent vectors  $\delta \mathbf{x}'_{,s}$ ,  $\partial \mathbf{x}'_{,s}$ , and  $\partial \delta \mathbf{x}'_{,s}$ . The interested reader could substitute (63) and (64) within (51) to write expressions for  $\pi_{\text{int}\delta}$  and  $\partial \pi_{\text{int}\delta}$  that are the counterparts of (55) in terms of true characteristic differential vectors. However, let us address now the issue of a consistent interpolation along a beam element.

## 7. Slender beam element

Ensuring path-independence and frame-invariance is an important asset for an interpolating mechanism in finite element approximations. In nonlinear problems, the interpolation of the total rotation vector on a Euclidean vector space is path independent but not frame invariant [Crisfield and Jelenić 1999]. Interpolating either the incremental or the iterative rotation vector impairs path-independence, however Ibrahimbegović and Taylor [2002] proved that it may achieve frame-invariance. These issues were remarked on by Ritto-Corrêa and Camotim [2002] and were also discussed in [Merlini and Morandini 2004b] with reference to multicoordinate domains. It was ascertained that interpolating the orientations instead of their motions is the key to ensuring path-independence, and that averaging relative orientations in a given configuration is the key to ensuring frame-invariance. An interpolation scheme for a beam element in line with these concepts and ensuring frame-invariance and path-independence was proposed by Jelenić and Crisfield [1999].

In Section 7.1 we discuss the basic features that an interpolating model fit for rotational kinematics should conform to, and in Section 7.2 we outline our general interpolation scheme that ensures path-independence and frame-invariance.

**7.1. Interpolating model.** A consistent interpolation for a beam element should provide a section orientation (function of the beam abscissa  $s$ ) consistent with the special manifold the rotations belong to. Accordingly, any interpolated orientation  $\alpha$  in the reference configuration must be allowed to be regarded as the result of a relative rotation  $\tilde{\Phi}_J(s)$  from the orientation  $\alpha_J$  of whichever node  $J$  (with  $J = 1, 2, \dots, N$  and  $N$  the number of nodes). The same must be true for the current configuration, so multiplicative expressions for the interpolated orientations,

$$\alpha = \tilde{\Phi}_J \alpha_J \quad \text{and} \quad \alpha' = \tilde{\Phi}'_J \alpha'_J, \quad (65)$$

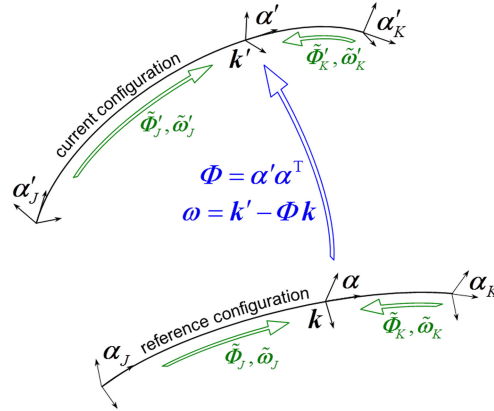
must hold. In (65) the relative rotations are transcendental functions  $\tilde{\Phi}_J(s)$  and  $\tilde{\Phi}'_J(s)$  of the beam abscissa. The derivatives of these functions along the beam, namely  $\tilde{\Phi}_J^T \tilde{\Phi}_{J,s} = (\tilde{\Phi}_J^T \tilde{\omega}_J) \times$  and  $\tilde{\Phi}_J'^T \tilde{\Phi}'_{J,s} = (\tilde{\Phi}_J'^T \tilde{\omega}'_J) \times$ , define relative angular strains  $\tilde{\omega}_J(s)$  and  $\tilde{\omega}'_J(s)$ . The interpolated curvatures  $k$  and  $k'$  are immediately found to coincide with these relative strains:

$$k = \tilde{\omega}_J \quad \text{and} \quad k' = \tilde{\omega}'_J. \quad (66)$$

An interpolation scheme provides the relative-rotation functions  $\tilde{\Phi}_J(s)$  and  $\tilde{\Phi}'_J(s)$  and the characteristic vectors  $\tilde{\omega}_J(s)$  and  $\tilde{\omega}'_J(s)$  of their derivatives. Once these quantities become known and orientations and curvatures are obtained in both the reference and the current configurations, the section rotation and the angular strain can immediately be recovered as  $\Phi = \alpha' \alpha^T$  and  $\omega = k' - \Phi k$  (Figure 2). This way, the correct equivalent of an isoparametric element for rotational kinematics is achieved, in the sense that the reference and the current orientation fields (not the rotation field) are approximated using the same interpolating model [Merlini and Morandini 2011b].

The interpolating model is completed with a linearization procedure that yields the appropriate variation variables (virtual, incremental, and mixed virtual-incremental variations). In the current configuration, the reference nodal orientations  $\alpha_J$  are rotated by the unknowns  $\Phi_J$ :

$$\alpha'_J = \Phi_J \alpha_J. \quad (67)$$



**Figure 2.** Interpolating model: recovering rotations and angular strains.

The interpolated orientation is a multiplicative combination of two successive rotations,  $\Phi_J$  in (67) followed by  $\tilde{\Phi}'_J$  in (65)<sub>2</sub>, resulting in  $\alpha' = \tilde{\Phi}'_J \Phi_J \alpha_J$ . A careful linearization of the compound rotation  $\tilde{\Phi}'_J \Phi_J$  is required in order to derive consistent orientation variations.

According to (2), the variations of either  $\Phi_J$  or  $\tilde{\Phi}'_J$  are characterized by specific differential vectors. The virtual, incremental and mixed virtual-incremental variations of the discrete rotation  $\Phi_J$  (or alternatively the discrete orientation  $\alpha'_J$ ) define, as usual, the differential rotation vectors  $\varphi_{J\delta}$ ,  $\varphi_{J\partial}$ , and  $\varphi_{J\partial\delta}$ . The relevant formulae are not displayed here, but are easily written from (52) (or alternatively (60)), simply by replacing  $\Phi$  with  $\Phi_J$  (or  $\alpha'$  with  $\alpha'_J$ ). In a similar way, variations of the relative-rotation field  $\tilde{\Phi}'_J$  define the virtual, incremental, and mixed virtual-incremental relative rotation vectors  $\tilde{\varphi}'_{J\delta}$ ,  $\tilde{\varphi}'_{J\partial}$ , and  $\tilde{\varphi}'_{J\partial\delta}$  and relative angular differential strains  $\tilde{\omega}'_{J\delta}$ ,  $\tilde{\omega}'_{J\partial}$ , and  $\tilde{\omega}'_{J\partial\delta}$ . The relevant formulae can be written from (52) and (58) after replacing  $\Phi$  with  $\tilde{\Phi}'_J$ .

Differentiating  $\alpha' = \tilde{\Phi}'_J \Phi_J \alpha_J$  yields a relationship between the local variation variables ( $\varphi_\delta$ ,  $\varphi_\partial$ ,  $\varphi_{\partial\delta}$ ,  $k'_\delta$ ,  $k'_\partial$ ,  $k'_{\partial\delta}$ ), the relative variation variables ( $\tilde{\varphi}'_{J\delta}$ ,  $\tilde{\varphi}'_{J\partial}$ ,  $\tilde{\varphi}'_{J\partial\delta}$ ,  $\tilde{\omega}'_{J\delta}$ ,  $\tilde{\omega}'_{J\partial}$ ,  $\tilde{\omega}'_{J\partial\delta}$ ), and the nodal variation variables ( $\varphi_{J\delta}$ ,  $\varphi_{J\partial}$ ,  $\varphi_{J\partial\delta}$ ). After cumbersome algebraic manipulations (detailed in [Merlini 2002]), one obtains

$$\begin{aligned} \varphi_\delta &= \tilde{\varphi}'_{J\delta} + \tilde{\Phi}'_J \varphi_{J\delta}, \\ \varphi_\partial &= \tilde{\varphi}'_{J\partial} + \tilde{\Phi}'_J \varphi_{J\partial}, \\ \varphi_{\partial\delta} &= \tilde{\varphi}'_{J\partial\delta} + \tilde{\Phi}'_J \varphi_{J\partial\delta} - \frac{1}{2} ((\tilde{\Phi}'_J \varphi_{J\partial}) \times \tilde{\varphi}'_{J\delta} + (\tilde{\Phi}'_J \varphi_{J\delta}) \times \tilde{\varphi}'_{J\partial}), \end{aligned} \quad (68)$$

and

$$\begin{aligned} k'_\delta &= \tilde{\omega}'_{J\delta} - \frac{1}{2} (\tilde{\Phi}'_J \varphi_{J\delta}) \times \tilde{\omega}'_J, \\ k'_\partial &= \tilde{\omega}'_{J\partial} - \frac{1}{2} (\tilde{\Phi}'_J \varphi_{J\partial}) \times \tilde{\omega}'_J, \\ k'_{\partial\delta} &= \tilde{\omega}'_{J\partial\delta} - \frac{1}{2} (\tilde{\Phi}'_J \varphi_{J\partial\delta}) \times \tilde{\omega}'_J - \frac{1}{2} ((\tilde{\Phi}'_J \varphi_{J\partial}) \times \tilde{\omega}'_{J\delta} + (\tilde{\Phi}'_J \varphi_{J\delta}) \times \tilde{\omega}'_{J\partial}) \\ &\quad + \frac{1}{4} ((\tilde{\Phi}'_J \varphi_{J\partial}) \times (\tilde{\varphi}'_{J\delta} + \tilde{\Phi}'_J \varphi_{J\delta}) \times + (\tilde{\Phi}'_J \varphi_{J\delta}) \times (\tilde{\varphi}'_{J\partial} + \tilde{\Phi}'_J \varphi_{J\partial}) \times) \tilde{\omega}'_J \\ &\quad - (\tilde{\Phi}'_J \varphi_{J\partial} \otimes \tilde{\varphi}'_{J\delta} + \tilde{\Phi}'_J \varphi_{J\delta} \otimes \tilde{\varphi}'_{J\partial} + \tilde{\Phi}'_J \varphi_{J\partial} \otimes \tilde{\Phi}'_J \varphi_{J\delta})^S \cdot \tilde{\omega}'_J. \end{aligned} \quad (69)$$

The reader should notice the relevance of considering the third characteristic differential vector defined in (2)<sub>3</sub>, which is present here as  $k'_{\partial\delta}$  and  $\tilde{\omega}'_{J\partial\delta}$ .

**7.2. A multiplicative interpolation scheme.** It is worth noting that the findings in the previous subsection are perfectly independent of any particular parameterization of the rotation tensor and of any particular interpolation scheme. They are just the natural premise for an interpolation to be consistent with the traits of the manifold the rotations belong to. In this subsection instead, we introduce the algorithm we use to implement a frame-invariant and path-independent multiplicative interpolation scheme for rotational kinematics.

The idea underlying the interpolation scheme is quite simple. The sought orientation is a weighted average between the nodal orientations, in the sense that the relative orientations with respect to each node are weighted in conformity with the position of the section within the beam element. This concept is achieved by satisfying a condition of multiplicative nature,

$$\sum_{J=1}^N W_J \ln(\boldsymbol{\alpha}\boldsymbol{\alpha}_J^T) = \mathbf{0},$$

where  $W_J(s)$  are (normalized) weight functions and the skew-symmetric tensors  $\ln(\boldsymbol{\alpha}\boldsymbol{\alpha}_J^T) = \tilde{\boldsymbol{\varphi}}_J \times$  (recall (65)<sub>1</sub>) are built with the relative rotation vectors from each node. Note that this condition is the exact transposition of the condition of additive nature  $\sum_{J=1}^N W_J(\mathbf{x} - \mathbf{x}_J) = \mathbf{0}$ , which underlies the commonly used interpolation on the Euclidean position-vector space and simply results in  $\mathbf{x} = \sum_{J=1}^N W_J \mathbf{x}_J$ . In contrast, the interpolation condition for rotational kinematics is a nonlinear implicit equation that cannot be solved in general in closed form, and therefore requires an iterative procedure. This interpolation scheme, together with a fast algorithm to solve the nonlinear condition, has been proposed in [Merlini and Morandini 2004b] with reference to multicoordinate domains.

In the current configuration, the interpolation algorithm reads

$$\sum_{J=1}^N W_J \ln(\boldsymbol{\alpha}'\boldsymbol{\alpha}'_J{}^T) = \mathbf{0}, \quad \sum_{J=1}^N W_J(\mathbf{x}' - \mathbf{x}'_J) = \mathbf{0}. \quad (70)$$

The linearization of the multiplicative condition (70)<sub>1</sub> is a quite subtle task. A short account is outlined below, but any details can be recovered from [Merlini and Morandini 2004b] and references therein. The reader should notice that so far, in the whole formulation aimed at building a beam element, we have not yet referred to any parameterization of the rotation tensor. Only now is it time to introduce a parameterization. We resort to the natural vectorial parameterization (10) and its linearization discussed in Section 5, where the differential maps (36)–(38) are obtained. Recalling (65)<sub>2</sub> and the exponential map  $\tilde{\boldsymbol{\Phi}}'_J = \exp(\tilde{\boldsymbol{\varphi}}'_J \times)$ , (70)<sub>1</sub> can be rewritten as a vectorial equation as

$$\sum_{J=1}^N W_J \tilde{\boldsymbol{\varphi}}'_J = \mathbf{0}. \quad (71)$$

Differentiating (71) up to third order yields seven algebraic equations for the unknowns  $\delta\tilde{\boldsymbol{\varphi}}'_J$ ,  $\partial\tilde{\boldsymbol{\varphi}}'_J$ ,  $\partial\delta\tilde{\boldsymbol{\varphi}}'_J$ ,  $\tilde{\boldsymbol{\varphi}}'_{J,s}$ ,  $\delta\tilde{\boldsymbol{\varphi}}'_{J,s}$ ,  $\partial\tilde{\boldsymbol{\varphi}}'_{J,s}$ , and  $\partial\delta\tilde{\boldsymbol{\varphi}}'_{J,s}$ . After inverting (36)–(38) for  $\delta\boldsymbol{\varphi}$ ,  $\partial\delta\boldsymbol{\varphi}$ , and  $\partial\delta\delta\boldsymbol{\varphi}$  and using the results within the seven equations, the unknowns of the equation set are turned to the relative variation variables  $\tilde{\boldsymbol{\varphi}}'_{J\delta}$ ,  $\tilde{\boldsymbol{\varphi}}'_{J\partial}$ ,  $\tilde{\boldsymbol{\varphi}}'_{J\partial\delta}$ ,  $\tilde{\boldsymbol{\omega}}'_J$ ,  $\tilde{\boldsymbol{\omega}}'_{J\delta}$ ,  $\tilde{\boldsymbol{\omega}}'_{J\partial}$ , and  $\tilde{\boldsymbol{\omega}}'_{J\partial\delta}$ . Finally, after solving (68), (66)<sub>2</sub>, and (69) for  $\tilde{\boldsymbol{\varphi}}'_{J\delta}$ ,  $\tilde{\boldsymbol{\varphi}}'_{J\partial}$ ,  $\tilde{\boldsymbol{\varphi}}'_{J\partial\delta}$ ,  $\tilde{\boldsymbol{\omega}}'_J$ ,  $\tilde{\boldsymbol{\omega}}'_{J\delta}$ ,  $\tilde{\boldsymbol{\omega}}'_{J\partial}$ , and  $\tilde{\boldsymbol{\omega}}'_{J\partial\delta}$  and substituting the results within the seven equations, the unknowns are changed to the local variation variables  $\boldsymbol{\varphi}_\delta$ ,  $\boldsymbol{\varphi}_\partial$ ,  $\boldsymbol{\varphi}_{\partial\delta}$ ,  $\mathbf{k}'$ ,  $\mathbf{k}'_\delta$ ,  $\mathbf{k}'_\partial$ , and  $\mathbf{k}'_{\partial\delta}$ .

The equations are solved in cascade and yield an expression for the curvature,

$$\mathbf{k}' = -\left(\sum_{K=1}^N W_K \tilde{\Gamma}'_{K^{-1}}\right)^{-1} \sum_{J=1}^N W_{J,s} \tilde{\varphi}'_J, \quad (72)$$

and six expressions for the local variation variables,

$$\begin{aligned} \boldsymbol{\varphi}_\delta &= \sum_{J=1}^N \tilde{Y}'_J \cdot \boldsymbol{\varphi}_{J\delta}, & \boldsymbol{\varphi}_\partial &= \sum_{K=1}^N \tilde{Y}'_K \cdot \boldsymbol{\varphi}_{K\partial}, \\ \boldsymbol{\varphi}_{\partial\delta} &= \sum_{J=1}^N \tilde{Y}'_J \cdot \boldsymbol{\varphi}_{J\partial\delta} + \sum_{J=1}^N \sum_{K=1}^N \tilde{y}'_{JK} : \boldsymbol{\varphi}_{J\delta} \otimes \boldsymbol{\varphi}_{K\partial}, \end{aligned} \quad (73)$$

and

$$\begin{aligned} \mathbf{k}'_\delta &= \sum_{J=1}^N \tilde{\mathcal{X}}'_J : \boldsymbol{\varphi}_{J\delta} \otimes \mathbf{I}, & \mathbf{k}'_\partial &= \sum_{K=1}^N \tilde{\mathcal{X}}'_K : \boldsymbol{\varphi}_{K\partial} \otimes \mathbf{I}, \\ \mathbf{k}'_{\partial\delta} &= \sum_{J=1}^N \tilde{\mathcal{X}}'_J : \boldsymbol{\varphi}_{J\partial\delta} \otimes \mathbf{I} + \sum_{J=1}^N \sum_{K=1}^N \tilde{\mathcal{Z}}'_{JK} : \boldsymbol{\varphi}_{J\delta} \otimes \boldsymbol{\varphi}_{K\partial} \otimes \mathbf{I}. \end{aligned} \quad (74)$$

Equations (73) and (74) are interpolating functions, linear in the nodal variation variables  $\boldsymbol{\varphi}_{J\delta}$ ,  $\boldsymbol{\varphi}_{J\partial}$ , and  $\boldsymbol{\varphi}_{J\partial\delta}$ . The tensor-valued coefficients  $\tilde{Y}'_J$ ,  $\tilde{y}'_{JK}$ ,  $\tilde{\mathcal{X}}'_J$ , and  $\tilde{\mathcal{Z}}'_{JK}$  (see [Merlini and Morandini 2004b]) are nonlinear functions of the current nodal unknowns and thus need to be computed dynamically in the solution process. The curvature in the reference configuration is computed once for all as in (72).

The linearization of translational kinematics descends straightforwardly from (70)<sub>2</sub>. Since no double variations  $\partial\delta\mathbf{x}'_j$  of free variables make sense, and the interpolation is linear with the free variables  $\mathbf{x}'_j$ , it follows

$$\delta\mathbf{x}'_{,s} = \sum_{J=1}^N W_{J,s} \delta\mathbf{x}'_J, \quad \partial\mathbf{x}'_{,s} = \sum_{J=1}^N W_{J,s} \partial\mathbf{x}'_J, \quad (75)$$

and  $\partial\delta\mathbf{x}'_{,s}$  is discarded.

**7.3. Element implementation.** Equations (72)–(75) are now used within (63)–(64) to provide appropriate interpolating functions for the kinematical strain corotational variations, and the latter are finally used within (51). After integration over the element domain, the internal virtual work contribution from each element to the linearized functional is obtained in the form

$$\begin{aligned} \Pi_{\text{int}\delta}^e &= \int_{s^e} \pi_{\text{int}\delta} \, ds^e = \left\{ \boldsymbol{\varphi}_{J\delta} \right\}^T \cdot \left\{ \begin{array}{c} \mathbf{F}_{\boldsymbol{\varphi}J}^e \\ \mathbf{F}_{\mathbf{x}J}^e \end{array} \right\}, \\ \partial\Pi_{\text{int}\delta}^e &= \int_{s^e} \partial\pi_{\text{int}\delta} \, ds^e = \boldsymbol{\varphi}_{J\partial\delta} \cdot \mathbf{F}_{\boldsymbol{\varphi}J}^e + \left\{ \boldsymbol{\varphi}_{J\delta} \right\}^T \cdot \left[ \begin{array}{cc} \mathbf{K}_{\boldsymbol{\varphi}\boldsymbol{\varphi}JK}^e & \mathbf{K}_{\boldsymbol{\varphi}\mathbf{x}JK}^e \\ \mathbf{K}_{\boldsymbol{\varphi}\mathbf{x}JK}^{eT} & \mathbf{K}_{\mathbf{x}\mathbf{x}JK}^e \end{array} \right] \cdot \left\{ \begin{array}{c} \boldsymbol{\varphi}_{K\partial} \\ \partial\mathbf{x}'_K \end{array} \right\}. \end{aligned} \quad (76)$$

However, (76) cannot be assembled within the global structure model in the present form. The nodal mixed multipliers  $\boldsymbol{\varphi}_{J\partial\delta}$  do not have a correspondence in finite element formulations, and must be solved for the separate virtual and incremental nodal variation variables  $\boldsymbol{\varphi}_{J\delta}$  and  $\boldsymbol{\varphi}_{J\partial}$ . Actually, this resolution is feasible at each single node, where the rotation vector  $\boldsymbol{\varphi}_J$  is a truly free variable, so that  $\partial\delta\boldsymbol{\varphi}_J$  does not

exist at all. By applying the differential maps (36) and (37) to the total nodal rotation  $\Phi_J$  and discarding the term in  $\partial\delta\varphi_J$ , the resolution formula

$$\varphi_{J\partial\delta} = \underline{\Gamma_J} \cdot \partial\delta \overline{\varphi_J} + \Gamma_{J'}^{S123} : \Gamma_J^{-1} \varphi_{J\partial} \otimes \Gamma_J^{-1} \varphi_{J\delta} \quad (77)$$

is made available. (Note that  $\Gamma_J$  and  $\Gamma_{J'}$  in (77) are global tensors, built with the total nodal rotation vector  $\varphi_J$ .) The first term of  $\partial\Pi_{\text{int}\delta}^e$  can then be unfolded into a form like the second term and added to the latter. Thus, a workable expression for the increment of the internal virtual work is obtained; it provides a symmetrical and fully nonlinear tangent matrix, built with the interpolating functions (72)–(74), which are nonlinear functions of the current nodal orientations. The global tangent matrix can be assembled and the linearized problem is solved for the incremental unknowns  $\varphi_{K\partial}$  and  $\partial\mathbf{x}'_K$ . The nodal variables are then updated consistently as for  $\alpha'_K \leftarrow \exp(\varphi_{K\partial} \times) \alpha'_K$  and  $\mathbf{x}'_K \leftarrow \mathbf{x}'_K + \partial\mathbf{x}'_K$ .

It may be worth stressing that the double variation  $\partial\delta\varphi$  of the local rotation vector cannot be discarded as if  $\varphi$  were a free variable. This follows from the differential maps (36) and (37), since the local variation variables  $\varphi_\delta$ ,  $\varphi_\partial$ , and  $\varphi_{\partial\delta}$  are computed — consistently with the assumed interpolating model — from (73).

The framework of this nonlinear slender beam element, and in particular the interpolation methodology, are the same as in our solid element [Merlini and Morandini 2004b; 2005] and shell element [Merlini and Morandini 2011a]. The frame-invariance and path-independence of our interpolation scheme was discussed and proved in those papers. In those papers, actually, we adopted a particular modeling of the continuum — called *helicoïdal modeling* — where the angular and linear kinematic fields are coupled into a single field of orthogonal dual tensors, to which however the multiplicative interpolation discussed above applies as is. Incidentally, a working slender beam element based on helicoïdal modeling is ready and will be published separately.

## 8. Numerical tests

A two-node and a three-node beam element have been implemented in our own finite-element code, formerly developed for nonlinear solid and shell elements. In the two-node case, the interpolation discussed in Section 7 provides an element having a constant curvature  $\mathbf{k}$  along the element length. In this case, the linearization of the interpolation is much simpler, and manageable expressions for the corotational variations (63) and (64) can be obtained in the form of linear functions of the nodal variation variables (similarly to (73)–(75)), without the need of the third differential map (38). However, both the two-node and the three-node elements are built with the general form of the interpolation discussed above. Standard Lagrange polynomials are used for the weight functions  $W_J(s)$ . As a common practice in beam element technology, shear locking is avoided by resorting to reduced integration; we adopt Gaussian quadrature with one integration point for the two-node element and two integration points for the three-node element.

Two benchmark problems in nonlinear beam analysis illustrate the performance of the proposed element. In all computations, the convergence of the Newton–Raphson iterations is checked against a tolerance of  $10^{-2}$  on the maximum absolute value of the residual. No explicit units are reported here, but it is understood that all measures are associated with a coherent system, for example, SI.

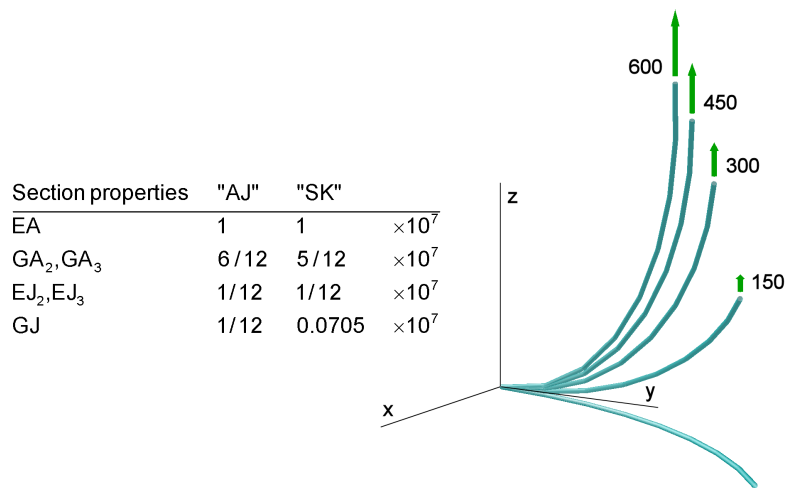
**8.1. Bending of 45° curved cantilever.** The cantilever bend was introduced by Bathe and Bolourchi [1979] and then repeated by almost every author of new nonlinear beam elements. The circular arc of

radius 100 spans  $45^\circ$  in the horizontal plane and is loaded at the free tip by a vertical force up to the final value 600 (Figure 3). The beam has a square cross-section of size  $1 \times 1$  and elastic material with Young's modulus  $E = 10^7$  and Poisson's ratio  $\nu = 0$ . After [Ritto-Corrêa and Camotim 2002], two sets of cross-section elastic properties are considered, according to a different choice of the effective shear areas and torsional constant. The set denoted by AJ takes unchanged areas  $A_2 = A_3 = 1$  for shear and polar moment of inertia  $J = J_2 + J_3$  for torsion; the set denoted by SK takes effective shear areas  $A_2 = A_3 = 5/6$  and torsional constant  $J = 0.141$ ; the values of both sets are listed in Figure 3.

Several meshes are analyzed, from coarse ones (4 elements) to refined ones (128 elements). All analyses are requested to seek convergence in a single load-step, but an algorithm of automatic step control makes the step size shrink or stretch dynamically. The resulting numbers of steps and iterations, together with the coordinates of the free tip at the final load 600, are listed in Tables 1 and 2. Some values available from recent literature are also compared in the tables, while earlier results — for which, however, it is not always certain which cross-section elastic properties were used — can be found in the cited papers.

**8.2. Cantilever beam twisted to a helical form.** A clamped slender beam is bent to a helical shape by a force and coaxial couple at the free tip, as illustrated in Figure 4 for a couple-to-force ratio  $M/F = 4\pi$ . At increasing load, the beam coils into narrower and narrower circles while the tip crosses the circles planes alternately from one side to the opposite side. At the final load  $M = 200\pi$  and  $F = 50$ , the helix develops in 10 circles and the out-of-plane displacement is opposite to the applied force. This example was introduced by Ibrahimbegović [1997] to test analyses with space rotations exceeding  $2\pi$ , and was then used by other authors [Zupan and Saje 2003; 2004; Mäkinen 2007; Zupan et al. 2009].

Five meshes with two-node and three-node elements are considered (Table 3). The load is applied in 100 equal steps, but an automatic step control may halve a step if necessary; this happens three to five times in any of the computations. However, the computations converge quite quickly, as can be seen



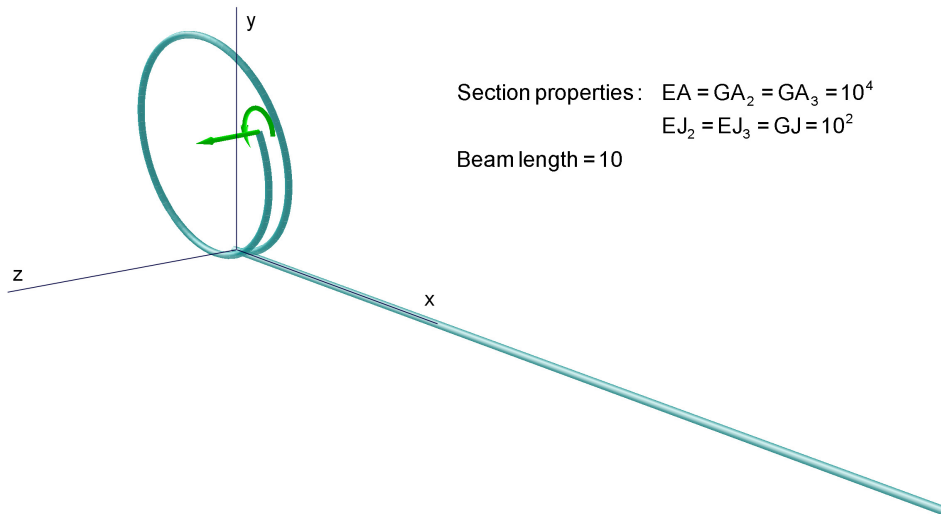
**Figure 3.**  $45^\circ$  bend: reference and deformed configurations of the SK beam model with eight two-node elements.

Element type	# of elements	# of steps	# of iterations	Tip coordinates at load 600		
				x	y	z
2-node <sup>1</sup>	8	6 equal	38	15.7426	47.2606	53.3730
2-node <sup>2</sup>	8	6 equal	54	15.67	47.29	53.37
2-node	8	6	41	15.7423	47.2600	53.3742
2-node	16	2	26	15.6991	47.1779	53.4495
2-node	32	6	41	15.6884	47.1573	53.4685
2-node	64	6	40	15.6857	47.1522	53.4733
2-node	128	6	40	15.6850	47.1509	53.4745
3-node	4	8	51	15.6837	47.1553	53.4671
3-node	8	6	42	15.6848	47.1507	53.4744
3-node	16	6	40	15.6848	47.1504	53.4748
3-node	32	6	41	15.6848	47.1504	53.4749
3-node	64	5	35	15.6848	47.1504	53.4749

**Table 1.** 45° bend: AJ model computation data. Data from <sup>(1)</sup> [Ritto-Corrêa and Camotim 2002] and <sup>(2)</sup> [Ghosh and Roy 2009].

from Table 3. To check the relevance of the resolution (77) of the nodal mixed variation variables, a test was run with the recovering of the first term of  $\partial \Pi_{\text{int}}^e$  in (76)<sub>2</sub> disabled, but it failed untimely at 3% of the final load.

The displacement of the free tip in the direction of the applied force is plotted in Figure 5 versus the fraction of the total load. Among the five meshes, only the response of the coarse mesh of twenty four



**Figure 4.** Helical beam: reference and deformed configuration of the model with ninety six two-node elements at 13% of the final load.



Element type	# of elements	# of steps	# of Iterations	Tip coordinates at load 600		
				x	y	z
2-node <sup>1</sup>	8	6 equal	38	15.6213	47.0142	53.4980
4-node <sup>3</sup>	1	4 equal	7–13 /step	15.57	46.95	53.51
4-node <sup>3</sup>	2	4 equal	7–13 /step	15.56	46.89	53.61
4-node <sup>3</sup>	8	4 equal	7–13 /step	15.56	46.89	53.61
3-int-pt <sup>4</sup>	8	6 equal	30	15.61	46.89	53.60
2-node	8	6	41	15.6211	47.0137	53.4991
2-node	16	2	26	15.5754	46.9259	53.5786
2-node	32	6	41	15.5640	46.9038	53.5988
2-node	64	6	40	15.5613	46.8982	53.6038
2-node	128	6	40	15.5606	46.8968	53.6050
3-node	4	8	51	15.5251	46.8361	53.6350
3-node	8	6	42	15.5410	46.8592	53.6295
3-node	16	6	40	15.5502	46.8766	53.6190
3-node	32	6	41	15.5551	46.8863	53.6125
3-node	64	5	35	15.5577	46.8913	53.6091

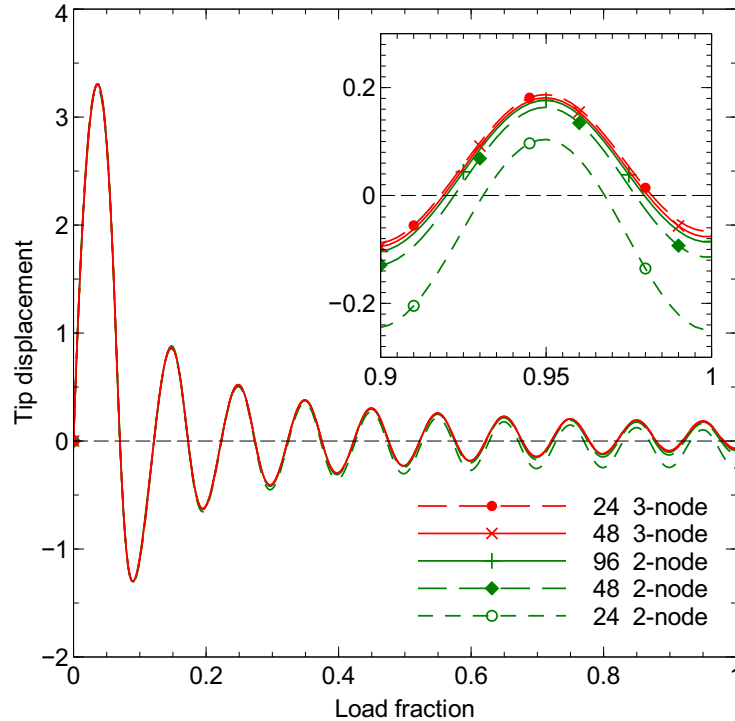
**Table 2.** 45° bend: SK model computation data. Data from <sup>(1)</sup> [Ritto-Corrêa and Camotim 2002], <sup>(3)</sup> [Kapania and Li 2003], and <sup>(4)</sup> [Zupan et al. 2009].

Element type	# of elements	# of steps	# of iterations	Average iterations per step
2-node	24	105	600	5.7
2-node	48	103	548	5.3
2-node	96	104	519	5.0
3-node	24	104	575	5.5
3-node	48	104	537	5.2

**Table 3.** Helical beam: computation data.

two-node elements is clearly distinguishable. These results are in good agreement with the papers cited above: though data for a tabular comparison are not available, a graphical superposition confirms that the curves in the published plots lie within the band between the curves in Figure 5. A series of coordinates of the free tip, as computed at completion of each successive coil, is listed in Table 4.

It is worth stressing that the computations are insensitive to rotations exceeding  $2\pi$ . This is due to the formulation implemented, which does not need to store and update total rotation vectors, hence eludes the discontinuities arising at rotation angles which are multiples of  $2\pi$ ; instead, the nodal orientation tensors are stored and updated via the incremental rotation tensors at each iteration. Of course, the proposed formulation is not able to handle the extreme case of relative nodal orientations exceeding  $2\pi$  within a single element; this causes the computation to stall, as punctually happened at just 95% of the final load when testing a mesh of twelve three-node elements.



**Figure 5.** Helical beam: tip displacement in the direction of the applied force.

### 9. Closing remarks

When tackling nonlinear problems of computational solid mechanics involving rotation fields, sooner or later a parameterization of the rotation tensor is introduced to allow a numerical solution. Thus, we may discern two subsequent stages: a first stage where an intrinsic, parameterization-free formulation is

Load fraction	x	y	z
0.0	10	0	0
0.1	0.413441	$52.67 \times 10^{-3}$	-1.078310
0.2	0.115254	$8.381 \times 10^{-3}$	-0.600272
0.3	0.052562	$2.606 \times 10^{-3}$	-0.402270
0.4	0.029900	$1.122 \times 10^{-3}$	-0.296987
0.5	0.018846	$589.0 \times 10^{-6}$	-0.230692
0.6	0.013505	$344.8 \times 10^{-6}$	-0.185107
0.7	0.009904	$219.0 \times 10^{-6}$	-0.151404
0.8	0.007593	$147.4 \times 10^{-6}$	-0.124996
0.9	0.005902	$101.7 \times 10^{-6}$	-0.103637
1.0	0.005318	$78.40 \times 10^{-6}$	-0.085524

**Table 4.** Helical beam: tip coordinates of the beam model with ninety six two-node elements.

developed, hopefully in a consistent way with the properties of the special orthogonal group the rotations belong to; and a second, parameterization-driven stage. In the sample beam problem we consider in this paper, an effort is made to keep intrinsic the formulation for as long as possible and to introduce a parameterization as late as possible.

After completing a consistent formulation and linearization of the beam variational mechanics, a consistent nonlinear, but properly linearized interpolating model provides an objective finite element approximation. Differential rotation vectors up to third order are used along these steps and lead to an approximate linearized virtual functional that includes the term  $\boldsymbol{\varphi}_{J\partial\delta} \cdot \mathbf{F}_{\varphi J}^e$ , which is unusual for the finite element method. All these steps are performed without resorting to any parameterization, so the outlined formulation is consistent and intrinsic. The parameterization of the rotation tensor is resorted to only twice in the discrete approximation: when linearizing the algorithm that interpolates the section orientation among the nodal orientations, and when solving the nodal mixed multipliers  $\boldsymbol{\varphi}_{J\partial\delta}$  while assembling the beam elements.

The proposed differentiations of the rotation tensor are applied here to the one-dimensional case of slender beams just as an example. However, they are also valid for other continuum mechanics problems, and proved to be a valuable tool for shells and solids, where two and three-dimensional gradients are involved.

### Appendix: Tensor reference guide

A short reference guide, in the form of a bare collection of tensor notations and rules, is appended here to help in following the mathematical developments in the paper. Hereafter,  $\mathbf{a}$  denotes a vector,  $\mathbf{A}$  a second-order tensor,  $\mathcal{A}$  a third-order tensor, and  $\mathbb{A}$  a fourth-order tensor. A three-dimensional space is understood, with  $\mathbf{g}_j$  and  $\mathbf{g}^j$  ( $j = 1, 2, 3$ ) reciprocal triads of base vectors. Repeated covariant and contravariant indexes entail summation from 1 to 3 (Einsteinian rule).

*Polyadic representation.* For example:

$$\mathbb{A} = \mathbb{A}_l \otimes \mathbf{g}^l = \mathbb{A}_{kl} \otimes \mathbf{g}^k \otimes \mathbf{g}^l = \mathbb{A}_{jkl} \otimes \mathbf{g}^j \otimes \mathbf{g}^k \otimes \mathbf{g}^l = \mathbb{A}_{ijkl} \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \otimes \mathbf{g}^l,$$

where components  $\mathbb{A}_l$ ,  $\mathbb{A}_{kl}$ ,  $\mathbb{A}_{jkl}$ , and  $\mathbb{A}_{ijkl}$  are third-order tensors, second-order tensors, vectors, and scalars, respectively.

*Dot operators.* Multiple dots saturate the neighboring polyadic legs in order left-to-right. For example:

$$\mathbb{A} \dot{\dot{\dot{\mathbb{B}}}} = \mathbb{A}_{jkl} \otimes \mathbf{g}^j \otimes \mathbf{g}^k \otimes \mathbf{g}^l \dot{\dot{\dot{\mathbb{B}}}}^{pqr} \mathbf{g}_p \otimes \mathbf{g}_q \otimes \mathbf{g}_r$$

produces vector  $\mathbb{A}_{jkl} \mathbb{B}^{jkl}$ . A single dot is often understood.

*Transpose tensors.* For example:

$$\begin{aligned} \mathbf{A} &= \mathbf{A}_{ij} \mathbf{g}^i \otimes \mathbf{g}^j, & \mathbf{A}^T &= \mathbf{A}_{ij} \mathbf{g}^j \otimes \mathbf{g}^i; \\ \mathcal{A} &= \mathcal{A}_{ijk} \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k = \mathcal{A}^{\text{T312T231}} = \mathcal{A}^{\text{T231T312}}, \\ \mathcal{A}^{\text{T231}} &= \mathcal{A}_{ijk} \mathbf{g}^j \otimes \mathbf{g}^k \otimes \mathbf{g}^i = \mathcal{A}^{\text{T312T312}}, & \mathcal{A}^{\text{T312}} &= \mathcal{A}_{ijk} \mathbf{g}^k \otimes \mathbf{g}^i \otimes \mathbf{g}^j = \mathcal{A}^{\text{T231T231}}, \\ \mathcal{A}^{\text{T132}} &= \mathcal{A}_{ijk} \mathbf{g}^i \otimes \mathbf{g}^k \otimes \mathbf{g}^j, & \mathcal{A}^{\text{T321}} &= \mathcal{A}_{ijk} \mathbf{g}^k \otimes \mathbf{g}^j \otimes \mathbf{g}^i, & \mathcal{A}^{\text{T213}} &= \mathcal{A}_{ijk} \mathbf{g}^j \otimes \mathbf{g}^i \otimes \mathbf{g}^k. \end{aligned}$$

*Symmetric tensors.* Symmetric second-order tensor:

$$\underline{\underline{A}}^{12} \quad \text{if} \quad \mathbf{A} = \mathbf{A}^T.$$

Simple-symmetric third-order tensors:

$$\underline{\underline{\mathcal{A}}}^{123} \quad \text{if} \quad \mathcal{A} = \mathcal{A}^{T132}, \quad \underline{\underline{\mathcal{A}}}^{123} \quad \text{if} \quad \mathcal{A} = \mathcal{A}^{T321}, \quad \underline{\underline{\mathcal{A}}}^{123} \quad \text{if} \quad \mathcal{A} = \mathcal{A}^{T213};$$

double-symmetric third-order tensor:

$$\underline{\underline{\mathcal{A}}}^{123} \quad \text{if} \quad \mathcal{A} = \mathcal{A}^{T231} = \mathcal{A}^{T312};$$

and full-symmetric third-order tensor:

$$\underline{\underline{\underline{\mathcal{A}}}}^{123} \quad \text{if} \quad \mathcal{A} = \mathcal{A}^{T231} = \mathcal{A}^{T312} = \mathcal{A}^{T132} = \mathcal{A}^{T321} = \mathcal{A}^{T213}.$$

Such symmetries are used as well for the rightmost polyadic legs of fourth-order tensors.

*Tensor additive decompositions.* Second-order tensor decomposition with a symmetric part:

$$\mathbf{A} = \mathbf{A}^S + \mathbf{A}^A, \quad \mathbf{A}^S = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T), \quad \mathbf{A}^A = \frac{1}{2}(\mathbf{A} - \mathbf{A}^T) = \mathbf{a} \times, \quad (\text{A.1})$$

where  $\mathbf{a} = \text{ax } \mathbf{A} = \frac{1}{2} \mathbf{g}^j \times \mathbf{A} \mathbf{g}_j = \frac{1}{2} \mathbf{I}^\times : \mathbf{A}$  is the axial vector of  $\mathbf{A}$ . Third-order tensor decompositions with simple-symmetric parts:

$$\begin{aligned} \mathcal{A} &= \underline{\underline{\mathcal{A}}}^{S123} + \underline{\underline{\mathcal{A}}}^{A123}, & \underline{\underline{\mathcal{A}}}^{S123} &= \frac{1}{2}(\mathcal{A} + \underline{\underline{\mathcal{A}}}^{T132}), & \underline{\underline{\mathcal{A}}}^{A123} &= \frac{1}{2}(\mathcal{A} - \underline{\underline{\mathcal{A}}}^{T132}), \\ \mathcal{A} &= \underline{\underline{\mathcal{A}}}^{S123} + \underline{\underline{\mathcal{A}}}^{A123}, & \underline{\underline{\mathcal{A}}}^{S123} &= \frac{1}{2}(\mathcal{A} + \underline{\underline{\mathcal{A}}}^{T321}), & \underline{\underline{\mathcal{A}}}^{A123} &= \frac{1}{2}(\mathcal{A} - \underline{\underline{\mathcal{A}}}^{T321}), \\ \mathcal{A} &= \underline{\underline{\mathcal{A}}}^{S123} + \underline{\underline{\mathcal{A}}}^{A123}, & \underline{\underline{\mathcal{A}}}^{S123} &= \frac{1}{2}(\mathcal{A} + \underline{\underline{\mathcal{A}}}^{T213}), & \underline{\underline{\mathcal{A}}}^{A123} &= \frac{1}{2}(\mathcal{A} - \underline{\underline{\mathcal{A}}}^{T213}); \end{aligned} \quad (\text{A.2})$$

third-order tensor decomposition with a double-symmetric part:

$$\mathcal{A} = \underline{\underline{\mathcal{A}}}^{S123} + \underline{\underline{\mathcal{A}}}^{A123}, \quad \begin{cases} \underline{\underline{\mathcal{A}}}^{S123} = \frac{1}{3}(\mathcal{A} + \underline{\underline{\mathcal{A}}}^{T231} + \underline{\underline{\mathcal{A}}}^{T312}), \\ \underline{\underline{\mathcal{A}}}^{A123} = \frac{1}{3}(2\mathcal{A} - \underline{\underline{\mathcal{A}}}^{T231} - \underline{\underline{\mathcal{A}}}^{T312}); \end{cases} \quad (\text{A.3})$$

and third-order tensor decomposition with a full-symmetric part:

$$\mathcal{A} = \underline{\underline{\underline{\mathcal{A}}}}^{S123} + \underline{\underline{\underline{\mathcal{A}}}}^{A123}, \quad \begin{cases} \underline{\underline{\underline{\mathcal{A}}}}^{S123} = \frac{1}{6}(\mathcal{A} + \underline{\underline{\mathcal{A}}}^{T231} + \underline{\underline{\mathcal{A}}}^{T312} + \underline{\underline{\mathcal{A}}}^{T132} + \underline{\underline{\mathcal{A}}}^{T321} + \underline{\underline{\mathcal{A}}}^{T213}), \\ \underline{\underline{\underline{\mathcal{A}}}}^{A123} = \frac{1}{6}(5\mathcal{A} - \underline{\underline{\mathcal{A}}}^{T231} - \underline{\underline{\mathcal{A}}}^{T312} - \underline{\underline{\mathcal{A}}}^{T132} - \underline{\underline{\mathcal{A}}}^{T321} - \underline{\underline{\mathcal{A}}}^{T213}). \end{cases} \quad (\text{A.4})$$

Such decompositions are used as well for the rightmost polyadic legs of fourth-order tensors. For example:

$$\mathbb{A} = \mathbb{A}^{\underline{\underline{\underline{S}}1234}} + \mathbb{A}^{\underline{\underline{\underline{A}}1234}}, \quad \begin{cases} \mathbb{A}^{\underline{\underline{\underline{S}}1234}} = \frac{1}{6}(\mathbb{A} + \mathbb{A}^{T1342} + \mathbb{A}^{T1423} + \mathbb{A}^{T1243} + \mathbb{A}^{T1432} + \mathbb{A}^{T1324}), \\ \mathbb{A}^{\underline{\underline{\underline{A}}1234}} = \frac{1}{6}(5\mathbb{A} - \mathbb{A}^{T1342} - \mathbb{A}^{T1423} - \mathbb{A}^{T1243} - \mathbb{A}^{T1432} - \mathbb{A}^{T1324}). \end{cases} \quad (\text{A.5})$$

Unitary tensors. Second-order tensor identity:

$$\mathbf{I} = \mathbf{g}_j \otimes \mathbf{g}^j.$$

Third-order Ricci's tensor:

$$\mathbf{I}^\times = \mathbf{g}_j \times \otimes \mathbf{g}^j = \mathbf{I}^{\times T312} = \mathbf{g}_j \otimes \mathbf{g}^j \times = \mathbf{I}^{\times T231} = -\mathbf{I}^{\times T213} = -\mathbf{I}^{\times T321} = -\mathbf{I}^{\times T132}. \quad (\text{A.6})$$

Fourth-order unitary tensors:

$$\begin{aligned} \mathbb{I} &= \mathbf{g}_j \otimes \mathbf{g}_k \otimes \mathbf{g}^j \otimes \mathbf{g}^k = \check{\mathbb{I}}^{T1243} = (\mathbf{I} \otimes \mathbf{I})^{T1324}, \\ \check{\mathbb{I}} &= \mathbf{g}_j \otimes \mathbf{g}_k \otimes \mathbf{g}^k \otimes \mathbf{g}^j = \mathbb{I}^{T1243} = (\mathbf{I} \otimes \mathbf{I})^{T1342}, \\ \mathbf{I} \otimes \mathbf{I} &= \mathbf{g}_j \otimes \mathbf{g}^j \otimes \mathbf{g}_k \otimes \mathbf{g}^k = \mathbb{I}^{T1324} = \check{\mathbb{I}}^{T1423}, \end{aligned}$$

where, according to the standards introduced by Del Piero [1979], the fourth-order identity also writes  $\mathbb{I} = \mathbf{I} \boxtimes \mathbf{I}$ , and  $\check{\mathbb{I}}$  is the transposer  $\mathbb{T}$  (the tensor product  $\boxtimes$  between second-order tensors converts, in our notation, as  $\mathbf{A} \boxtimes \mathbf{B} = (\mathbf{A} \otimes \mathbf{B})^{T1324}$ ); other useful fourth-order unitary tensors are the symmetrizer  $\mathbb{S} = \frac{1}{2}(\mathbb{I} + \check{\mathbb{I}})$  and the skew-symmetrizer

$$\mathbb{W} = \frac{1}{2}(\mathbb{I} - \check{\mathbb{I}}) = \frac{1}{2}\mathbf{I}^\times \mathbf{I}^\times = \frac{1}{2}\mathbf{g}_j \times \otimes \mathbf{g}^j \times.$$

Some useful tensor identities.

$$\mathbf{a} \times = \mathbf{I}^\times \mathbf{a} = \mathbf{a} \cdot \mathbf{I}^\times, \quad (\text{A.7})$$

$$\mathbf{a} \times \mathbf{b} = \mathbf{I}^\times \mathbf{a} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{I}^\times \mathbf{b} = \mathbf{I}^\times : \mathbf{b} \otimes \mathbf{a}, \quad (\text{A.8})$$

$$\mathbf{a} \times \mathbf{b} \times = \mathbf{b} \otimes \mathbf{a} - \mathbf{a} \cdot \mathbf{b} \otimes \mathbf{I}, \quad (\text{A.9})$$

$$\mathbf{a} \times : \mathbf{b} \times = 2\mathbf{a} \cdot \mathbf{b}, \quad (\text{A.10})$$

$$\mathbf{a} \times \mathbf{a} \times \mathbf{b} \times \mathbf{a} \times = \mathbf{a} \times \mathbf{b} \times \mathbf{a} \times \mathbf{a} \times, \quad (\text{A.11})$$

$$(\mathbf{a} \times \mathbf{b}) \times = \mathbf{I}^\times \mathbf{I}^\times : \mathbf{b} \otimes \mathbf{a} = \mathbf{b} \otimes \mathbf{a} - \mathbf{a} \otimes \mathbf{b} = \mathbf{a} \times \mathbf{b} \times - \mathbf{b} \times \mathbf{a} \times, \quad (\text{A.12})$$

$$\mathbf{I}^\times \mathbf{a} \times - \mathbf{a} \times \mathbf{I}^\times = \mathbf{I} \otimes \mathbf{a} - \mathbf{a} \otimes \mathbf{I}, \quad (\text{A.13})$$

$$\mathbf{a} \times \mathbf{a} \times \mathbf{I}^\times \mathbf{a} \times = \mathbf{a} \times \mathbf{I}^\times \mathbf{a} \times \mathbf{a} \times, \quad (\text{A.14})$$

$$\mathbf{a} \times \mathbf{c} \times \mathbf{b} \times + \mathbf{b} \times \mathbf{c} \times \mathbf{a} \times = -((\mathbf{a} \otimes \mathbf{b} + \mathbf{b} \otimes \mathbf{a}) \cdot \mathbf{c}) \times, \quad (\text{A.15})$$

$$\mathbf{I}^\times \mathbf{A} = (\mathbf{A}^T \mathbf{I}^\times)^{T231}, \quad \mathbf{A}^T \mathbf{I}^\times = (\mathbf{I}^\times \mathbf{A})^{T312}, \quad (\text{A.16})$$

$$((\mathbf{I}^\times \mathbf{A})^{T132} \mathbf{A})^{T132} = -(\mathbf{I}^\times \mathbf{A})^{T132} \mathbf{A}. \quad (\text{A.17})$$

## References

- [Altenbach et al. 2010] J. Altenbach, H. Altenbach, and V. A. Eremeyev, "On generalized Cosserat-type theories of plates and shells: a short review and bibliography", *Arch. Appl. Mech.* **80** (2010), 73–92.
- [Argyris 1982] J. Argyris, "An excursion into large rotations", *Comput. Methods Appl. Mech. Eng.* **32**:1-3 (1982), 85–155.
- [Atluri and Cazzani 1995] S. N. Atluri and A. Cazzani, "Rotations in computational solid mechanics", *Arch. Comput. Methods Eng.* **2**:1 (1995), 49–138.
- [Bathe and Bolourchi 1979] K.-J. Bathe and S. Bolourchi, "Large displacement analysis of three-dimensional beam structures", *Int. J. Numer. Methods Eng.* **14** (1979), 961–986.

- [Bauchau and Trainelli 2003] O. A. Bauchau and L. Trainelli, “The vectorial parameterization of rotation”, *Nonlinear Dyn.* **32**:1 (2003), 71–92.
- [Bauer et al. 2010] S. Bauer, M. Schäfer, P. Grammenoudis, and C. Tsakmakis, “Three-dimensional finite elements for large deformation micropolar elasticity”, *Comput. Methods Appl. Mech. Eng.* **199**:41-44 (2010), 2643–2654.
- [Borri et al. 1990] M. Borri, F. Mello, and S. N. Atluri, “Variational approaches for dynamics and time-finite-elements: numerical studies”, *Comput. Mech.* **7** (1990), 49–76.
- [Borri et al. 2000] M. Borri, L. Trainelli, and C. L. Bottasso, “On representations and parameterizations of motion”, *Multibody Syst. Dyn.* **4**:2-3 (2000), 129–193.
- [Cardona and Géradin 1988] A. Cardona and M. Géradin, “A beam finite element non-linear theory with finite rotations”, *Int. J. Numer. Methods Eng.* **26**:11 (1988), 2403–2438.
- [Cheng and Gupta 1989] H. Cheng and K. C. Gupta, “An historical note on finite rotations”, *J. Appl. Mech. (ASME)* **56**:1 (1989), 139–145.
- [Crisfield and Jelenić 1999] M. A. Crisfield and G. Jelenić, “Objectivity of strain measures in the geometrically exact three-dimensional beam theory and its finite-element implementation”, *Proc. R. Soc. Lond. A.* **455**:1983 (1999), 1125–1147.
- [Del Piero 1979] G. Del Piero, “Some properties of the set of fourth-order tensors, with application to elasticity”, *J. Elasticity* **9**:3 (1979), 245–261.
- [Dui et al. 2006] G.-S. Dui, Z.-D. Wang, and M. Jin, “Derivatives on the isotropic tensor functions”, *Sci. China G Phys. Mech. Astronom.* **49** (2006), 321–334.
- [Fung 2004] T. C. Fung, “Computation of the matrix exponential and its derivatives by scaling and squaring”, *Int. J. Numer. Methods Eng.* **59**:10 (2004), 1273–1286.
- [Ghosh and Roy 2009] S. Ghosh and D. Roy, “A frame-invariant scheme for the geometrically exact beam using rotation vector parametrization”, *Comput. Mech.* **44** (2009), 103–118.
- [Grekova and Zhilin 2001] E. Grekova and P. Zhilin, “Basic equations of Kelvin’s medium and analogy with ferromagnets”, *J. Elasticity* **64**:1 (2001), 29–70.
- [Ibrahimbegović 1995] A. Ibrahimbegović, “On the finite element implementation of geometrically non-linear Reissner’s beam theory: 3D curved beam element”, *Comput. Methods Appl. Mech. Eng.* **122** (1995), 11–26.
- [Ibrahimbegović 1997] A. Ibrahimbegović, “On the choice of finite rotation parameters”, *Comput. Methods Appl. Mech. Eng.* **149**:1-4 (1997), 49–71.
- [Ibrahimbegović and Taylor 2002] A. Ibrahimbegović and R. L. Taylor, “On the role of frame-invariance in structural mechanics models at finite rotations”, *Comput. Methods Appl. Mech. Eng.* **191**:45 (2002), 5159–5176.
- [Ibrahimbegović et al. 1995] A. Ibrahimbegović, F. Frey, and I. Kozar, “Computational aspects of vector-like parametrization of three-dimensional finite rotations”, *Int. J. Numer. Methods Eng.* **38**:21 (1995), 3653–3673.
- [Itskov 2002] M. Itskov, “The derivative with respect to a tensor: some theoretical aspects and applications”, *Z. Angew. Math. Mech.* **82**:8 (2002), 535–544.
- [Itskov 2003] M. Itskov, “Computation of the exponential and other isotropic tensor functions and their derivatives”, *Comput. Methods Appl. Mech. Eng.* **192**:35-36 (2003), 3985–3999.
- [Itskov and Aksel 2002] M. Itskov and N. Aksel, “A closed-form representation for the derivative of non-symmetric tensor power series”, *Int. J. Solids Struct.* **39**:24 (2002), 5963–5978.
- [Jelenić and Crisfield 1999] G. Jelenić and M. A. Crisfield, “Geometrically exact 3D beam theory: implementation of a strain-invariant finite element for statics and dynamics”, *Comput. Methods Appl. Mech. Eng.* **171**:1-2 (1999), 141–171.
- [Jog 2008] C. S. Jog, “The explicit determination of the logarithm of a tensor and its derivatives”, *J. Elasticity* **93**:2 (2008), 141–148.
- [Kapania and Li 2003] R. K. Kapania and J. Li, “A formulation and implementation of geometrically exact curved beam elements incorporating finite strains and finite rotations”, *Comput. Mech.* **30** (2003), 444–459.
- [Lu 2004] J. Lu, “Exact expansions of arbitrary tensor functions  $\mathbf{F}(\mathbf{A})$  and their derivatives”, *Int. J. Solids Struct.* **41**:2 (2004), 337–349.

- [Mäkinen 2007] J. Mäkinen, “Total Lagrangian Reissner’s geometrically exact beam element without singularities”, *Int. J. Numer. Methods Eng.* **70**:9 (2007), 1009–1048.
- [Mäkinen 2008] J. Mäkinen, “Rotation manifold  $SO(3)$  and its tangential vectors”, *Comput. Mech.* **42**:6 (2008), 907–919.
- [Merlini 1997] T. Merlini, “A variational formulation for finite elasticity with independent rotation and Biot-axial fields”, *Comput. Mech.* **19** (1997), 153–168.
- [Merlini 2002] T. Merlini, “Differentiation of rotation and rototranslation”, Scientific Report DIA-SR 02-16, Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano, 2002, Available at <http://www.aero.polimi.it/diasr/02-16.pdf>.
- [Merlini and Morandini 2004a] T. Merlini and M. Morandini, “The helicoidal modeling in computational finite elasticity. Part I: Variational formulation”, *Int. J. Solids Struct.* **41**:18-19 (2004), 5351–5381.
- [Merlini and Morandini 2004b] T. Merlini and M. Morandini, “The helicoidal modeling in computational finite elasticity. Part II: Multiplicative interpolation”, *Int. J. Solids Struct.* **41**:18-19 (2004), 5383–5409. Erratum in *Int. J. Solids Struct.* **42** (2005), 1269.
- [Merlini and Morandini 2005] T. Merlini and M. Morandini, “The helicoidal modeling in computational finite elasticity. Part III: Finite element approximation for non-polar media”, *Int. J. Solids Struct.* **42** (2005), 6475–6513.
- [Merlini and Morandini 2011a] T. Merlini and M. Morandini, “Computational shell mechanics by helicoidal modeling, II: Shell element”, *J. Mech. Mater. Struct.* **6** (2011), 693–728.
- [Merlini and Morandini 2011b] T. Merlini and M. Morandini, “Consistency issues in shell elements for geometrically nonlinear problems”, pp. 355–376 in *Shell-like structures: nonclassical theories and applications*, edited by H. Altenbach and V. Eremeyev, Advanced Structured Materials **15**, Springer, Heidelberg, 2011.
- [Najfeld and Havel 1995] I. Najfeld and T. F. Havel, “Derivatives of the matrix exponential and their computation”, *Adv. Appl. Math.* **16**:3 (1995), 321–375.
- [Ortiz et al. 2001] M. Ortiz, R. A. Radovitzky, and E. A. Repetto, “The computation of the exponential and logarithmic mappings and their first and second linearizations”, *Int. J. Numer. Methods Eng.* **52**:12 (2001), 1431–1441.
- [Pietraszkiewicz and Badur 1983] W. Pietraszkiewicz and J. Badur, “Finite rotations in the description of continuum deformation”, *Int. J. Eng. Sci.* **21**:9 (1983), 1097–1115.
- [Reissner 1973] E. Reissner, “On one-dimensional, large-displacement, finite-strain beam theory”, *Stud. Appl. Math.* **52** (1973), 87–95.
- [Ritto-Corrêa and Camotim 2002] M. Ritto-Corrêa and D. Camotim, “On the differentiation of the Rodrigues formula and its significance for the vector-like parameterization of Reissner–Simo beam theory”, *Int. J. Numer. Methods Eng.* **55**:9 (2002), 1005–1032.
- [Rosati 1999] L. Rosati, “Derivatives and rates of the stretch and rotation tensors”, *J. Elasticity* **56**:3 (1999), 213–230.
- [Simo 1985] J. C. Simo, “A finite strain beam formulation. The three-dimensional dynamic problem. Part I”, *Comput. Methods Appl. Mech. Eng.* **49** (1985), 55–70.
- [Simo and Vu-Quoc 1986] J. C. Simo and L. Vu-Quoc, “A three-dimensional finite-strain rod model. Part II: Computational aspects”, *Comput. Methods Appl. Mech. Eng.* **58** (1986), 79–116.
- [Simo et al. 1992] J. C. Simo, D. D. Fox, and T. J. R. Hughes, “Formulations of finite elasticity with independent rotations”, *Comput. Methods Appl. Mech. Eng.* **95**:2 (1992), 277–288.
- [de Souza Neto 2001] E. A. de Souza Neto, “The exact derivative of the exponential of an unsymmetric tensor”, *Comput. Methods Appl. Mech. Eng.* **190**:18-19 (2001), 2377–2383.
- [de Souza Neto 2004] E. A. de Souza Neto, “On general isotropic tensor functions of one tensor”, *Int. J. Numer. Methods Eng.* **61**:6 (2004), 880–895.
- [Wang and Dui 2007] Z.-Q. Wang and G.-S. Dui, “On the derivatives of a subclass of isotropic tensor functions of a nonsymmetric tensor”, *Int. J. Solids Struct.* **44**:16 (2007), 5369–5379.
- [Zupan and Saje 2003] D. Zupan and M. Saje, “Finite-element formulation of geometrically exact three-dimensional beam theories based on interpolation of strain measures”, *Comput. Methods Appl. Mech. Eng.* **192**:49-50 (2003), 5209–5248.
- [Zupan and Saje 2004] D. Zupan and M. Saje, “Rotational invariants in finite element formulation of three-dimensional beam theories”, *Comput. Struct.* **82** (2004), 2027–2040.

[Zupan et al. 2009] E. Zupan, M. Saje, and D. Zupan, “The quaternion-based three-dimensional beam theory”, *Comput. Methods Appl. Mech. Eng.* **198**:49-52 (2009), 3944–3956.

Received 27 Nov 2012. Revised 1 Mar 2013. Accepted 1 Mar 2013.

TEODORO MERLINI: *Dipartimento di Scienze e Tecnologie Aerospaziali, Politecnico di Milano, Campus Bovisa, via La Masa 34, 20156 Milano, Italy*

MARCO MORANDINI: `marco.morandini@polimi.it`

*Dipartimento di Scienze e Tecnologie Aerospaziali, Politecnico di Milano, Campus Bovisa, via La Masa 34, 20156 Milano, Italy*



## PREDICTING THE EFFECTIVE STIFFNESS OF CELLULAR AND COMPOSITE MATERIALS WITH SELF-SIMILAR HIERARCHICAL MICROSTRUCTURES

YI MIN XIE, ZHI HAO ZUO, XIAODONG HUANG AND XIAOYING YANG

Many natural and man-made materials exhibit self-similar hierarchical microstructures on several length scales. The effective macroscopic mechanical properties of such materials or composites are affected by the number of hierarchical levels and the topology of microstructures. Although the effective mechanical properties can be determined numerically using homogenization techniques, the computational costs can become prohibitively high as the level of hierarchy increases. This paper proposes an analytical approach to predicting the effective stiffness of a class of materials and structures with self-similar hierarchical microstructures. For each microstructural configuration, a simple relationship between the effective stiffness and the hierarchical level is established and verified against results of finite element analysis or data in the literature. It is found that the simple relationships we have developed provide quite accurate stiffness predictions of various hierarchical materials and composites including the Menger sponge. For composites, the predicted effective stiffness is accurate even when one of the phases is near its incompressibility limit, with its Poisson ratio close to 0.5. Inspired by the Menger sponge and informed by our topology optimization result, we propose a lighter yet stiffer “cross sponge”.

### 1. Introduction

Hierarchical solids contain structural elements which themselves have structures on more than one length scale [Lakes 1993]. Multilevel structural hierarchy, as observed in many living organisms, seems to be a universal strategy adopted by natural evolution for realizing remarkable properties and functions [Currey 1984; Aizenberg et al. 2005; Zhang et al. 2011].

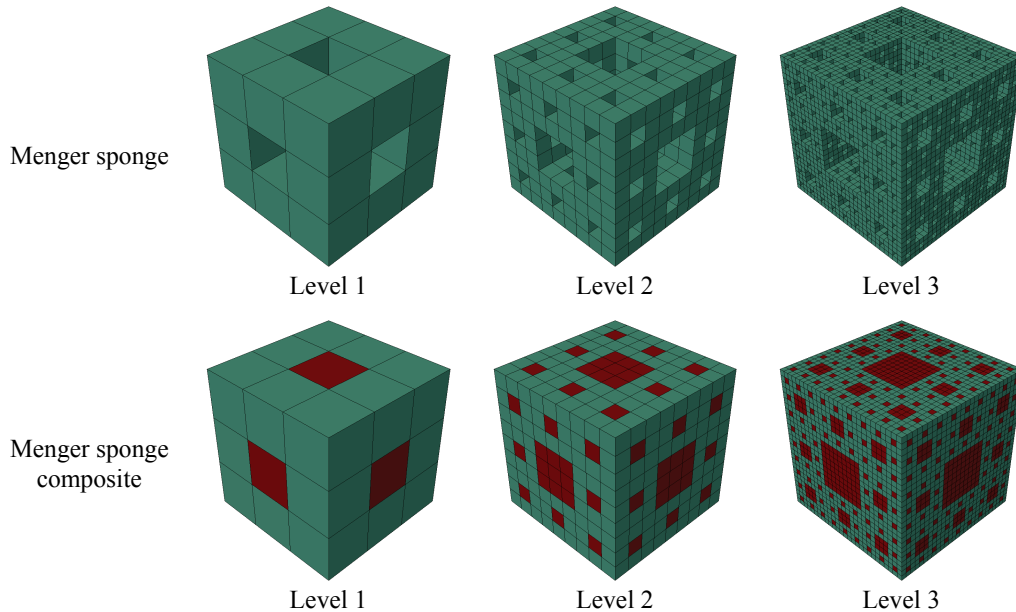
According to Currey [1977; 1984] and Gao [2010], a sea shell has two to three levels of lamellar structure while bone has seven levels of structural hierarchy. It has been observed from many biological materials that the microstructures at different hierarchical levels often exhibit striking self-similarity [Jäger and Fratzl 2000; Puxkandl et al. 2002; Gao 2010; Zhang et al. 2011]. A mineralized tendon fiber, for example, has four levels of hierarchy with a highly ordered, self-similar structure at every level [Puxkandl et al. 2002; Zhang et al. 2011].

Among 3D solids with self-similar hierarchical structures or microstructures, the Menger sponge shown in Figure 1 is perhaps the most famous example. Starting from a solid cubic that is divided equally into  $3 \times 3 \times 3$  subcubes, the Menger sponge is created by simply removing the seven subcubes at the body and face centers of each remaining solid cube from the previous level. If the voids are replaced with an inclusion material, the Menger sponge becomes a composite with two phases. The Menger

---

This work is supported by the Australian Research Council's Discovery Projects (DP1094401).

*Keywords:* Menger sponge, cellular material, composite material, structural hierarchy, homogenization.

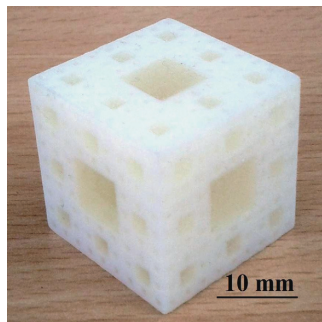


**Figure 1.** The first three levels of Menger sponges and Menger sponge composites.

sponge presents self-similar structural hierarchy since each subcube has the same topology as that of the cube of the previous level.

Figure 2 shows a level-three Menger sponge composite we have fabricated using the Connex350 3D printer [Stratasys 2012] which allows the simultaneous printing of two different materials. The transparent inclusion material is a rubber-like elastomer and the white matrix material is an acrylic-based glassy polymer.

There has been extensive research on the estimation of the macroscopic effective mechanical properties of load-carrying media [Watt et al. 1976; Sánchez-Palencia 1980; Coussy 1991; Tan et al. 1999; Liu et al. 2009]. For this purpose, numerical methods have been commonly applied; for instance, Hashin and Shtrikman [1962; 1963] proposed a variation approach for polycrystals and multiphase media, Cleary



**Figure 2.** A level-three Menger sponge composite fabricated using a multimaterial 3D printer.

et al. [1980] used techniques of self-consistent approximations for heterogeneous media, Day et al. [1992] implemented a discretized-spring scheme for a sheet with circular inclusions, Garboczi and Day [1995] analyzed the mechanical properties of 3D composite using a finite element (FE) algorithm, and Poutet et al. [1996] applied a multiple-scale expansion scheme to random porous media including some hierarchical media. Apart from these, numerical homogenization techniques (for example, [Bakhvalov and Panasenko 1984; Nemat-Nasser and Hori 1993]) have been frequently used. Similarly, Steven [1997] and Tan et al. [2000] set up appropriate boundary conditions on a material base cell to attain elastic properties by using finite element analysis.

There has been limited work on the prediction of the effective physical and mechanical properties of hierarchical materials and structures. Thovert et al. [1990] studied the thermal conductivity of regular fractals and Poutet et al. [1996] examined elastic constants of a variety of porous media. However, due to the limitations of computer capacities of the time, only very low levels of the hierarchy were solved. This is because, as the level of hierarchy increases, computational cost can become prohibitively high. Oshmyan et al. [2001] studied the elastic properties of a special kind of self-similar hierarchical structures (2D Sierpinski-like structures) using FE-based simulations. They also considered Sierpinski composites with rigid inclusions. Our current work will cover general 3D self-similar hierarchical cellular structures as well as composites with deformable inclusions. Simple formulas for predicting the elastic moduli have been derived using the renormalization argument [Bergman and Kantor 1984; Poutet et al. 1996]. From the renormalization argument, Poutet et al. [1996] found that the ratio of the Young's moduli between two adjacent levels should be 2/3. Since then, this result has been accepted and cited by other researchers [Picu and Soare 2009]. We shall show in this paper that the prediction from the renormalization argument is inaccurate with an error of more than 20%. Indeed, even if one compares the numerical results of [Poutet et al. 1996] with the renormalization prediction, one can see differences of the same magnitude. In other words, the renormalization argument can be highly inaccurate. Therefore a new and more accurate approach to the prediction of the effective stiffness of self-similar hierarchical media needs be established. This will be the focus and the main contribution of the present paper. In addition, we shall propose a lighter yet stiffer hierarchical material than the Menger sponge.

## 2. Preliminary basics

**2.1. Terminology for mechanical properties.** In the literature, various terms are used to express the same item, such as the stiffness constitutive matrix, which can be denoted as a tensor  $C$  or  $E$ . The term stiffness may also refer to different mechanical quantities. For clarity, we shall first briefly define the terminology for mechanical properties. The following formula expresses Hooke's law in stiffness form with  $C$  as the stiffness matrix:

$$\sigma = C \cdot \epsilon \quad \text{or} \quad \begin{Bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{13} \end{Bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{bmatrix} \begin{Bmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ \epsilon_{12} \\ \epsilon_{23} \\ \epsilon_{13} \end{Bmatrix}. \quad (1)$$

Very often, Hooke's law may be expressed using the engineering constants. Since the hierarchical media addressed in this paper are all symmetric with respect to the three middle planes, the orthotropic elasticity can be defined as

$$\begin{Bmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ \epsilon_{12} \\ \epsilon_{23} \\ \epsilon_{13} \end{Bmatrix} = \begin{bmatrix} 1/E_{11} & -\nu_{12}/E_{22} & -\nu_{13}/E_{33} & 0 & 0 & 0 \\ & 1/E_{22} & -\nu_{23}/E_{33} & 0 & 0 & 0 \\ & & 1/E_{33} & 0 & 0 & 0 \\ & & & 1/G_{12} & 0 & 0 \\ \text{Sym.} & & & & 1/G_{23} & 0 \\ & & & & & 1/G_{13} \end{bmatrix} \begin{Bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{13} \end{Bmatrix}, \quad (2)$$

where  $E_{11}$ ,  $E_{22}$ , and  $E_{33}$  denote the effective Young's moduli in the three axial directions;  $\nu_{12}$ ,  $\nu_{23}$ , and  $\nu_{13}$  the Poisson's ratios; and  $G_{12}$ ,  $G_{23}$ , and  $G_{13}$  the shear moduli. The stiffness discussed in this paper refers to the moduli  $E_{11}$ ,  $E_{22}$ , and  $E_{33}$ . Note that these moduli can also be calculated from the elastic constants as [Poutet et al. 1996]

$$E_{11} = C_{11} - \frac{2C_{12}^2}{C_{11} + C_{12}}. \quad (3)$$

Further, for a material with cubic symmetry, that is, with three mutually perpendicular symmetry planes ( $E_{11} = E_{22} = E_{33} = E$ ,  $\nu_{12} = \nu_{23} = \nu_{13} = \nu$ , and  $G_{12} = G_{23} = G_{13} = G$ ), such as the Menger sponge, the effective stiffness refers to the unique effective Young's modulus  $E$ .

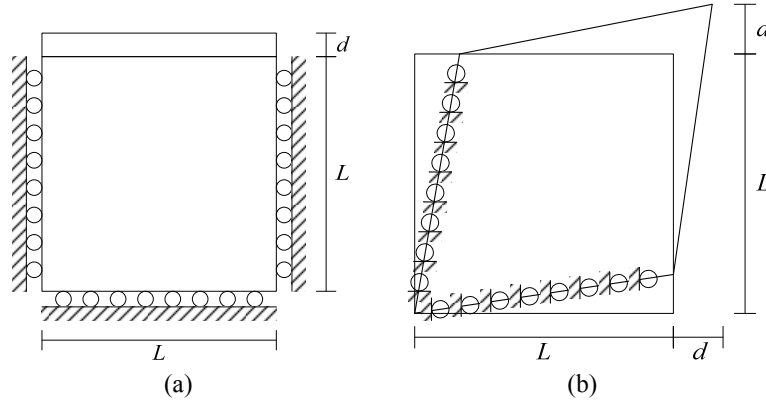
**2.2. Numerical homogenization through designed boundaries on unit cell.** A simple description of the numerical homogenization technique used in this study is presented here for 3D macroscopic mechanical characterization of an arbitrary structure or material microstructure. More details can be found in [Steven 1997; Tan et al. 2000].

The symmetry of the elastic stiffness matrix in (1) leaves 21 unknown constants to be determined. With six specified strain vectors, each with one strain component being unit and others being zero, we can calculate six elastic matrix constants easily. For example,

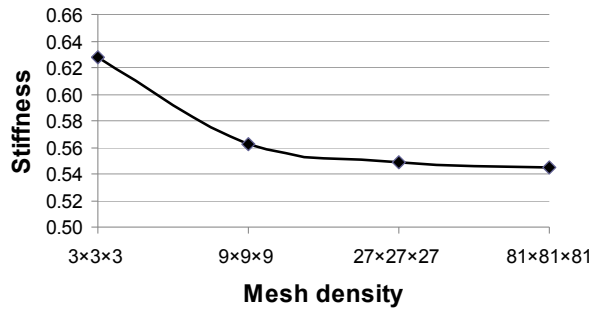
$$\begin{Bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{13} \end{Bmatrix} = \begin{Bmatrix} C_{11} \\ C_{21} \\ C_{31} \\ C_{41} \\ C_{51} \\ C_{61} \end{Bmatrix}, \quad \text{using the input strain } \boldsymbol{\epsilon}^* = \begin{Bmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ \epsilon_{12} \\ \epsilon_{23} \\ \epsilon_{13} \end{Bmatrix} = \begin{Bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix}. \quad (4)$$

In the numerical realization, six finite element analyses are required with unit strains expressed as prescribed displacements on the boundaries in order to get the corresponding stresses determined from the reaction forces. Two types of boundary conditions are involved: one with normal strain and the other with shear strain. Figure 3 demonstrates the two types of boundary conditions.

This simple and straightforward approach is capable of obtaining highly accurate results for the macroscopic mechanical properties of materials and structures, as suggested in [Steven 1997; Tan et al. 2000]. However, a sufficiently fine mesh is needed if the unit cell contains complex geometries such as arbitrary cavities or multiple phases. As an example, Figure 4 illustrates the effective stiffness obtained from



**Figure 3.** Boundary conditions used in the numerical homogenization: with (a) normal strain and (b) shear strain.



**Figure 4.** Numerical results of the stiffness of a level-one Menger sponge using different finite element meshes.

numerical homogenization using various finite element meshes. The chosen microstructure is the level-one Menger sponge shown in Figure 1 with a base material of  $E_0 = 1$  and  $\nu = 0.2$ . The convergence of the stiffness results through mesh refinement is shown below. It is seen that even for the relatively simple shape of the level-one Menger sponge, a dense mesh (at least  $27 \times 27 \times 27$  or, to be more prudent,  $81 \times 81 \times 81$ ) should be used in order to obtain an accurate stiffness result.

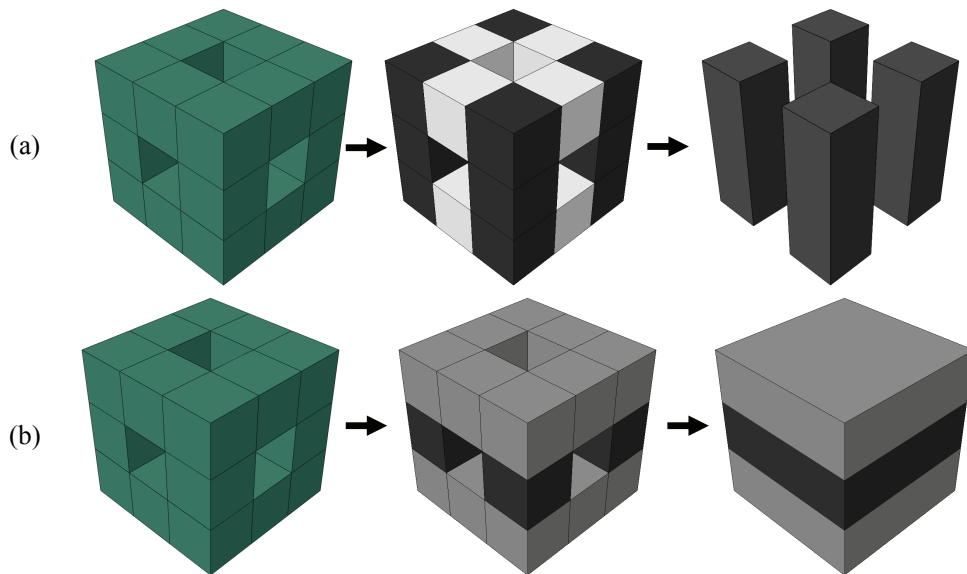
### 3. Derivation of a stiffness prediction scheme

**3.1. The renormalization argument.** Several simplified models were presented in [Poutet et al. 1996] for stiffness predictions of various hierarchical media, including the Menger sponge. The basic idea of the renormalization argument is to decompose the Menger sponge into three homogeneous layers. Each layer’s stiffness is simply assumed to be proportional to the cross-sectional area. The stiffness for the whole sponge can then be calculated through a simple force and deformation relationship. However, the prediction failed to match with the numerical results, in which the stiffness ratio between two adjacent levels ranged from 0.53 to 0.58, significantly different from the renormalization prediction of  $2/3$ .

If we look carefully at the simplified model of the Menger sponge within the renormalization argument, we will find that the force from one face to the opposite face is not transmitted evenly through the three layers, that is, stress concentrations will be found at the actual connections between layers. In this sense, the simplified model overestimates the effective stiffness.

**3.2. Stiffness prediction through a bounded estimation scheme.** Due to the nonuniform geometry, it is difficult or inaccurate to use only one simplified model to describe the force and deformation relationship of a hierarchical material such as the Menger sponge. However, one may set up multiple simplified models to predict the real value within a certain range. In this vein, we propose a scheme to calculate the stiffness based on the under and overestimates.

As an example, the Menger sponge can be simplified in two different ways, as shown in Figure 5. In the first model, eight white-colored blocks in the top and bottom layers are ignored, as shown in Figure 5(a). This is based on the assumption that under vertical pressure loading on the top and bottom surfaces these eight blocks are only weakly stressed. In this case, the Menger sponge is simplified and transformed into a parallel connection of four columns in the force direction. Since this parallel decomposition ignores the existence of the removed parts and thus ignores their stiffness, this simplified model underestimates the actual stiffness of the Menger sponge. The underestimate of stiffness is approximated as proportional to the cross-sectional area. Note that each constitutive block of a level- $n$  Menger sponge is a level- $(n - 1)$  Menger sponge due to the self-similarity, and that the stiffness of a full cube model with all such blocks is naturally equal to the stiffness of a level- $(n - 1)$  Menger sponge. Since four out of the nine columns of the full cube model are left, the stiffness of a level- $n$  Menger sponge can be simply determined as  $4/9$  of that of the level- $(n - 1)$  Menger sponge. In (5) the subscript denotes the level and the superscript



**Figure 5.** Simplified models of the level-one Menger sponge: (a) parallel connection model and (b) serial-connection model.

denotes the underestimate:

$$E_n^- = \frac{4E_{n-1}}{9}. \quad (5)$$

The aforementioned renormalization argument is referred to for the second model, which consists of three homogenized layers. This decomposition using serial connections, as in Figure 5(b), overestimates the stiffness as mentioned earlier and as supported by the numerical results from [Poutet et al. 1996]. The effective Young's moduli of each layer can be simplified as  $8E_{n-1}/9$ ,  $4E_{n-1}/9$ , and  $8E_{n-1}/9$ , respectively. With  $p$  being the pressure acting on the top and bottom faces of the Menger sponge, and  $a$  being the thickness of each layer, the normal strain is calculated as

$$\begin{aligned} \epsilon_{xx} &= \frac{p}{E_n} = \frac{d_{xx}}{3a} = \frac{1}{3} \left( \frac{d_{xx,1} + d_{xx,2} + d_{xx,3}}{a} \right) = \frac{1}{3} (\epsilon_{xx,1} + \epsilon_{xx,2} + \epsilon_{xx,3}) \\ &= \frac{1}{3} \left( \frac{p}{8E_{n-1}/9} + \frac{p}{4E_{n-1}/9} + \frac{p}{8E_{n-1}/9} \right), \end{aligned} \quad (6)$$

where the second subscript in the displacement and strain denotes the layer number. Hence we obtain the effective stiffness of the overestimation model as

$$E_n^+ = \frac{2E_{n-1}}{3}. \quad (7)$$

The Reuss [1929] bound and the Voigt [1889] bound have been believed to be the lower and upper bounds of the effective stiffness of composites. These two bounds correspond to the serial and parallel-connection models, respectively, which seems contrary to the proposal in this paper. However, being different from the Reuss and Voigt models, the model decomposition in this paper is valid also for cellular hierarchical materials, and building blocks are deliberately removed or added to form a weaker or stronger simplified model for analytical calculation of the stiffness. As a result, artificial upper and lower estimates are obtained that actually work as the bounds for the effective stiffness.

The overestimated stiffness is obviously larger than the underestimated stiffness. The real value of the stiffness is deemed to lie between the above under and overestimates. Hence we can set up the following framework for an effective approximation using an interpolation scheme:

$$E_n = (1 - \alpha)E_n^- + \alpha E_n^+, \quad \text{with } 0 \leq \alpha \leq 1, \quad (8)$$

and particularly for the Menger sponge:

$$E_n = \frac{2}{3}\alpha E_{n-1} + \frac{4}{9}(1 - \alpha)E_{n-1} = \frac{4+2\alpha}{9} E_{n-1}. \quad (9)$$

The interpolation parameter  $\alpha$  in the above formula applies to all levels of the hierarchy in theory. Therefore  $\alpha$  can be simply calculated by substituting the stiffnesses of levels-zero and one Menger sponges. A level-zero Menger sponge refers to the solid, isotropic base material. The stiffness of the level-one Menger sponge is obtained numerically using the numerical homogenization approach introduced in Section 2. Due to the cavities in the level-one Menger sponge, a very fine finite element mesh  $3^4 \times 3^4 \times 3^4$  is used in finite element analysis (FEA) to get accurate results, as mentioned earlier. With  $E_0 = 1$  and  $\nu_0 = 0.2$ , the FEA result gives  $E_1 = 0.5450$ . From (9),  $\alpha$  is obtained as 0.4525. Therefore, the stiffness

Level	Numerical	$E_n^-$	$E_n^+$	Predicted	Difference	[Poutet et al. 1996]	RA
1	0.5450	0.4444	0.6666	0.5450	0.00%	0.53298	0.6666
2	0.3077	0.1975	0.4444	0.2970	3.48%	0.29771	0.4444
3	0.1718	0.0878	0.2963	0.1619	5.76%	0.1737	0.2963
4	0.0960	0.0390	0.1975	0.0882	8.13%	N/A	0.1975

**Table 1.** Summary of stiffnesses of Menger sponges of different levels, with  $E_0 = 1$  and  $\nu_0 = 0.2$ .

prediction formula is further simplified as

$$E_n = 0.5450E_{n-1} \quad \text{or} \quad E_n = 0.5450^n E_0. \quad (10)$$

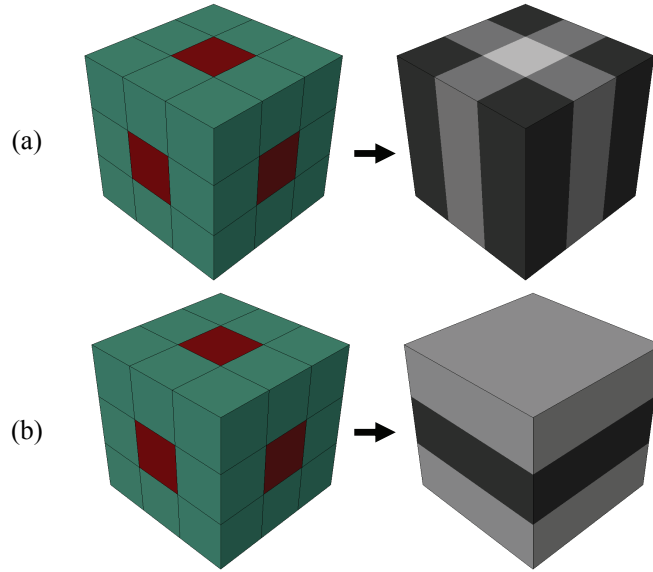
In order to examine the accuracy of the proposed prediction of effective stiffness, further levels of the Menger sponge have been solved using numerical homogenization. The numerical results are summarized in Table 1 with comparison to the under and overestimated stiffnesses, the predicted stiffness values using (10), the numerical results of [Poutet et al. 1996], and the predictions of the renormalization argument (RA).

Through the hierarchical level increasing, the stiffness under and overestimates differ more and more from the numerical results due to the built-up errors caused by the model simplifications. Nevertheless, the bounds give a practical estimate of the effective stiffness before any accurate homogenization technique is applied. The renormalization argument actually gives an upper bound of the effective stiffness. The predicted effective stiffnesses match well with the numerical results, and with the previous numerical results of [Poutet et al. 1996]. In contrast, the renormalization argument gives values that are at least 20% different from the numerical results. Due to limited computer power in the 1990s, Poutet et al. [1996] only solved the first three levels of the Menger sponge. In our study, the meshes we have used for levels-one to three Menger sponges are  $3^4 \times 3^4 \times 3^4$ ,  $3^4 \times 3^4 \times 3^4$ , and  $3^5 \times 3^5 \times 3^5$ , respectively. Since the Menger sponge is cubic symmetric, a one-eighth model is used for level four with a  $(2 \times 3^5) \times (2 \times 3^5) \times (2 \times 3^5)$  mesh to balance the numerical accuracy and computational cost. Note that due to voids in the Menger sponges, the actual numbers of solid elements in the above four meshes are  $3^4 \times 3^4 \times 3^4 \times 20/27$ ,  $3^4 \times 3^4 \times 3^4 \times (20/27)^2$ ,  $3^5 \times 3^5 \times 3^5 \times (20/27)^3$ , and  $(2 \times 3^5) \times (2 \times 3^5) \times (2 \times 3^5) \times (20/27)^4$ . It is seen that the basic building block — a unit structure with a level-one Menger sponge pattern — is much more coarsely meshed in the level-four Menger sponge than in the level-one Menger sponge. In other words, an extremely fine mesh,  $3^7 \times 3^7 \times 3^7$ , should be used for level four in order to achieve an equal numerical accuracy to level one. However even a supercomputer would have difficulty in handling such a fine mesh. Nevertheless, if such a fine mesh is achieved, the numerical stiffness is expected to decrease as indicated in Figure 4. In this case, it is envisaged that the numerical results will match the predicted values even better.

#### 4. Stiffness prediction for hierarchical composites

Stiffness of hierarchical composites can be predicted in the same way as for hierarchical media of a single base material as discussed in Section 3. Taking the Menger sponge composite in Figure 6 for example,





**Figure 6.** Simplified models for a Menger sponge composite of level one: (a) parallel-connection model and (b) serial-connection model.

two simplified models are similarly introduced for stiffness estimates. The composite is assumed to have two base materials, with  $E_m$  being the Young's modulus for the matrix base material (green) and  $E_i$  for the inclusion material (red). Note that a level- $n$  Menger sponge composite of stiffness  $E_n$  consists of 20 blocks of level- $(n - 1)$  Menger sponge composite of  $E_{n-1}$  and seven blocks of pure inclusion base material of  $E_i$ .

The first simplified model decomposes the level- $n$  Menger sponge composite into nine parallel columns. The Young's moduli for the four corner columns and the central column are  $E_{n-1}$  and  $E_i$ , respectively, while the Young's modulus of the remaining four columns is estimated by a simple force and deformation relationship. Suppose a pressure  $p$  is applied on one of these four columns that is composed of three blocks with Young's moduli  $E_{n-1}$ ,  $E_i$ , and  $E_{n-1}$ , respectively, the strain will be calculated as  $(p/E_{n-1} + p/E_i + p/E_{n-1})/3$ . On the other hand, the strain can also be calculated as the pressure divided by the homogenized Young's modulus of the whole column. Therefore the effective Young's modulus for these columns can be obtained based on a simple calculation and given as  $3E_{n-1}E_i/(E_{n-1} + 2E_i)$ . Finally, we have the first estimate of the effective stiffness of the whole composite as

$$E_n^- = \frac{1}{9} \left( 4E_{n-1} + E_i + 12 \frac{E_{n-1}E_i}{E_{n-1} + 2E_i} \right). \quad (11)$$

Similarly as in the single base material case, the second simplified model decomposes a level- $n$  Menger sponge composite into three serial-connected homogenized layers. With a simplification of stiffness proportional to the cross-sectional areas, the estimated Young's moduli for the three layers are  $8E_{n-1}/9 + E_i/9$ ,  $4E_{n-1}/9 + 5E_i/9$ , and  $8E_{n-1}/9 + E_i/9$ , respectively. With  $p$  being the pressure applied to the top and bottom faces of the composite, the total deformation in the force direction is calculated similarly as

Level	Numerical	$E_n^-$	$E_n^+$	Predicted	Difference
1	0.5647	0.4586	0.6712	0.5647	0.00%
2	0.3311	0.2177	0.4520	0.3228	2.51%
3	0.2000	0.1101	0.3059	0.1884	5.80%

**Table 2.** Summary of stiffnesses of Menger sponge composites with  $E_0 = 1$ ,  $\nu_0 = 0.2$ ,  $E_i = 0.01$ , and  $\nu_i = 0.4$ .

Numerical	Predicted	Difference
$4.3616 \times 10^{10}$	$4.3616 \times 10^{10}$	0.00%
$2.7862 \times 10^{10}$	$2.7428 \times 10^{10}$	1.56%
$1.8235 \times 10^{10}$	$1.7490 \times 10^{10}$	4.09%

**Table 3.** Summary of stiffnesses of Menger sponge composites with  $E_0 = 70 \times 10^9$ ,  $\nu_0 = 0.33$ ,  $E_i = 1 \times 10^9$ , and  $\nu_i = 0.49$ .

in (6):

$$\epsilon_{xx} = \frac{p}{E_n} = \frac{1}{3} \left( \frac{p}{8E_{n-1}/9 + E_i/9} + \frac{p}{4E_{n-1}/9 + 5E_i/9} + \frac{p}{8E_{n-1}/9 + E_i/9} \right). \quad (12)$$

Hence we obtain the second estimated stiffness as

$$E_n^+ = \frac{32E_{n-1}^2 + 44E_{n-1}E_i + 5E_i^2}{48E_{n-1} + 33E_i}. \quad (13)$$

Within the framework established by (8), we get the stiffness prediction

$$E_n = \frac{1}{9}(1 - \alpha) \left( 4E_{n-1} + E_i + 12 \frac{E_{n-1}E_i}{E_{n-1} + 2E_i} \right) + \alpha \frac{32E_{n-1}^2 + 44E_{n-1}E_i + 5E_i^2}{48E_{n-1} + 33E_i}. \quad (14)$$

It is worth pointing out that the Menger sponge is a special case of the Menger sponge composite. If the inclusion base material is replaced with void, that is,  $E_i = 0$ , the stiffness prediction in (14) reduces to that in (9).

The parameter  $\alpha$  is determined by the stiffnesses of levels-zero and one Menger sponge composites, where accurate results can be obtained with feasible fine meshes. In fact, no computation is required for level zero as it refers to the solid matrix base material. In the first case, it is assumed that the matrix base material has  $E_0 = 1$  and  $\nu_0 = 0.2$ , and the inclusion base material has  $E_i = 0.01$  and  $\nu_i = 0.4$ . The FEA result gives  $E_1 = 0.5647$ . From (14),  $\alpha$  is obtained as 0.4987. Table 2 presents the stiffnesses of various Menger sponge composites composed of these two base materials, with comparison between the numerical values and the predicted stiffnesses from (14). Meshes of  $3^3 \times 3^3 \times 3^3$ ,  $3^4 \times 3^4 \times 3^4$ , and  $3^5 \times 3^5 \times 3^5$  are used for levels one to three, respectively. In another case, the composite is assumed to have aluminum as the matrix base material, with  $E_0 = 70 \times 10^9$  and  $\nu_0 = 0.33$ , and silicon rubber as the inclusion base material, with  $E_i = 1 \times 10^9$  and  $\nu_i = 0.49$ . For the latter case,  $\alpha$  is obtained as 0.7599 and the results summary is given in Table 3.

More computational effort is needed for composites than for single-base material media because the finite element model is always a full model without any voids. Therefore only the first three levels of the Menger sponge composites are solved. Again through the hierarchical level increasing, the stiffness under and overestimates expand to make a wider and wider range into which the numerical result falls. The predicted stiffness values match well with the numerical results, even when the inclusion base material is almost incompressible, with its Poisson's ratio close to 0.5. This finding is of special importance because when one of the phases of the composite is near its incompressibility limit, it is usually necessary to use a different and more sophisticated analytical model in order to predict the effective stiffness accurately [Liu et al. 2006; 2009]. However, our proposed simple formula in (14) does not have this problem.

### 5. Stiffness prediction for a different self-similar hierarchical medium

Similar to the Menger sponge, whose basic topology can be defined based on a cube divided into  $3 \times 3 \times 3$  blocks, various other self-similar hierarchical media can be created by specifying different recursive rules for block removal. We propose an alternative hierarchical "sponge" that is obtained by removing eight corner blocks as well as the central one, as shown in Figure 7. Since this hollow material presents a cross shape in all three faces, we name it the "hollow cross sponge". The base material is selected as  $E_0 = 1$  and  $\nu_0 = 0.2$ .

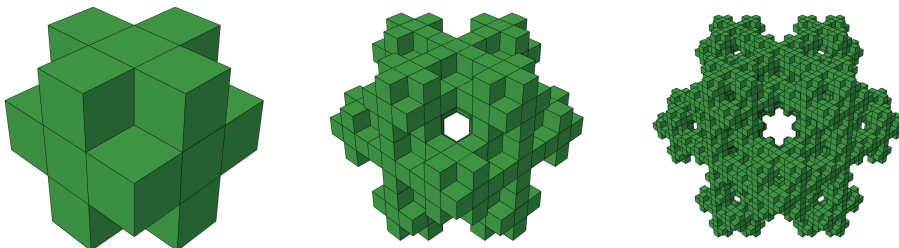
To predict the stiffness of such a hierarchical material, two simplified models in Figure 8 with parallel and serial connections are employed, as we have done with the Menger sponge.

In the first simplified model, the whole structure is decomposed into four columns since the other parts form no direct connection in the axial force direction. The parts removed make the simplified model underestimate the actual stiffness. Similarly as in the case described in (5), four out of nine columns of the full-cube model are left and the stiffness of this underestimation model is easily determined based on the cross-sectional area:

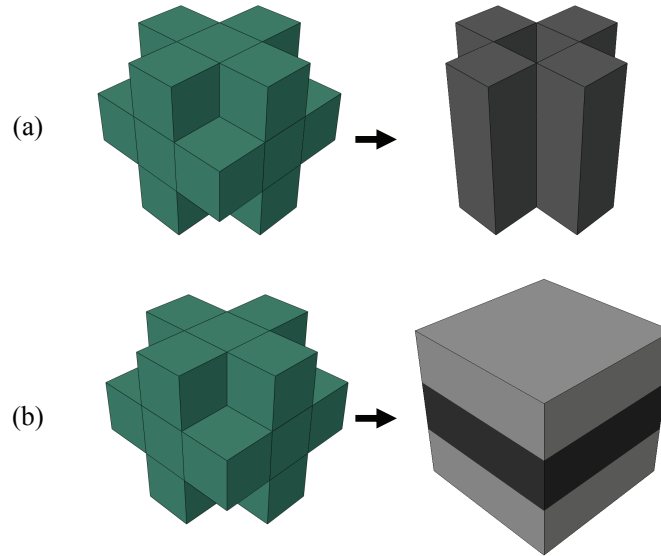
$$E_n^- = \frac{4E_{n-1}}{9}. \quad (15)$$

For the serial-connection model, the Young's moduli of the three homogenized layers are estimated as  $5E_{n-1}/9$ ,  $8E_{n-1}/9$ , and  $5E_{n-1}/9$ , respectively, based on the areas. In a simple derivation of the force and deformation relationship as in (6), the overestimated stiffness of the whole sponge is found as

$$E_n^+ = \frac{40E_{n-1}}{63}. \quad (16)$$



**Figure 7.** The first three levels of an alternative hierarchical material: the hollow cross sponge.



**Figure 8.** Simplified models of the hollow cross sponge, with all eight corner blocks and the central block removed: (a) parallel-connection model and (b) serial-connection model.

In the framework set up in (8), the effective stiffness of the sponge is predicted as

$$E_n = \frac{4}{9}(1 - \alpha)E_{n-1} + \frac{40}{63}\alpha E_{n-1} = \frac{28+12\alpha}{63}E_{n-1}. \quad (17)$$

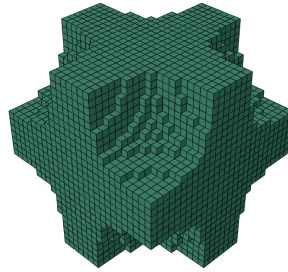
Assume that the base material has  $E_0 = 1$  and  $\nu_0 = 0.2$ . The parameter  $\alpha$  can be determined from  $E_1$ , which is found to be equal to 0.5087 from the numerical homogenization applied to the level-one sponge in Figure 7. This results in  $\alpha = 0.3373$ . Table 4 shows the stiffnesses of various levels of the hollow cross sponge using  $E_0 = 1$  and  $\nu_0 = 0.2$  for the purpose of examining the accuracy of the predicted effective stiffness. It is seen that the under and overestimates give effective bounds for the effective stiffness, and that the predictions from the simple formula in (17) are reasonably accurate.

## 6. Proposing a lighter yet stiffer material than the Menger sponge

The stiffness of a hierarchical medium depends on the topology and the amount of base material. To find a design with a higher stiffness for a predefined amount of material usage, topology optimization techniques can be employed. Figure 9 shows an optimal material base cell for maximizing the effective

Level	Numerical	$E_n^-$	$E_n^+$	Predicted	Difference
1	0.5087	0.4444	0.6349	0.5087	0.00%
2	0.2538	0.1975	0.4031	0.2588	1.98%
3	0.1253	0.0878	0.2560	0.1316	5.05%
4	0.0617	0.0390	0.1625	0.0670	8.63%

**Table 4.** Summary of stiffnesses of the first four levels of the hollow cross sponge.



**Figure 9.** Optimal topology for maximizing the effective stiffness for a given amount of volume.

stiffness we have obtained using the ESO/BESO topology optimization technique [Xie and Steven 1993; 1997; Huang and Xie 2007; 2010; Yang et al. 2013]. The optimization is based on the volume fraction of  $19/27$  and the cubic symmetric constraint, that is, the effective stiffnesses in the three directions are set equal to each other. The elastic properties of the base material are  $E_0 = 1$  and  $\nu_0 = 0.2$ . The effective stiffness of this design is 0.5864.

Informed by the above optimal topology, we propose to construct a new sponge as shown in Figure 10. Starting from a cube with  $3 \times 3 \times 3$  blocks, we remove the eight corner blocks. This will create a design of a similar topology to, and with the same volume as, the optimal solution shown in Figure 9. Since this model presents a cross in a 2D view from all three faces, we name it the “cross sponge”. From this basic configuration, we can recursively create a series of self-similar sponges of any level.

The under and overestimates of the cross sponge stiffness can be derived from the two simplified models shown in Figure 11:

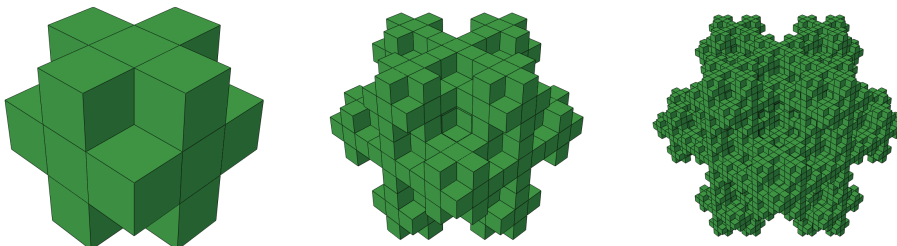
$$E_n^- = \frac{5E_{n-1}}{9}, \tag{18}$$

$$E_n^+ = \frac{15E_{n-1}}{23}, \tag{19}$$

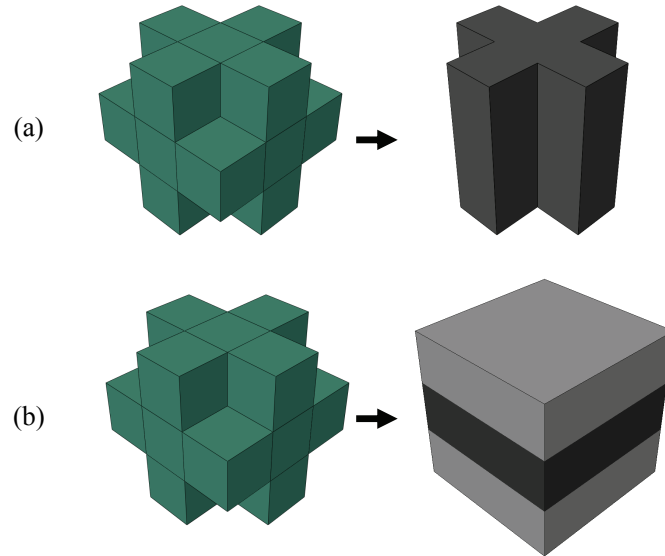
and the stiffness prediction formula is defined as

$$E_n = \frac{5}{9}(1 - \alpha)E_{n-1} + \frac{15}{23}\alpha E_{n-1} = \frac{115 + 20\alpha}{207}E_{n-1}. \tag{20}$$

The parameter  $\alpha$  is then determined from the level-zero (assuming  $E_0 = 1$  and  $\nu_0 = 0.2$ ) and level-one sponges ( $E_1 = 0.5816$  from numerical results). This results in  $\alpha = 0.2696$ . The stiffness of the cross



**Figure 10.** The proposed cross sponge, which is lighter yet stiffer than the Menger sponge.



**Figure 11.** Simplified models of the cross sponge: (a) parallel-connection model and (b) serial-connection model.

Level	Numerical	$E_n^-$	$E_n^+$	Predicted	Difference
1	0.5816	0.5556	0.6522	0.5816	0.00%
2	0.3321	0.3086	0.4253	0.3383	1.87%
3	0.1914	0.1715	0.2774	0.1967	2.77%
4	0.1085	0.0953	0.1809	0.1144	5.44%

**Table 5.** Summary of stiffnesses of the first four levels of the cross sponge.

sponge is higher than that of the Menger sponge (0.5816 versus 0.5450) even though the cross sponge is lighter (19/27 versus 20/27). Table 5 compares the stiffnesses of various levels of the cross sponge. It is seen that the predicted stiffnesses using the simple formula in (20) are quite accurate.

## 7. Concluding remarks

This paper has presented an analytical approach for stiffness prediction of hierarchical media. Unlike the previous renormalization argument, which uses only one simplified model and thus yields highly inaccurate stiffness predictions, the proposed approach employs two simplified models that under and overestimate the effective stiffness. The predicted stiffness is calculated as an interpolation between the under and overestimates. As the prediction presents the relationship between the stiffness of two adjacent levels, the interpolation parameter is determined based on the two lowest levels for which accurate numerical values can be easily obtained. The proposed approach for stiffness prediction is applicable to a whole class of self-similar hierarchical media, including composites, as demonstrated by a range of examples in this paper. The simple formulas developed have been found to give accurate predictions of stiffnesses of self-similar hierarchical media of various levels. For composites, the predicted effective

stiffness is accurate even when one of the phases is near its incompressibility limit, with its Poisson's ratio close to 0.5.

Inspired by the Menger sponge and informed by our topology optimization result, we have proposed a lighter yet stiffer cross sponge.

It is interesting to note that Lakes [1993] predicted that the effective stiffness of hierarchical materials should follow the relationship

$$\frac{E_n}{E_0} = k^n \left[ \frac{\rho_n}{\rho_0} \right]^r, \quad (21)$$

where  $k$  and  $r$  are coefficients depending on the type of structure,  $\rho$  is the density, and  $\rho_0$  is the density of the solid phase. If we look at (9), (17), and (20), we will find that the proposed stiffness prediction formulas agree with Lakes' assumption, given in (21). Take the Menger sponge as an example. The mass density can be calculated as

$$\rho_n = \left( \frac{20}{27} \right)^n \rho_0. \quad (22)$$

Thus (10) may be transformed into the same format as that of Lakes' formula if we assign  $r = 1$  and  $k = 0.7358$ :

$$\frac{E_n}{E_0} = 0.7358^n \left[ \frac{\rho_n}{\rho_0} \right] = 0.5450^n. \quad (23)$$

In fact, according to the proposed scheme based on the simplified models in this paper, the stiffness prediction for hierarchical media with one base material can be generalized into

$$\frac{E_n}{E_0} = \beta^n, \quad (24)$$

where  $\beta$  is a constant. One can determine the  $\beta$  value to make the prediction explicit; on the other hand, one may also use the effective stiffnesses of the simplified models proposed in this paper as bounds for a rough estimation.

## References

- [Aizenberg et al. 2005] J. Aizenberg, J. C. Weaver, M. S. Thanawala, V. C. Sundar, D. E. Morse, and P. Fratzl, "Skeleton of *Euplectella* sp.: structural hierarchy from the nanoscale to the macroscale", *Science* **309**:5732 (2005), 275–278.
- [Bakhvalov and Panasenko 1984] N. Bakhvalov and G. Panasenko, *Осреднение процессов в периодических средах*, Nauka, Moscow, 1984. Translated as *Homogenisation: averaging processes in periodic media*, Mathematics and its Applications (Soviet Series) **36**, Kluwer Academic, Dordrecht, 1989.
- [Bergman and Kantor 1984] D. J. Bergman and Y. Kantor, "Critical properties of an elastic fractal", *Phys. Rev. Lett.* **53**:6 (1984), 511–514.
- [Cleary et al. 1980] M. P. Cleary, S.-M. Lee, and I.-W. Chen, "Self-consistent techniques for heterogeneous media", *J. Eng. Mech. (ASCE)* **106** (1980), 861–887.
- [Coussy 1991] O. Coussy, *Mécanique des milieux poreux*, Editions Technip, Paris, 1991.
- [Currey 1977] J. D. Currey, "Mechanical properties of mother of pearl in tension", *Proc. R. Soc. Lond. B* **196** (1977), 443–463.
- [Currey 1984] J. D. Currey, *The mechanical adaptations of bones*, Princeton University Press, 1984.
- [Day et al. 1992] A. R. Day, K. A. Snyder, E. J. Garboczi, and M. F. Thorpe, "The elastic moduli of a sheet containing circular holes", *J. Mech. Phys. Solids* **40** (1992), 1031–1051.

- [Gao 2010] H. Gao, “Learning from nature about principles of hierarchical materials”, pp. 65–68 in *Proceedings of the 3rd International Nanoelectronics Conference* (Hong Kong, 2010), edited by P. K. Chu, IEEE, Piscataway, NJ, 2010.
- [Garboczi and Day 1995] E. J. Garboczi and A. R. Day, “An algorithm for computing the effective linear elastic properties of heterogeneous materials: 3-D results for composites with equal phase Poisson ratios”, *J. Mech. Phys. Solids* **43**:9 (1995), 1349–1362.
- [Hashin and Shtrikman 1962] Z. Hashin and S. Shtrikman, “A variational approach to the theory of the elastic behaviour of polycrystals”, *J. Mech. Phys. Solids* **10** (1962), 343–352.
- [Hashin and Shtrikman 1963] Z. Hashin and S. Shtrikman, “A variational approach to the theory of the elastic behaviour of multiphase materials”, *J. Mech. Phys. Solids* **11** (1963), 127–140.
- [Huang and Xie 2007] X. Huang and Y. M. Xie, “Convergent and mesh-independent solutions for the bi-directional evolutionary structural optimization method”, *Finite Elem. Anal. Des.* **43**:14 (2007), 1039–1049.
- [Huang and Xie 2010] X. Huang and Y. M. Xie, *Evolutionary topology optimization of continuum structures: methods and applications*, Wiley, Chichester, 2010.
- [Jäger and Fratzl 2000] I. Jäger and P. Fratzl, “Mineralized collagen fibrils: a mechanical model with a staggered arrangement of mineral particles”, *Biophys. J.* **79**:4 (2000), 1737–1746.
- [Lakes 1993] R. Lakes, “Materials with structural hierarchy”, *Nature* **361**:6412 (1993), 511–515.
- [Liu et al. 2006] B. Liu, L. Zhang, and H. Gao, “Poisson ratio can play a crucial role in mechanical properties of biocomposites”, *Mech. Mater.* **38**:12 (2006), 1128–1142.
- [Liu et al. 2009] B. Liu, X. Feng, and S.-M. Zhang, “The effective Young’s modulus of composites beyond the Voigt estimation due to the Poisson effect”, *Compos. Sci. Technol.* **69**:13 (2009), 2198–2204.
- [Nemat-Nasser and Hori 1993] S. Nemat-Nasser and M. Hori, *Micromechanics: overall properties of heterogeneous materials*, North-Holland Series in Applied Mathematics and Mechanics **37**, Elsevier, Amsterdam, 1993.
- [Oshmyan et al. 2001] V. G. Oshmyan, S. A. Patlazhan, and S. A. Timan, “Elastic properties of Sierpinski-like carpets: finite element based formulation”, *Phys. Rev. E* **64**:5 (2001), Article ID #056108.
- [Picu and Soare 2009] R. C. Picu and M. A. Soare, “Mechanics of materials with self-similar hierarchical microstructure”, pp. 295–332 in *Multiscale modeling in solid mechanics: computational approaches*, edited by U. Galvanetto and M. H. F. Aliabadi, Imperial College Press, London, 2009.
- [Poutet et al. 1996] J. Poutet, D. Manzoni, F. Hage-Chehade, C. J. Jacquin, M. J. Bouteca, J.-F. Thovert, and P. M. Adler, “The effective mechanical properties of random porous media”, *J. Mech. Phys. Solids* **44**:10 (1996), 1587–1620.
- [Puxkandl et al. 2002] R. Puxkandl, I. Zizak, O. Paris, J. Keckes, W. Tesch, S. Bernstorff, P. Purslow, and P. Fratzl, “Viscoelastic properties of collagen: synchrotron radiation investigations and structural model”, *Phil. Trans. R. Soc. B* **357**:1418 (2002), 191–197.
- [Reuss 1929] A. Reuss, “Berechnung der Fließgrenze von Mischkristallen auf Grund der Plastizitätsbedingung für Einkristalle”, *Z. Angew. Math. Mech.* **9**:1 (1929), 49–58.
- [Sánchez-Palencia 1980] E. Sánchez-Palencia, *Non-homogeneous media and vibration theory*, Lecture Notes in Physics **127**, Springer, Berlin, 1980.
- [Steven 1997] G. P. Steven, “Homogenization of multicomponent orthotropic materials using FEA”, *Commun. Numer. Methods Eng.* **13**:7 (1997), 517–531.
- [Stratasys 2012] “Objet350 Connex multi-material 3D printer”, Stratasys, Eden Prairie, MN, 2012, <http://www.stratasys.com/3d-printers/design-series/precision/objet-connex350>.
- [Tan et al. 1999] P. Tan, L. Tong, and G. P. Steven, “Micromechanics models for the elastic constants and failure strengths of plain weave composites”, *Compos. Struct.* **47**:1–4 (1999), 797–804.
- [Tan et al. 2000] P. Tan, L. Tong, and G. P. Steven, “Behavior of 3D orthogonal woven CFRP composites, II: FEA and analytical modeling approaches”, *Compos. A Appl. Sci. Manuf.* **31**:3 (2000), 273–281.
- [Thovert et al. 1990] J. F. Thovert, F. Wary, and P. M. Adler, “Thermal conductivity of random media and regular fractals”, *J. Appl. Phys.* **68** (1990), 3872–3883.



- [Voigt 1889] W. Voigt, "Ueber die Beziehung zwischen den beiden Elasticitätsconstanten isotroper Körper", *Ann. Physik* **274**:12 (1889), 573–587.
- [Watt et al. 1976] J. P. Watt, G. F. Davies, and R. J. O'Connell, "The elastic properties of composite materials", *Rev. Geophys.* **14**:4 (1976), 541–563.
- [Xie and Steven 1993] Y. M. Xie and G. P. Steven, "A simple evolutionary procedure for structural optimization", *Comput. Struct.* **49**:5 (1993), 885–896.
- [Xie and Steven 1997] Y. M. Xie and G. P. Steven, *Evolutionary structural optimization*, Springer, London, 1997.
- [Yang et al. 2013] X. Y. Yang, X. Huang, J. H. Rong, and Y. M. Xie, "Design of 3D orthotropic materials with prescribed ratios for effective Young's moduli", *Comput. Mater. Sci.* **67** (2013), 229–237.
- [Zhang et al. 2011] Z. Zhang, Y.-W. Zhang, and H. Gao, "On optimal hierarchy of load-bearing biological materials", *Proc. R. Soc. Lond. B* **278**:1705 (2011), 519–525.

Received 27 Dec 2012. Revised 20 May 2013. Accepted 25 Jun 2013.

YI MIN XIE: [mike.xie@rmit.edu.au](mailto:mike.xie@rmit.edu.au)

*Centre for Innovative Structures and Materials, School of Civil, Environmental and Chemical Engineering, RMIT University, GPO Box 2476, Melbourne VIC 3001, Australia*

ZHI HAO ZUO: [zhihao.zuo@rmit.edu.au](mailto:zhihao.zuo@rmit.edu.au)

*Centre for Innovative Structures and Materials, School of Civil, Environmental and Chemical Engineering, RMIT University, GPO Box 2476, Melbourne VIC 3001, Australia*

XIAODONG HUANG: [huang.xiaodong@rmit.edu.au](mailto:huang.xiaodong@rmit.edu.au)

*Centre for Innovative Structures and Materials, School of Civil, Environmental and Chemical Engineering, RMIT University, GPO Box 2476, Melbourne VIC 3001, Australia*

XIAOYING YANG: [xiaoying.yang@rmit.edu.au](mailto:xiaoying.yang@rmit.edu.au)

*Centre for Innovative Structures and Materials, School of Civil, Environmental and Chemical Engineering, RMIT University, GPO Box 2476, Melbourne VIC 3001, Australia*



## ON ACOUSTOELASTICITY AND THE ELASTIC CONSTANTS OF SOFT BIOLOGICAL TISSUES

PHAM CHI VINH AND JOSE MERODIO

We analyze the acoustoelastic study of material moduli that appear in the constitutive relations that characterize the response of anisotropic nonlinearly elastic bodies, in particular, materials reinforced with one set of fibers along one direction. Studies dealing with acoustoelastic coefficients in incompressible solids modeled by means of strain-energy density functions expanded up to different orders in terms of the Green strain tensor can be found in the literature. In this paper, we connect that analysis and the parallel one developed from the general theory of nonlinear elasticity which is based on strain energies that depend on the right Cauchy–Green deformation tensor. Establishing this relation explicitly will improve understanding of the mechanical properties of soft biological tissues among other materials.

### 1. Introduction

Determination of the acoustoelastic coefficients in incompressible solids has very recently attracted a lot of attention since these analyses give an opportunity to capture the mechanical properties of such materials, among other applications [Bigoni et al. 2007; 2008; Destrade et al. 2010b].

An incompressible transversely isotropic model has recently been analyzed by Destrade et al. [2010a] in which the strain-energy density is given by

$$W = \mu I_2 + \frac{1}{3} A I_3 + \alpha_1 I_4^2 + \alpha_2 I_5 + \alpha_3 I_2 I_4 + \alpha_4 I_4^3 + \alpha_5 I_4 I_5, \quad (1)$$

where

$$I_2 = \text{tr}(\mathbf{E}^2), \quad I_3 = \text{tr}(\mathbf{E}^3), \quad I_4 = \mathbf{M} \cdot (\mathbf{E}\mathbf{M}), \quad I_5 = \mathbf{M} \cdot (\mathbf{E}^2\mathbf{M}), \quad (2)$$

$\mathbf{E}$  is the Green strain tensor,  $\mathbf{M}$  is the unit vector that gives the undeformed fiber direction, and  $\mu, \alpha_1, \alpha_2$  and  $A, \alpha_3, \alpha_4, \alpha_5$  are second and third-order elastic constants, respectively (the order is given by the exponent of  $\mathbf{E}$ ). To evaluate the elastic constants Destrade et al. established formulas for the velocity waves. The formulas are given as first-order polynomials in terms of the elongation  $e_1$ , which is defined by  $\lambda = 1 + e_1$ , where  $\lambda$  is the principal stretch in the direction that gives both the fiber direction and the direction of uniaxial tension. The speeds of infinitesimal waves do provide a basis for the acoustoelastic evaluation of the material constants [Destrade and Ogden 2010].

Soft biological tissues are anisotropic solids due to the presence of oriented collagen fiber bundles [Holzapfel et al. 2000; Destrade et al. 2010a]. To make the model (1) more general and able to capture

---

*Keywords:* incompressible transversely isotropic elastic solids, soft biological tissue, wave velocity, elastic constants.

soft biological tissue mechanical behavior a fourth-order incompressible strain-energy function has been analyzed in [Vinh and Merodio 2013], namely

$$W = \mu I_2 + \frac{1}{3} A I_3 + \alpha_1 I_4^2 + \alpha_2 I_5 + \alpha_3 I_2 I_4 + \alpha_4 I_4^3 + \alpha_5 I_4 I_5 + \alpha_6 I_2^2 + \alpha_7 I_2 I_4^2 + \alpha_8 I_2 I_5 + \alpha_9 I_4^4 + \alpha_{10} I_5^2 + \alpha_{11} I_3 I_4, \quad (3)$$

where  $\alpha_6, \dots, \alpha_{11}$  are fourth-order elastic constants. The results show that linear corrections to the acoustoelastic wave speed formulas involve second and third-order constants, and that quadratic corrections involve second, third, and fourth-order constants, in agreement with [Hoger 1999]. Indeed, this is precisely the rationale behind the considered expansions (1) and (3) and the reason to develop the acoustoelastic wave speed formulas in terms of constitutive models that depend on the invariants of the Green strain tensor. Nevertheless, in this paper we develop acoustoelastic wave speed formulas in terms of constitutive models that depend on the invariants of the right Cauchy–Green tensor.

The model (3) has 13 elastic constants. It would be perfectly justifiable to question the efficacy of a model that depends on such a number of elastic constants. It is not easy to determine the structure of these material constants by any correlation with experiments. This makes a careful scrutiny of the nature of these constitutive models necessary in order to determine which of these elastic constants must be retained in the development of models. The models have to be well understood, and we can only do this if we analyze their structure. In passing, we mention that there has been lately in the literature some controversy regarding the use of planar tests to characterize anisotropic nonlinearly elastic materials. For a general discussion refer to [Holzapfel and Ogden 2009].

The more available formulations there exist in the literature to characterize elastic materials the more possibilities researchers have to capture the structure of the constitutive models. Furthermore, while physical acousticians are interested in third-order constants for anisotropic solids, workers in nonlinear elasticity, and, in particular, in soft biological tissue, use finite extensions involving fourth-order constants. In addition, soft biological tissue is modeled using general nonlinear elasticity theory expressed in terms of the right Cauchy–Green deformation tensor (see [Holzapfel et al. 2000]), among other formulations. Therefore, to develop the acoustoelastic wave speed formulas in terms of general nonlinear elasticity theory, that is, in terms of the right Cauchy–Green tensor, may improve our understanding of the mechanical response of soft tissue, among other materials. It is to this aspect of the problem that this study is directed.

In this paper we only address the subtle differences between the two approaches. To the best of our knowledge this relation has not been explored in the literature and we believe that the cumbersome technical details of the analysis are worthy of investigation. Our purpose is twofold: on the one hand, we illustrate the analysis for these nonisotropic elastic energy functions; on the other, we connect the acoustoelastic formulations of both material models, the one depending on the Green strain tensor and the one depending on the right Cauchy–Green tensor.

The layout of the paper is as follows. In Section 2, we introduce briefly the main governing equations. Section 3 is devoted to the acoustoelastic analysis of constitutive models that depend on the invariants of the right Cauchy–Green strain tensor. In particular, the equations governing infinitesimal motions superimposed on a finite deformation have been used to establish formulas for the velocity of (plane homogeneous) shear bulk waves propagating in general soft biological tissues subject to uniaxial tension

or compression. Furthermore, the analysis connects with the constitutive model (3). In Section 4 we give some conclusions.

## 2. Overview of the main equations

We consider an elastic body whose initial geometry defines a *reference configuration*, which we denote by  $\mathcal{B}_r$ , and a finitely deformed equilibrium configuration  $\mathcal{B}_0$ . The position vectors of representative particles in  $\mathcal{B}_r$  and  $\mathcal{B}_0$  are denoted by  $\mathbf{X}$  and  $\mathbf{x}$ , respectively. It is well known that  $\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$ , where  $t$  is time. The deformation gradient tensor associated with the deformation  $\mathcal{B}_r \rightarrow \mathcal{B}_0$  is denoted by  $\mathbf{F}$ .

**2.1. Material model.** Soft tissue is modeled as an incompressible transversely isotropic elastic solid. The most general transversely isotropic nonlinear elastic strain-energy function  $\Omega$  depends on  $\mathbf{F}$  through the right Cauchy–Green tensor  $\mathbf{C}$ , which is  $\mathbf{C} = \mathbf{F}^T \mathbf{F}$ , and we therefore consider  $\Omega$  to depend on the invariants of the tensor  $\mathbf{C}$ . It is well known that  $\mathbf{E} = (\mathbf{C} - \mathbf{I})/2$ , where  $\mathbf{I}$  is the identity tensor. The isotropic invariants of  $\mathbf{C}$  most commonly used are the *principal* invariants, defined by

$$I_1^* = \text{tr } \mathbf{C}, \quad I_2^* = \frac{1}{2}[(\text{tr } \mathbf{C})^2 - \text{tr}(\mathbf{C}^2)], \quad I_3^* = \det \mathbf{C}. \quad (4)$$

The (anisotropic) invariants associated with  $\mathbf{M}$  and  $\mathbf{C}$  are usually taken as (see, for instance, [Merodio and Saccomandi 2006])

$$I_4^* = \mathbf{M} \cdot (\mathbf{C}\mathbf{M}), \quad I_5^* = \mathbf{M} \cdot (\mathbf{C}^2 \mathbf{M}). \quad (5)$$

It follows that for incompressible materials  $\Omega = \Omega(I_1^*, I_2^*, I_4^*, I_5^*)$  since  $I_3^* = 1$ .

In the Appendix several expressions that are needed in this analysis are given. The corresponding Cauchy stress tensor for  $\Omega$  using the relations (A.4) and (A.6) yields

$$\boldsymbol{\sigma} = -p\mathbf{I} + 2\Omega_1 \mathbf{B} + 2\Omega_2 (I_1^* \mathbf{B} - \mathbf{B}^2) + 2\Omega_4 \mathbf{m} \otimes \mathbf{m} + 2\Omega_5 (\mathbf{m} \otimes \mathbf{B}\mathbf{m} + \mathbf{B}\mathbf{m} \otimes \mathbf{m}), \quad (6)$$

where  $p$  is the hydrostatic pressure arising from the incompressibility constraint,  $\mathbf{B} = \mathbf{F}\mathbf{F}^T$ , and  $\mathbf{m} = \mathbf{F}\mathbf{M}$  gives the deformed fiber direction. This expression in indicial notation is

$$\sigma_{ij} = -p\delta_{ij} + 2\Omega_1 B_{ij} + 2\Omega_2 (I_1^* \delta_{ij} - B_{i\gamma} B_{j\gamma}) + 2\Omega_4 m_i m_j + 2\Omega_5 (B_{j\gamma} m_\gamma m_i + B_{i\gamma} m_\gamma m_j). \quad (7)$$

It follows using (7) that

$$p = 2\Omega_1, \quad \Omega_4 + 2\Omega_5 = 0, \quad (8)$$

in the reference configuration where  $\mathbf{F} = \mathbf{I}$  and the Cauchy stress components are zero.

**2.2. Linearized equations of motion.** The linearized equations of motion for incompressible materials are summarized below. For a complete derivation see [Ogden and Singh 2011]. From  $\mathcal{B}_0$ , the linearized incremental form of the equations of motion and the incompressibility constraint, in component form, are written as

$$\mathcal{A}_0 p_{iqj} u_{j,pq} - p^*_{,i} = \rho u_{i,tt}, \quad u_{i,i} = 0, \quad (9)$$

respectively, where  $u_i(\mathbf{x}, t)$  is the small time-dependent displacement increment, a comma indicates differentiation with respect to the spatial coordinate or with respect to  $t$ ,  $p^*$  is the time-dependent pressure increment, and

$$\mathcal{A}_{0piqj} = F_{p\alpha} F_{q\beta} \frac{\partial^2 \Omega}{\partial F_{i\alpha} \partial F_{j\beta}}.$$

The subscript 0 indicates the so-called pushed-forward quantity from the initial reference configuration to the finitely deformed equilibrium configuration. We give its specialization to the situation in which there is no finite deformation and  $\mathcal{B}_0$  coincides with  $\mathcal{B}_r$ . The elasticity tensor  $\mathcal{A}_{0piqj}$  is given in (A.7). Under these conditions, customarily, the subscript 0 on  $\mathcal{A}_0$  is omitted, and in what follows we do so.

**2.3. Homogeneous plane waves.** We apply the equation of motion and the incompressibility condition to the analysis of homogeneous plane waves. In particular, we consider the incremental displacement  $\mathbf{u}$  and Lagrange multiplier  $p^*$  to have the forms

$$\mathbf{u} = f(\mathbf{n} \cdot \mathbf{x} - vt)\mathbf{d}, \quad p^* = g(\mathbf{n} \cdot \mathbf{x} - vt), \quad (10)$$

where  $\mathbf{d}$  is a constant unit (polarization) vector, the unit vector  $\mathbf{n}$  is the direction of propagation of the plane wave,  $v$  is the wave speed,  $f$  is a function that need not be made explicit but is subject to the restriction  $f'' \neq 0$ , and  $g$  is a function related to  $f$ . A prime on  $f$  or  $g$  indicates differentiation with respect to its argument.

Substitution of (10) into (9) then yields

$$[\mathbf{Q}(\mathbf{n})\mathbf{d} - \rho v^2 \mathbf{d}]f'' - g'\mathbf{n} = \mathbf{0}, \quad \mathbf{d} \cdot \mathbf{n} = 0, \quad (11)$$

where the (symmetric) *acoustic tensor*  $\mathbf{Q}(\mathbf{n})$  is defined by

$$Q_{ij}(\mathbf{n}) = \mathcal{A}_{piqj} n_p n_q. \quad (12)$$

The elasticity tensor  $\mathcal{A}_{piqj}$  is given in (A.7). Now, we just give the main result. For a complete derivation see [Ogden and Singh 2011]. It follows that for a given  $\mathbf{n}$  and  $\mathbf{d}$  the wave speed  $v$  is obtained from

$$\rho v^2 = [\mathbf{Q}(\mathbf{n})\mathbf{d}] \cdot \mathbf{d}. \quad (13)$$

### 3. An approach to finding formulas for the speeds of homogeneous plane waves using general nonlinear elasticity theory

We now describe the loading and geometric case that will be used in the analysis that follows. Consider a rectangular block of a soft transversely isotropic incompressible elastic solid whose faces in the unstressed state are parallel to the  $(X_1, X_2)$ ,  $(X_2, X_3)$ , and  $(X_3, X_1)$ -planes and with the fiber direction  $\mathbf{M}$  parallel to the  $X_1$ -direction. Suppose that the sample is under uniaxial tension or compression with the direction of tension parallel to the  $X_1$ -axis. It is easy to see that the sample is subject to  $x_1 = \lambda_1 X_1$ ,  $x_2 = \lambda_2 X_2$ , and  $x_3 = \lambda_3 X_3$ , and whence

$$\mathbf{F} = \text{diag}(\lambda_1, \lambda_2, \lambda_3), \quad (14)$$

in which

$$\lambda_1 = \lambda, \quad \lambda_2 = \lambda_3 = \lambda^{-1/2}, \quad \lambda > 0, \quad (15)$$

where  $\lambda_k$  are the principal stretches of deformation. Note that the faces of the deformed block are parallel to the  $(x_1, x_2)$ ,  $(x_2, x_3)$ , and  $(x_3, x_1)$ -planes.

The analysis can now focus on different cases. We consider motion in the  $(x_1, x_2)$ -plane (a plane that contains the deformed fiber direction) with  $n_1 = \cos \theta = c$ ,  $n_2 = \sin \theta = s$ ,  $d_1 = -\sin \theta$ , and  $d_2 = \cos \theta$ , where  $\theta$  is the angle between  $\mathbf{n}$  and the  $x_1$ -direction. The wave speed is, under these conditions, obtained using (13), (12), and (A.7) and can be written as

$$\begin{aligned} \rho v^2 &= \mathcal{A}_{piqj} n_p n_q d_i d_j \\ &= \mathcal{A}_{1212} c^4 + 2(\mathcal{A}_{1222} - \mathcal{A}_{1112}) c^3 s + (\mathcal{A}_{1111} + \mathcal{A}_{2222} - 2\mathcal{A}_{1122} - 2\mathcal{A}_{1221}) c^2 s^2 \\ &\quad + 2(\mathcal{A}_{1121} - \mathcal{A}_{2221}) c s^3 + \mathcal{A}_{2121} s^4. \end{aligned} \quad (16)$$

We now assume the situation described in (14) and investigate propagation in the fiber direction, that is,  $\mathbf{n}$  coincides with the deformed fiber direction, which initially is in the  $X_1$ -direction. In this case, the relevant term in (16) is

$$\begin{aligned} \mathcal{A}_{1212} &= 2\Omega_1 \lambda_1^2 + 2\Omega_2 (I_1^* \lambda_1^2 - \lambda_1^2 \lambda_2^2 - \lambda_1^4) + 2\Omega_4 m_1^2 + 2\Omega_5 [2m_1 B_{1\gamma} m_\gamma + \lambda_2^2 m_1^2 + \lambda_1^2 m_2^2] \\ &\quad + 4\Omega_{44} m_1^2 m_2^2 + 2\Omega_{55} (m_1 B_{2\gamma} m_\gamma + m_2 B_{1\gamma} m_\gamma)^2 + 4\Omega_{45} 2m_1 m_2 (m_1 B_{2\gamma} m_\gamma + m_2 B_{1\gamma} m_\gamma) \\ &\quad + 2\Omega_1 \lambda_1^2 + 2\Omega_2 \lambda_1^2 \lambda_3^2 + 2\Omega_4 m_1^2 + 2\Omega_5 [2m_1 B_{1\gamma} m_\gamma + \lambda_1^2 m_2^2 + \lambda_2^2 m_1^2] \\ &\quad + 4\Omega_{44} m_1^2 m_2^2 + 4\Omega_{55} (m_1 B_{2\gamma} m_\gamma + m_2 B_{1\gamma} m_\gamma)^2 + 8\Omega_{45} m_1 m_2 (m_1 B_{2\gamma} m_\gamma + m_2 B_{1\gamma} m_\gamma). \end{aligned}$$

Considering, further, that in this case the components of  $\mathbf{m}$  are  $m_1 = \lambda_1$  and  $m_2 = 0$  the wave speed  $v_{12}$  is

$$\rho v_{12}^2 = \mathcal{A}_{1212} = 2\Omega_1 \lambda_1^2 + 2\Omega_2 \lambda_1^2 \lambda_3^2 + 2\Omega_4 \lambda_1^2 + 2\Omega_5 (2\lambda_1^4 + \lambda_1^2 \lambda_2^2). \quad (17)$$

On the other hand, and with a parallel argument, the wave speed for the analysis of propagation in the perpendicular-to-the-fiber direction denoted by  $v_{21}$ , that is,  $\mathbf{n}$  is perpendicular to the deformed fiber direction, which initially is in the  $X_1$ -direction, yields using (16)

$$\rho v_{21}^2 = \mathcal{A}_{2121} = 2\Omega_1 \lambda_2^2 + 2\Omega_2 \lambda_2^2 \lambda_3^2 + 2\Omega_5 \lambda_1^2 \lambda_2^2. \quad (18)$$

It follows using (7), particularized for the (uniaxial) conditions at hand, (17), and (18) that

$$\mathcal{A}_{1212} - \mathcal{A}_{2121} = \sigma_1 - \sigma_2. \quad (19)$$

We do not pursue here the study of these relations. For further details, we refer to [Ogden and Singh 2011], which also establishes connections between the identities given with the formulation developed here and the identities developed by Biot [1965] with his formulation.

Now, we consider  $\lambda_1 = 1 + e_1$  for  $e_1$  sufficiently small. By incompressibility  $\lambda_2 = \lambda_3 = \lambda_1^{-1/2}$ , and it follows that  $\lambda_1^2 \lambda_3^2 = \lambda_1^2 \lambda_2^2 = \lambda_1 = 1 + e_1$ . Using these expressions, (4), and (5), the invariants  $I_i^*$  in terms of  $e_1$  are

$$\begin{aligned} I_1^* &= \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = \lambda_1^2 + 2\lambda_2^2 = 1 + 2e_1 + e_1^2 + 2(1 + e_1)^{-1} = 3 + 3e_1^2, \\ I_2^* &= \lambda_1^{-2} + 2\lambda_2^{-2} = 1 - 2e_1 + 3e_1^2 + 2(1 + e_1) = 3 + 3e_1^2, \\ I_4^* &= \lambda_1^2 = 1 + 2e_1 + e_1^2, \\ I_5^* &= \lambda_1^4 = 1 + 4e_1 + 6e_1^2. \end{aligned} \quad (20)$$

Expansion of (17) requires the expansion of the different derivatives  $\Omega_i$ ,  $i = 1, 2, 4, 5$ , in (17). One finds using the chain rule and (20) that

$$\begin{aligned}\Omega_1 &= \Omega_1^{(o)} + e_1 \frac{d}{de_1} \Omega_1 \Big|_{e_1=0} + \frac{1}{2} e_1^2 \frac{d^2}{de_1^2} \Omega_1 \Big|_{e_1=0} \\ &= \Omega_1^{(o)} + e_1 [2\Omega_{14}^{(o)} + 4\Omega_{15}^{(o)}] + \frac{1}{2} e_1^2 \frac{d}{de_1} [6\Omega_{11} e_1 + 6\Omega_{12} e_1 + \Omega_{14}(2+2e_1) + \Omega_{15}(4+12e_1)] \Big|_{e_1=0} \\ &= \Omega_1^{(o)} + 2e_1 (\Omega_{14}^{(o)} + 2\Omega_{15}^{(o)}) + \frac{1}{2} e_1^2 (6\Omega_{11}^{(o)} + 6\Omega_{12}^{(o)} + 2\Omega_{14}^{(o)} + 2\Omega_{15}^{(o)} + 4\Omega_{144}^{(o)} + 16\Omega_{145}^{(o)} + 16\Omega_{155}^{(o)}),\end{aligned}$$

where  $\Omega_1^{(o)}$  is the value of  $\Omega_1$  in the reference configuration. Furthermore, it is easy to obtain that

$$\Omega_i = \Omega_i^{(o)} + 2(\Omega_{i4}^{(o)} + 2\Omega_{i5}^{(o)})e_1 + (3\Omega_{i1}^{(o)} + 3\Omega_{i2}^{(o)} + \Omega_{i4}^{(o)} + 6\Omega_{i5}^{(o)} + 2\Omega_{i44}^{(o)} + \Omega_{i45}^{(o)} + 8\Omega_{i55}^{(o)})e_1^2, \quad (21)$$

where  $i$  can take the values 1, 2, 4, and 5.

Use of (21) and (17) yields for the wave speed, disregarding terms of order higher than 2 in  $e_1$ ,

$$\begin{aligned}\rho v_{12}^2 &= 2\Omega_1^{(o)} + 2\Omega_2^{(o)} + 2\Omega_4^{(o)} + 6\Omega_5^{(o)} + e_1 [4\Omega_{11}^{(o)} + 2\Omega_{12}^{(o)} + 4\Omega_{14}^{(o)} + 18\Omega_{15}^{(o)} + 4\Omega_{144}^{(o)} + 8\Omega_{155}^{(o)} \\ &\quad + 4\Omega_{24}^{(o)} + 8\Omega_{25}^{(o)} + 4\Omega_{44}^{(o)} + 20\Omega_{45}^{(o)} + 24\Omega_{55}^{(o)}] + e_1^2 [2\Omega_1^{(o)} + 2\Omega_4^{(o)} + 24\Omega_5^{(o)} \\ &\quad + 8(\Omega_{14}^{(o)} + 2\Omega_{15}^{(o)}) + 4(\Omega_{24}^{(o)} + 2\Omega_{25}^{(o)}) + 8(\Omega_{44}^{(o)} + \Omega_{45}^{(o)}) + 36(\Omega_{45}^{(o)} + 2\Omega_{55}^{(o)}) + 6\Omega_{11}^{(o)} + 12\Omega_{12}^{(o)} \\ &\quad + 8\Omega_{14}^{(o)} + 30\Omega_{15}^{(o)} + 6\Omega_{22}^{(o)} + 8\Omega_{24}^{(o)} + 30\Omega_{25}^{(o)} + 2\Omega_{44}^{(o)} + 18\Omega_{45}^{(o)} + 36\Omega_{55}^{(o)} + 4\Omega_{144}^{(o)} + 16\Omega_{145}^{(o)} \\ &\quad + 16\Omega_{155}^{(o)} + 4\Omega_{244}^{(o)} + 16\Omega_{245}^{(o)} + 16\Omega_{255}^{(o)} + 4\Omega_{444}^{(o)} + 30\Omega_{445}^{(o)} + 64\Omega_{455}^{(o)} + 48\Omega_{555}^{(o)}]. \quad (22)\end{aligned}$$

This formula gives the general acoustoelastic wave speed for constitutive models that depend on the invariants of the right Cauchy–Green tensor. This completes the first purpose of our analysis. Now, we focus on the second purpose, which is to connect this formulation and the one developed for constitutive models that depend on invariants of the Green strain tensor.

The relations between both invariant formulations  $I_i^*$  and  $I_i$  are established through  $C$  and the well-known Cayley–Hamilton theorem, that we write as

$$C^3 = I_1^* C^2 - I_2^* C + I, \quad \text{tr}(C^3) = I_1^*(I_1^{*2} - 2I_2^*) - I_1^* I_2^* + 3.$$

Hence, the invariants  $I_i$  in terms of  $I_i^*$  are

$$\begin{aligned}I_2 &= \text{tr}(E^2) = \frac{1}{4} \text{tr}(C^2 - 2C + I) = \frac{1}{4}(I_1^{*2} - 2I_2^* - 2I_1^* + 3), \\ I_3 &= \text{tr}(E^3) = \frac{1}{8} \text{tr}(C^3 - 3C^2 + 3C + I) = \frac{1}{8}(I_1^{*3} - 3I_1^* I_2^* - 3I_1^{*2} + 6I_2^* + 3I_1^*), \\ I_4 &= M \cdot (EM) = \frac{1}{2}(I_4^* - 1), \\ I_5 &= M \cdot (E^2 M) = \frac{1}{4}(I_5^* - 2I_4^* + 1).\end{aligned} \quad (23)$$

Using (23) the constitutive model (3) (or any other model written in terms of the invariants  $I_i$ ) can be written in terms of the invariants  $I_i^*$  and the same constants  $\mu$ ,  $A$ , and  $\alpha_1, \dots, \alpha_{11}$ . Then, (22) for that



particular model after a lengthy but straightforward calculation yields

$$\begin{aligned} \rho v_{12}^2 = & \mu + \frac{1}{2}\alpha_2 + (3\mu + \frac{1}{4}A + 2\alpha_1 + \frac{5}{2}\alpha_2 + \alpha_3 + \frac{1}{2}\alpha_5)e \\ & + (5\mu + \frac{7}{4}A + 5\alpha_1 + 5\alpha_2 + 5\alpha_3 + 3\alpha_4 + \frac{15}{4}\alpha_5 + 3\alpha_6 + \alpha_7 + \frac{7}{4}\alpha_8 + \alpha_{10} + \frac{3}{4}\alpha_{11})e^2. \end{aligned} \quad (24)$$

This formula was obtained in [Vinh and Merodio 2013]. Furthermore, the result given in [Destrade et al. 2010a] is a special case of the approximation (24) when  $\alpha_k = 0$  and  $k = \overline{6, 11}$ .

#### 4. Conclusions

The motivation behind this analysis is the possibility of capturing the mechanical properties of soft transversely isotropic incompressible nonlinear elastic materials, such as certain soft biological tissues, using acoustoelasticity theory. The constitutive model is given as a strain-energy density function that depends on the invariants of the right Cauchy–Green tensor. The equations governing infinitesimal motions superimposed on a finite deformation have been used in conjunction with the constitutive law to examine the propagation of homogeneous plane waves. The speeds of homogeneous plane waves have been derived. Furthermore, the differences between this theoretical framework and the parallel one obtained for constitutive models that depend on the Green strain tensor have been highlighted. The use of both acoustoelastic wave speed framework formulations may help to scrutinize the nature of the elastic constants as well as to decide which elastic constants must be retained in the development of models.

#### Appendix: Derivatives of the invariants and the elasticity tensor

The expressions for the stress and elasticity tensors require the calculation of

$$\frac{\partial \Omega}{\partial \mathbf{F}} = \sum_{i=1}^N \Omega_i \frac{\partial I_i^*}{\partial \mathbf{F}} \quad (\text{A.1})$$

and

$$\frac{\partial^2 \Omega}{\partial \mathbf{F} \partial \mathbf{F}} = \sum_{i=1}^N \Omega_i \frac{\partial^2 I_i^*}{\partial \mathbf{F} \partial \mathbf{F}} + \sum_{i=1}^N \sum_{j=1}^N \Omega_{ij} \frac{\partial I_i^*}{\partial \mathbf{F}} \otimes \frac{\partial I_j^*}{\partial \mathbf{F}}, \quad (\text{A.2})$$

where we have used the shorthand notations  $\Omega_i = \partial \Omega / \partial I_i^*$ ,  $\Omega_{ij} = \partial^2 \Omega / \partial I_i^* \partial I_j^*$ ,  $i, j = 1, 2, \dots, N$ .

For the considered incompressible material the nominal stress is

$$\mathbf{S} = \frac{\partial \Omega}{\partial \mathbf{F}} - p \mathbf{F}^{-1} = \sum_{i=1, i \neq 3}^5 \Omega_i \frac{\partial I_i^*}{\partial \mathbf{F}} - p \mathbf{F}^{-1}, \quad (\text{A.3})$$

and the corresponding Cauchy stress is

$$\boldsymbol{\sigma} = \mathbf{F} \frac{\partial \Omega}{\partial \mathbf{F}} - p \mathbf{I} = \sum_{\substack{i=1 \\ i \neq 3}}^5 \Omega_i \mathbf{F} \frac{\partial I_i^*}{\partial \mathbf{F}} - p \mathbf{I}. \quad (\text{A.4})$$

The elasticity tensor is given by

$$\mathcal{A} = \frac{\partial^2 \Omega}{\partial \mathbf{F} \partial \mathbf{F}} = \sum_{\substack{1 \leq i \leq 5 \\ i \neq 3}} \Omega_i \frac{\partial^2 I_i^*}{\partial \mathbf{F} \partial \mathbf{F}} + \sum_{\substack{1 \leq i \leq 5 \\ i \neq 3}} \sum_{\substack{1 \leq j \leq 5 \\ j \neq 3}} \Omega_{ij} \frac{\partial I_i^*}{\partial \mathbf{F}} \otimes \frac{\partial I_j^*}{\partial \mathbf{F}}. \quad (\text{A.5})$$

This requires expressions for the derivatives of the invariants, which are

$$\begin{aligned} \frac{\partial I_1^*}{\partial F_{i\alpha}} &= 2F_{i\alpha}, & \frac{\partial I_2^*}{\partial F_{i\alpha}} &= 2(I_1^* F_{i\alpha} - F_{k\alpha} F_{k\delta} F_{i\delta}), & \frac{\partial I_4^*}{\partial F_{i\alpha}} &= 2M_\alpha F_{i\delta} M_\delta, \\ \frac{\partial I_5^*}{\partial F_{i\alpha}} &= 2(M_\alpha F_{i\delta} C_{\delta\gamma} M_\gamma + C_{\alpha\gamma} M_\gamma F_{i\delta} M_\delta), & \frac{\partial^2 I_1^*}{\partial F_{i\alpha} \partial F_{j\beta}} &= 2\delta_{ij} \delta_{\alpha\beta}, \\ \frac{\partial^2 I_2^*}{\partial F_{i\alpha} \partial F_{j\beta}} &= 2(2F_{i\alpha} F_{j\beta} - F_{i\beta} F_{j\alpha} + C_{\gamma\gamma} \delta_{ij} \delta_{\alpha\beta} - B_{ij} \delta_{\alpha\beta} - C_{\alpha\beta} \delta_{ij}), & \frac{\partial^2 I_4^*}{\partial F_{i\alpha} \partial F_{j\beta}} &= 2\delta_{ij} M_\alpha M_\beta, \\ \frac{\partial^2 I_5^*}{\partial F_{i\alpha} \partial F_{j\beta}} &= 2\delta_{ij} (M_\alpha C_{\beta\gamma} + M_\beta C_{\alpha\gamma}) M_\gamma + 2B_{ij} M_\alpha M_\beta \\ &\quad + 2\delta_{\alpha\beta} F_{i\gamma} M_\gamma F_{j\delta} M_\delta + 2(F_{i\beta} F_{j\gamma} M_\alpha + F_{j\alpha} F_{i\gamma} M_\beta) M_\gamma. \end{aligned} \quad (\text{A.6})$$

The pushed-forward quantity from the initial reference configuration to the finitely deformed equilibrium configuration of  $\mathcal{A}$  is denoted  $\mathcal{A}_0$ . We give its specialization to the situation in which there is no finite deformation and  $\mathcal{B}_0$  coincides with  $\mathcal{B}_r$ . The components of  $\mathcal{A}_0$  in the reference configuration for  $\Omega$  using (A.6), the chain rule, and the conditions (8) can be arranged in the form

$$\begin{aligned} &\mathcal{A}_{0piqj} \\ &= F_{p\alpha} F_{q\beta} \frac{\partial^2 \Omega}{\partial F_{i\alpha} \partial F_{j\beta}} \\ &= 2\Omega_1 \delta_{ij} B_{pq} + 2\Omega_2 (2B_{ip} B_{jq} - B_{iq} B_{jp} + I_1^* \delta_{ij} B_{pq} - B_{ij} B_{pq} - \delta_{ij} (B^2)_{pq}) + 2\Omega_4 \delta_{ij} m_p m_q \\ &\quad + 2\Omega_5 [\delta_{ij} (m_p B_{qr} m_r + m_q B_{pr} m_r) + B_{ij} m_p m_q + m_i m_j B_{pq} + B_{iq} m_j m_p + B_{pj} m_i m_q] \\ &\quad + 4\Omega_{11} B_{ip} B_{jq} + 4\Omega_{22} (I_1^* B_{ip} - (B^2)_{ip}) (I_1^* B_{jq} - (B^2)_{jq}) + 4\Omega_{44} m_i m_j m_p m_q \\ &\quad + 4\Omega_{55} [m_p B_{ir} m_r + m_i B_{pr} m_r] [m_q B_{jr} m_r + m_j B_{qr} m_r] \\ &\quad + 4\Omega_{12} [B_{ip} (I_1^* B_{jq} - (B^2)_{jq}) + B_{jq} (I_1^* B_{ip} - (B^2)_{ip})] + 4\Omega_{14} [B_{ip} m_j m_q + B_{jq} m_i m_p] \\ &\quad + 4\Omega_{15} [B_{ip} [m_q B_{jr} m_r + m_j B_{qr} m_r] + B_{jq} [m_i B_{pr} m_r + m_p B_{ir} m_r]] \\ &\quad + 4\Omega_{24} [(I_1^* B_{ip} - (B^2)_{ip}) m_j m_q + (I_1^* B_{jq} - (B^2)_{jq}) m_i m_p] \\ &\quad + 4\Omega_{25} [(I_1^* B_{ip} - (B^2)_{ip}) [m_q B_{jr} m_r + m_j B_{qr} m_r] + (I_1^* B_{jq} - (B^2)_{jq}) [m_i B_{pr} m_r + m_p B_{ir} m_r]] \\ &\quad + 4\Omega_{45} [m_i m_p [m_q B_{jr} m_r + m_j B_{qr} m_r] + m_j m_q [m_i B_{pr} m_r + m_p B_{ir} m_r]]. \end{aligned} \quad (\text{A.7})$$

### Acknowledgements

This work was supported by the Vietnam National Foundation for Science and Technology Development (NAFOSTED). Merodio acknowledges support from the Ministerio de Ciencia of Spain under the project reference DPI2011-26167.

## References

- [Bigoni et al. 2007] D. Bigoni, D. Capuani, P. Bonetti, and S. Colli, “A novel boundary element approach to time-harmonic dynamics of incremental nonlinear elasticity: the role of pre-stress on structural vibrations and dynamic shear banding”, *Comput. Methods Appl. Mech. Eng.* **196**:41-44 (2007), 4222–4249.
- [Bigoni et al. 2008] D. Bigoni, M. Gei, and A. B. Movchan, “Dynamics of a prestressed stiff layer on an elastic half space: filtering and band gap characteristics of periodic structural models derived from long-wave asymptotics”, *J. Mech. Phys. Solids* **56**:7 (2008), 2494–2520.
- [Biot 1965] M. A. Biot, *Mechanics of incremental deformations: theory of elasticity and viscoelasticity of initially stressed solids and fluids, including thermodynamic foundations and applications to finite strain*, Wiley, New York, 1965.
- [Destrade and Ogden 2010] M. Destrade and R. W. Ogden, “On the third- and fourth-order constants of incompressible isotropic elasticity”, *J. Acoust. Soc. Am.* **128** (2010), 3334–3343.
- [Destrade et al. 2010a] M. Destrade, M. D. Gilchrist, and R. W. Ogden, “Third- and fourth-order elasticity of biological soft tissues”, *J. Acoust. Soc. Am.* **127** (2010), 2103–2106.
- [Destrade et al. 2010b] M. Destrade, M. D. Gilchrist, and G. Saccomandi, “Third- and fourth-order constants of incompressible soft solids and the acousto-elastic effect”, *J. Acoust. Soc. Am.* **127** (2010), 2759–2763.
- [Hoger 1999] A. Hoger, “A second order constitutive theory for hyperelastic materials”, *Int. J. Solids Struct.* **36**:6 (1999), 847–868.
- [Holzapfel and Ogden 2009] G. A. Holzapfel and R. W. Ogden, “On planar biaxial tests for anisotropic nonlinearly elastic solids: a continuum mechanical framework”, *Math. Mech. Solids* **14** (2009), 474–489.
- [Holzapfel et al. 2000] G. A. Holzapfel, T. C. Gasser, and R. W. Ogden, “A new constitutive framework for arterial wall mechanics and a comparative study of material models”, *J. Elasticity* **61**:1-3 (2000), 1–48.
- [Merodio and Saccomandi 2006] J. Merodio and G. Saccomandi, “Remarks on cavity formation in fiber-reinforced incompressible non-linearly elastic solids”, *Eur. J. Mech. A Solids* **25**:5 (2006), 778–792.
- [Ogden and Singh 2011] R. W. Ogden and B. Singh, “Propagation of waves in an incompressible transversely isotropic elastic solid with initial stress: Biot revisited”, *J. Mech. Mater. Struct.* **6** (2011), 453–477.
- [Vinh and Merodio 2013] P. C. Vinh and J. Merodio, “Wave velocity formulas to evaluate elastic constants of soft biological tissues”, *J. Mech. Mater. Struct.* **8** (2013), 51–64.

Received 1 Mar 2013. Revised 2 Apr 2013. Accepted 10 Apr 2013.

PHAM CHI VINH: pcvinh@vnu.edu.vn

*Faculty of Mathematics, Mechanics and Informatics, Hanoi University of Science, 334, Nguyen Trai Str., Thanh Xuan, Hanoi 1000, Vietnam*

JOSE MERODIO: merodioj@gmail.com

*Department of Continuum Mechanics and Structures, E.T.S. Ingeniería de Caminos, Canales e Puertos, Universidad Politécnica de Madrid, 28040 Madrid, Spain*



## IDENTIFICATION OF MULTILAYERED THIN-FILM STRESS FROM NONLINEAR DEFORMATION OF SUBSTRATE

KANG FU

Due to the enlargement of wafer size and the increase of product integrity, thin-film stress problems inevitably get into the range of geometric nonlinearity and are found in multilayered thin-film materials. In this work, multilayered thin-film materials are modeled as a large-deflection multilayered composite plate in the framework of geometrically nonlinear plate theory. Based on the principle of virtual work for a thin-film material plate with thin-film stresses of multiple layers as driving forces, a nonlinear plate finite element system for kinematic fields of thin-film materials, which includes in-plane displacements, cross section rotations, and out-of-plane deflection, is established. The least squares method with regularization applied for total or partial kinematic fields obtained by the finite element method solution versus those given by experiments leads to an iterative procedure for identification of the nonlinear multilayered thin-film stresses.

### 1. Introduction

Thin-film materials are widely used in the microelectronics, optoelectronics, and microelectromechanical systems industries. An understanding of the mechanical properties of thin-film materials plays an important role in quality control in component fabrication and in reliability assurance. Among all thin-film material mechanical properties, thin-film stress is a key factor which has to be determined. Thin-film stress in fabrication processes may introduce undesirable warping of wafers, peeling of thin films from substrates, and cracking of thin films. Thin-film stress may influence the functional properties of products, for example, the electronic properties of microelectronic circuits and the performance of microelectromechanical systems. Along with the enlargement of wafer diameter and the increase of product integrity, thin-film stress problems inevitably get into the range of geometric nonlinearity and are found in the multilayered thin-film materials.

Thin-film stress is introduced in thin-film materials during the thin-film material development process. The mismatch of the microstructures and the difference in thermoexpansion coefficients of the materials constituting substrates and thin films are the origins of thin-film stress. Thin-film stress is divided by origin into intrinsic and thermal thin-film stresses, which generally coexist. Although intrinsic and thermal thin-film stress can be determined, respectively, from microscopic and macroscopic models, the methods that determine thin-film stress directly from integrated thin-film materials remain the first choice owing to their simplicity in application and the sufficient precision they offer under certain conditions. In this category is the method which uses the measured curvature variation of thin-film material to evaluate thin-film stress, proposed first by Stoney [1909] and developed further in [Brenner and Senderoff 1949;

---

*MSC2000:* 74K35, 74G75, 74S05.

*Keywords:* thin-film stress, thin-film material, inverse problem, numerical method, nonlinearity.

Glang et al. 1965]. This method is valid in the case of small deformation and constant thin-film material thickness, as well as for isotropic and homogeneous thin-film stress distributions. The classic Stoney formula is used in many industrial applications [Flinn 1989; Nix 1989]. Extensions to the cases of multilayered thin-film materials, nonconstant thin-film material thicknesses, nonhomogeneous thin-film stress distributions, and geometric nonlinearity have been made in the last three decades [Harper and Wu 1990; Finot and Suresh 1996; Finot et al. 1997; Freund et al. 1999; Freund 2000; Giannakopoulos et al. 2001; Feng et al. 2006; Huang and Rosakis 2006]. In dealing with thin-film stress problems with geometric nonlinearity, Masters and Salamon [1993] and Salamon and Masters [1995] resorted to a Ritz-type method based on the potential energy minimization principle, where the relations of the thin-film stress to both the deflection and the curvature of thin-film materials are studied. The aforementioned methods are all analytic methods. In these cases, various types of restrictions must be retained in order to guarantee the possibility of obtaining analytic solutions. Less restricted methods are developed in the framework of finite element methods. The calculation of thin-film stress from measured thin-film material deformation based on an inverse solution of thin-film material problems with a finite element model is found in [Engelstad et al. 2005]. An inverse problem technique based on a linear plate finite element method that is used to identify thin-film stress is presented in [Fu 2012]. Many advantages can be found in the two numerical-type methods due to the inherent adaptability of the finite element method.

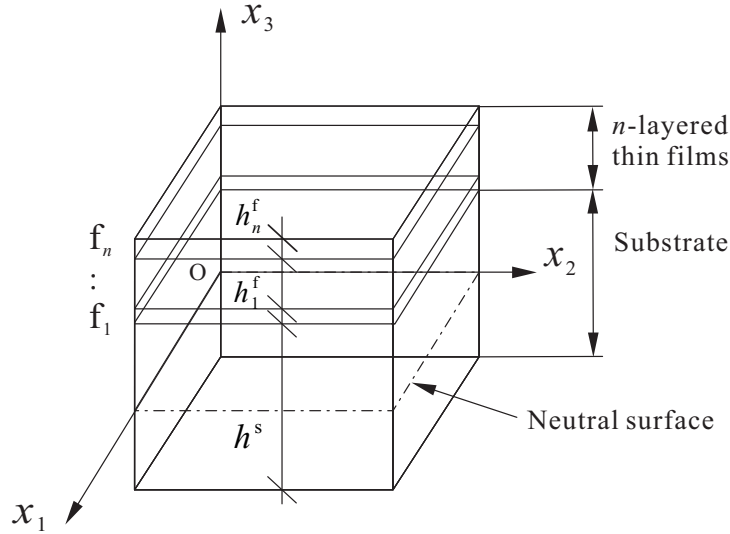
In this paper, the author's previous work is extended to determine the geometrically nonlinear thin-film stress of multilayered thin-film materials. The mechanical modeling of multilayered thin-film material as a special nonlinear composite plate is given in Section 2, where the principle of virtual work for multilayered thin-film stress in the range of geometric nonlinearity is applied at the end of the section. A nonlinear finite element system is established from the virtual work equation by finite element discretization in Section 3. In Section 4, a nonlinear inverse problem for determining thin-film stress using measured kinematic data of thin-film materials is set up via a least squares formula together with the solution scheme. Examples are provided in Section 5 for illustrating the effectiveness of the proposed approach in the determination of nonlinear multilayered thin-film stress. Section 6 contains conclusions.

In the text, Greek subscripts vary in the range  $\{1, 2\}$  and Latin subscripts take values in  $\{1, 2, 3\}$ . The summation convention applies to each repeated pair of indices in the formulas unless otherwise indicated.

## 2. Mechanical modeling of nonlinear multilayered thin-film materials

**2.1. Basic formulations.** Nonlinear multilayered thin-film material is here modeled as a special composite plate, which consists of a substrate and  $n$  thin-film layers of different materials. An infinitesimal portion of the thin-film material is illustrated in Figure 1, where  $h^s$  and  $h_i^f$  are the thicknesses of the substrate and the  $i$ -th thin-film layer for  $i = 1, \dots, n$ , respectively. For thin films,  $h_i^f \ll h^s$ . For the thin-film materials considered in this work, the thicknesses of the thin films do not have to be constant and no defaults such as cracks exist in the materials.

Suppose that the total macroscopic deformation of a thin-film material stems only from the nonconformal internal eigenstrains of the thin films and the substrate without the intervention of external loads, and it is the continuity requirement on the total deformation of the thin-film material that will introduce the elastic strains and corresponding stresses in the thin films as well as in the substrate.



**Figure 1.** Multilayered thin-film materials.

In modeling the structure of a thin-film material as a nonlinear multilayered thin or moderately thick plate [Reddy 2003; Roccabianca et al. 2010; 2011] under the hypothesis of large deflection [von Kármán 1910; Timoshenko and Woinowsky-Krieger 1959], the primary kinematic fields of the first-order shear deformable plate model used to represent the total deformation of the thin-film material are the in-plane displacements  $u_\alpha$ , the out-of-plane displacement  $u_3$ , and the plate cross section rotations  $\theta_\alpha$ , which are all defined on the neutral surface of the plate and are in reference to a configuration that is free of any strain.

In a Cartesian coordinate system, the origin of which is on the neutral surface of the thin-film material plate as shown in Figure 1, the geometrically nonlinear strain-displacement relations in the two longitudinal directions of the thin-film material plate are given by

$$\epsilon_{11} = \partial_1 u_1 + x_3 \partial_1 \theta_2 + (\partial_1 u_3 \partial_1 u_3)/2, \quad (1)$$

$$\epsilon_{22} = \partial_2 u_2 - x_3 \partial_2 \theta_1 + (\partial_2 u_3 \partial_2 u_3)/2, \quad (2)$$

$$2\epsilon_{12} = \partial_2 u_1 + \partial_1 u_2 - x_3 \partial_1 \theta_1 + x_3 \partial_2 \theta_2 + (\partial_1 u_3 \partial_2 u_3 + \partial_2 u_3 \partial_1 u_3)/2, \quad (3)$$

where  $\epsilon_{\alpha\beta}$  are the strain components in the two longitudinal directions of the thin-film material plate and  $x_3$  is the coordinate in the normal direction of the plate neutral surface, and, as in the following,  $\partial_\alpha(*) = \partial(*)/\partial x_\alpha$ . The components of the transversal shear strain are given by

$$\epsilon_{13} = (\partial_1 u_3 - \theta_2)/2, \quad (4)$$

$$\epsilon_{23} = (\partial_2 u_3 - \theta_1)/2. \quad (5)$$

The strain component formulations given by (1)–(5) are applied for the strain components of the substrate,  $\epsilon_{\alpha\beta}^s$ , or for those of the  $i$ -th thin-film layer,  $\epsilon_{\alpha\beta}^f$ ,  $i = 1, \dots, n$ , depending on whether the coordinate  $x_3$  is placed in the substrate or in the  $i$ -th thin-film layer. The application of the identical geometric relations of (1)–(5) to the substrate and the thin films implicates automatically the continuity of the total deformation

of the thin-film materials at the interfaces between all the adjacent thin-film layers as well as at that between the first thin-film layer and the substrate. Since the kinematic fields have a strain-free reference configuration, the total strains given by (1)–(5) consist of elastic strains and eigenstrains.

The deformations of the substrate and thin films given by (1)–(3) are total strains, which consist of eigenstrains,  $\epsilon_{\alpha\beta}^*$ , and elastic strains,  $\epsilon_{\alpha\beta}^e$ , for the substrate as well as for the thin films:

$$\epsilon_{\alpha\beta} = \epsilon_{\alpha\beta}^e + \epsilon_{\alpha\beta}^*. \quad (6)$$

The eigenstrains of the substrate,  $\epsilon_{\alpha\beta}^{s*}$ , are usually isotropic thermal strains, while those of the thin films,  $\epsilon_{\alpha\beta}^{f*}$ , consist of thermal strains and intrinsic strains. In the substrate and the thin films, the elastic strains are developed to compensate the nonconformal eigenstrains in order to ensure the continuity of the total deformation of the thin-film materials.

The stresses in the substrate and thin films are related to their mechanical conjugated elastic strains. For the sake of simplicity of presentation, the constitutive relations for the isotropic and linear elastic materials that constitute the substrate and thin films are discussed in this paper. For substrate materials, the stress components,  $\sigma_{\alpha\beta}^s$ , and the elastic strain components,  $\epsilon_{\alpha\beta}^{es}$ , are supposed to be related by a constitutive law of the pseudo-three-dimensional stress state given by

$$\sigma_{\alpha\alpha}^s = \frac{E^s}{1 - \nu^s 2} (\epsilon_{\alpha\alpha}^{es} + \nu^s \epsilon_{\beta\beta}^{es}), \quad (7)$$

$$\sigma_{\alpha\beta}^s = \frac{E^s}{1 + \nu^s} \epsilon_{\alpha\beta}^{es} \quad \text{for } \alpha \neq \beta, \quad (8)$$

$$\sigma_{\alpha 3}^s = \frac{\kappa E^s}{1 + \nu^s} \epsilon_{\alpha 3}^{es}, \quad (9)$$

where  $E^s$  and  $\nu^s$  are the Young's modulus and Poisson's ratio of the substrate materials, respectively, and  $\kappa$  is the shear force numerical corrector of the plate model. The transversal normal stress component  $\sigma_{33}$  is considered to vanish both in the substrate and the thin films. The transversal shear deformation and the transversal shear stress are considered to be negligible in the thin films, supposing that they are in the region that is sufficiently approximated to the thin-film material plate top or bottom surface. Therefore, the thin-film stress components,  $\sigma_{\alpha\beta}^{fi}$ , and the thin-film elastic strain components,  $\epsilon_{\alpha\beta}^{efi}$ , in the  $i$ -th thin-film layer are considered to be related by a constitutive law of plane stress state given by

$$\sigma_{\alpha\alpha}^{fi} = \frac{E^{fi}}{1 - \nu^{fi} 2} (\epsilon_{\alpha\alpha}^{efi} + \nu^{fi} \epsilon_{\beta\beta}^{efi}), \quad (10)$$

$$\sigma_{\alpha\beta}^{fi} = \frac{E^{fi}}{1 + \nu^{fi}} \epsilon_{\alpha\beta}^{efi} \quad \text{for } \alpha \neq \beta, \quad (11)$$

where  $E^{fi}$  and  $\nu^{fi}$  are the Young's modulus and Poisson's ratio of the material of the  $i$ -th thin-film layer, respectively.

Similarly to the previous work [Fu 2012], the principle of virtual work for geometrically nonlinear multilayered thin-film materials in its established equilibrium state is formulated as follows:

$$\int_{\Omega^s} \sigma_{kl}^s(u_\alpha, u_3, \theta_\alpha) \delta \epsilon_{kl}^s d\Omega^s + \sum_{i=1}^n \int_{\Omega^{fi}} \sigma_{\alpha\beta}^{fi} \delta \epsilon_{\alpha\beta}^{fi} d\Omega^{fi} = 0, \quad (12)$$



where  $\Omega^s$  and  $\Omega^{f_i}$  are the volumes of the substrate and the  $i$ -th thin-film layer, respectively. The identification of thin-film stress means that the thin-film stress components of all the thin-film layers  $\sigma_{\alpha\beta}^{f_i}$ ,  $i = 1, \dots, n$ , in the second term of the left-hand side of (12) are considered as independent variables, which have to be determined from some of the kinematic fields that are obtained from physical measurements.

**2.2. Identification-oriented formulations.** The use of the relation between the total kinematic fields of the thin-film materials and the thin-film stresses of all the thin-film layers of (12) is not sufficient to identify the thin-film stresses from the total kinematic fields, since total deformation of a thin-film material may correspond to different combinations of multilayered thin-film stresses. Two approaches that can be used to solve this problem are suggested in this subsection.

One approach is based on the use of a series of measured substrate deformations of thin-film material  $\Delta\bar{u}^{f_i}$ ,  $i = 1, \dots, n$ , of which the  $i$ -th measured deformation is related to the  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$ , for  $i = 1, \dots, n$ , respectively. In practice, there are two occasions in which the substrate deformation of thin-film material  $\Delta\bar{u}^{f_i}$  with respect to the  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$ , for  $i = 1, \dots, n$ , can be measured experimentally: One is that when the  $i$ -th thin-film layer is deposited on the  $(i - 1)$ -th thin-film layer during a normal layer-by-layer thin-film forming process, which is realized in order from the first thin-film layer to the  $n$ -th. This measurement may be carried out at an interval of every two normal deposition operations. Another occasion is that when the  $i$ -th thin-film layer is removed from the  $(i - 1)$ -th thin-film layer during a layer-by-layer thin-film sacrificing process, which is realized in order from the  $n$ -th thin-film layer to the first. This second process is a type of investigative test with the use of destructive techniques such as chemical etching or mechanical grinding upon fabricated multilayered thin-film material samples. Usually, these operations on the  $i$ -th thin-film layer should not change the intrinsic thin-film strains of the lower-ordered layers that are bound in in-processing thin-film materials. As for possible changes of the temperature strains of lower-ordered layer thin films or the substrate, they are easily estimated.

Notice that in the two measurement processes, the measured deformation of the thin-film material  $\Delta\bar{u}^{f_i}$  is a deformation induced by adding or removing the  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$  on an in-processing thin-film material. At this moment, the substrate deformation of the thin-film material  $\Delta\bar{u}^{f_i}$  depends not only on the  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$ , but also on all the thin-film stresses  $\sigma_{\alpha\beta}^{f_p}$ ,  $p = 1, \dots, i - 1$ , since the first through  $(i - 1)$ -th layer thin films are still bound on the substrate.

In accordance with these properties of the layer-by-layer deformation measurement processes on an in-processing thin-film material, the principle of virtual work for the in-processing thin-film material under the action both of the substrate stress  $\sigma_{kl}^s$  and thin-film stresses  $\sigma_{\alpha\beta}^{f_p}$ ,  $p = 1, \dots, i$ , is formulated as

$$\int_{\Omega^s} \sigma_{kl}^s \left( \sum_{p=1}^i \Delta\bar{u}^{f_p} \right) \delta\epsilon_{kl}^s d\Omega^s + \sum_{p=1}^i \int_{\Omega^{f_p}} \sigma_{\alpha\beta}^{f_p} \delta\epsilon_{\alpha\beta}^{f_p} d\Omega^{f_p} = 0, \quad (13)$$

where the summation of  $\Delta\bar{u}^{f_p}$  from 1 to  $i$  constitutes the total change of the kinematic fields induced by the thin-film stresses of all the thin-film layers  $\sigma_{\alpha\beta}^{f_p}$ ,  $p = 1, \dots, i$ . When the deformations of the thin-film material with respect to all the thin-film stresses  $\Delta\bar{u}^{f_i}$ ,  $i = 1, \dots, n$ , are in disposition after one of the aforementioned measurement processes has been completely carried out, the  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$ ,  $i = 1, \dots, n$ , can be evaluated according to (13) in order from the first thin-film layer to the  $n$ -th in a layer-by-layer cumulative way.

In the identification process of the  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$ , the  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$  is evaluated according to (13) from the measured deformations  $\Delta\bar{u}^{f_p}$ ,  $p = 1, \dots, i$ , and the thin-film stresses  $\sigma_{\alpha\beta}^{f_p}$ ,  $p = 1, \dots, i - 1$ . Meanwhile the thin-film stresses  $\sigma_{\alpha\beta}^{f_p}$ ,  $p = 1, \dots, i - 1$ , also have to be updated so as to take into account the influence on them of the  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$ . For this purpose, the measured deformations  $\Delta\bar{u}^{f_p}$ ,  $p = 1, \dots, i$ , and the thin-film stresses  $\sigma_{\alpha\beta}^{f_p}$ ,  $p = 1, \dots, i - 1$ , obtained in the previous layer evaluation are used at the beginning of the process to evaluate a first  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$ . Then, the  $p$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_p}$ , for  $p = 1, \dots, i - 1$ , is updated according to (6) and (10) and (11) from the  $p$ -th layer total strain  $\epsilon_{\alpha\beta}^{f_p}$  given by the kinematic fields obtained in the actual layer evaluation, and the  $p$ -th layer invariant eigenstrain  $\epsilon_{\alpha\beta}^{f_p*}$  obtained in the previous layer evaluations. The process proceeds iteratively until the  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$  is finally identified. At this moment, the  $i$ -th layer invariant eigenstrain  $\epsilon_{\alpha\beta}^{f_i*}$  is once for all evaluated from the  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$  and the actual  $i$ -th layer total strain  $\epsilon_{\alpha\beta}^{f_i}$  according to (6) and (10) and (11). The final thin-film stresses of all the thin-film layers are identified when this layer-by-layer cumulative process is finished at  $i = n$ .

Another approach is to identify the thin-film stresses of all the thin-film layers from the total deformation of thin-film materials with the use of (12), which is possible when the continuity condition of deformations between all the adjacent thin-film layers can be formulated with the use of thin-film eigenstrains. When the thin-film layers are bound together, the in-plane strain components must be equal on the interface of the adjacent  $i$ -th and  $(i + 1)$ -th thin-film layers owing to the deformation continuity requirement:

$$\epsilon_{\alpha\beta}^{f_i} = \epsilon_{\alpha\beta}^{f_{i+1}}. \quad (14)$$

By substituting (6) into (14), the continuity condition of (14) becomes

$$\epsilon_{\alpha\beta}^{ef_{i+1}} = \epsilon_{\alpha\beta}^{ef_i} - \Delta\epsilon_{\alpha\beta}^{*(i+1,i)}, \quad (15)$$

where  $\Delta\epsilon_{\alpha\beta}^{*(i+1,i)}$  is the difference of the eigenstrain components between the  $i$ -th and  $(i + 1)$ -th thin-film layers at their interface. By substituting (10) and (11) into (15), the transitional conditions for the thin-film stresses on the interface between the  $i$ -th and  $(i + 1)$ -th thin-film layers are found as

$$\sigma_{\alpha\alpha}^{f_{i+1}} = \frac{E^{f_{i+1}}}{[1 - (\nu^{f_{i+1}})^2]} \left\{ \frac{1}{E^{f_i}} [(1 - \nu^{f_i} \nu^{f_{i+1}}) \sigma_{\alpha\alpha}^{f_i} - (\nu^{f_i} - \nu^{f_{i+1}}) \sigma_{\beta\beta}^{f_i}] - (\Delta\epsilon_{\alpha\alpha}^{*(i+1,i)} + \nu^{f_{i+1}} \Delta\epsilon_{\beta\beta}^{*(i+1,i)}) \right\}, \quad (16)$$

$$\sigma_{\alpha\beta}^{f_{i+1}} = \frac{E^{f_{i+1}}}{(1 + \nu^{f_{i+1}})} \left[ \frac{(1 + \nu^{f_i})}{E^{f_i}} \sigma_{\alpha\beta}^{f_i} - \Delta\epsilon_{\alpha\beta}^{*(i+1,i)} \right] \quad \text{for } \alpha \neq \beta. \quad (17)$$

With the use of these transitional conditions on the thin-film stresses, the thin-film stresses of all the thin-film layers can be represented by a generic thin-film stress of a chosen principal thin-film layer. After doing this, (12) can be used to determine the generic thin-film stress from the integral deformation of the  $n$ -layered thin-film materials. The key in this approach is that all the differences of the eigenstrains between the  $i$ -th and  $(i + 1)$ -th thin-film layers  $\Delta\epsilon_{\alpha\beta}^{*(i+1,i)}$ ,  $i = 1, \dots, n - 1$ , must be known, for the acquisition of which one may turn to the solution of other important problems such as the thermal strains, experimental measurements, multiscale modeling and computations, etc., which are left to be the research subjects of other works.

In the present work, the second approach is used partially in the aforementioned layer-by-layer cumulative approach. Notice that once the thin-film stress of a layer is determined in the layer-by-layer

cumulative process, it can be used with the previously evaluated and newly updated thin-film stress of the lower-ordered layer to calculate the difference of eigenstrain components between the two layers according to (16) and (17). The eigenstrains as well as their differences between adjacent layers are generally unchanged or may be easily determined during the layer-by-layer cumulative evaluation process, and the layer-by-layer cumulative approach can use these properties to establish a more efficient solution strategy. The advantages of the application of these properties are illustrated in the following.

For an in-processing thin-film material with one thin-film layer at the beginning, both the substrate stress  $\sigma_{ij}^s$  and the first-layer thin-film stress  $\sigma_{\alpha\beta}^{f_1}$  are first evaluated as before from  $\Delta\bar{u}^{f_1}$ . Then, the difference of the eigenstrain components at the interface between the substrate and the first thin-film layer  $\Delta\epsilon_{\alpha\beta}^{*(1,0)}$  is determined from (16) and (17), in which the substrate stress  $\sigma_{\alpha\beta}^s$  at the one surface of the substrate that is in contact with the first thin-film layer is used as the lower-ordered layer stress.

Following the preceding step, in identifying the second-layer thin-film stress  $\sigma_{\alpha\beta}^{f_2}$  and in updating the substrate stress  $\sigma_{ij}^s$  and the first-layer thin-film stress  $\sigma_{\alpha\beta}^{f_1}$  from  $\Delta\bar{u}^{f_1}$  and  $\Delta\bar{u}^{f_2}$ , the actual first-layer thin-film stress  $\sigma_{\alpha\beta}^{f_1}$  in (13) is represented by the actual substrate stress  $\sigma_{ij}^s$  according to (16) and (17) with the use of the unchanged difference of the eigenstrain components at the interface between the substrate and the first thin-film layer  $\Delta\epsilon_{\alpha\beta}^{*(1,0)}$ . After the updated substrate stress  $\sigma_{ij}^s$  and the identified second-layer thin-film stress  $\sigma_{\alpha\beta}^{f_2}$  are determined from  $\Delta\bar{u}^{f_1}$  and  $\Delta\bar{u}^{f_2}$  according to (13), the updated first-layer thin-film stress  $\sigma_{\alpha\beta}^{f_1}$  is given according to (16) and (17) with the use of the updated substrate stress  $\sigma_{ij}^s$  and the unchanged difference of the eigenstrain components at the interface between the substrate and the first thin-film layer  $\Delta\epsilon_{\alpha\beta}^{*(1,0)}$ . The difference of the eigenstrain components at the interface between the first and second thin-film layers  $\Delta\epsilon_{\alpha\beta}^{*(2,1)}$  is then calculated according to (16) and (17) with the use of the updated first-layer thin-film stress  $\sigma_{\alpha\beta}^{f_1}$  and the just-identified second-layer thin-film stress  $\sigma_{\alpha\beta}^{f_2}$ .

Similarly, in identifying the  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$  and in updating the substrate stress  $\sigma_{ij}^s$  and the thin-film stresses  $\sigma_{\alpha\beta}^{f_p}$ ,  $p = 1, \dots, i - 1$ , from  $\Delta\bar{u}^{f_q}$ ,  $q = 1, \dots, i$ , all the actual thin-film stresses  $\sigma_{\alpha\beta}^{f_p}$ ,  $p = 1, \dots, i - 1$ , in (13) are represented successively by the actual substrate stress  $\sigma_{ij}^s$  according to (16) and (17), in which the differences of the eigenstrain components at the interfaces between the  $(p - 1)$ -th and  $p$ -th thin-film layers  $\Delta\epsilon_{\alpha\beta}^{*(p,p-1)}$ ,  $p = 1, \dots, i - 1$ , are already known. After the updated substrate stress  $\sigma_{ij}^s$  and the identified  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$  are determined according to (13) from  $\Delta\bar{u}^{f_q}$ ,  $q = 1, \dots, i$ , the thin-film stresses  $\sigma_{\alpha\beta}^{f_p}$ ,  $p = 1, \dots, i - 1$ , are updated with successive uses of (16) and (17) from the updated substrate stress  $\sigma_{ij}^s$  and the unchanged differences of the eigenstrain components at the interfaces between the  $(p - 1)$ -th and  $p$ -th thin-film layers  $\Delta\epsilon_{\alpha\beta}^{*(p,p-1)}$ ,  $p = 1, \dots, i - 1$ . The difference of the eigenstrain components at the interface between the  $(i - 1)$ -th and  $i$ -th thin-film layers  $\Delta\epsilon_{\alpha\beta}^{*(i,i-1)}$  is then calculated according to (16) and (17) with the use of the updated  $(i - 1)$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_{i-1}}$  and the just-identified  $i$ -th layer thin-film stress  $\sigma_{\alpha\beta}^{f_i}$ . The thin-film stresses of all the thin-film layers are ultimately identified when this process is finished at  $i = n$ .

Since the identification process for multilayered thin-film stress is similar to a successive identification process of single layered thin-film stress and the updating of the thin-film stresses of the lower-ordered layers is not in the core of the identification process, the efficiency of the method is largely improved.

### 3. Finite element discretization

Since the kinematic fields are defined on the plate neutral surface, only two-dimensional quadrilateral or triangular element meshes are needed to mesh the neutral surface of the thin-film material plates.

Thin-film materials may be considered as very thin or moderately thick plates depending on the ratio of thickness to characteristic longitudinal dimensions of the thin-film materials studied. In these cases, the suitable elements for the mechanical model of the multilayered thin-film materials presented in Section 2 are those of transversal shear locking-free plate elements without the appearance of “hourglass” phenomena. Many successful elements for this type of problem can be implemented in the present work with minor modifications. The MITC4 elements [Brezzi et al. 1989; Bucalem and Bathe 1993] used in work [Fu 2012] are here extended to study geometrically nonlinear problems with five kinematic fields. The resulting finite element equation from (12) or (13) can be written as

$$Ku = F\sigma, \quad (18)$$

where  $K$  is the global geometrically nonlinear finite element rigidity matrix;  $u$  is the column vector matrix containing all nodal degrees of freedom of the mesh;  $\sigma$  is the column vector matrix of multilayered thin-film stresses defined in elements or on nodes, which contains only the unknown thin-film stresses of the thin-film layer actually considered during the identification process; and  $F$  is the geometrically nonlinear matrix which transforms the stress column matrix vector  $\sigma$  into an equivalent nodal force column vector matrix of the finite element system, and which contains also the previously identified thin-film stresses of the lower-ordered thin-film layers if the cumulative layer-by-layer method that is presented in Section 2.2 is used. In the identification process of the thin-film stresses, the sensitivity of the kinematic fields to the thin-film stresses is needed in the next section. From (18), the system of the sensitivity is formulated as

$$K \frac{\partial u}{\partial \sigma_{\alpha\beta}^{f_i}} = F \frac{\partial \sigma}{\partial \sigma_{\alpha\beta}^{f_i}}. \quad (19)$$

Equations (18) and (19) are geometrically nonlinear systems dependent on  $u$ , which have to be solved by an iterative procedure. In the same iteration, the factorized matrix  $K$  can be used for the solution processes both of the kinematic nodal values  $u$  and their sensitivity to the thin-film stresses.

#### 4. Identification procedure of geometrically nonlinear multilayered thin-film stress

The determination of the thin-film stresses from the kinematic fields is an inverse problem from the point of view of (18). It is in general impossible to obtain the thin-film stresses  $\sigma$  from the nodal degrees of freedom  $u$  with the direct use of (18) since (18) is usually not well-posed mathematically and only a part of the nodal degrees of freedom  $u$  can, in fact, be provided by the experimental measurements. In order to overcome these difficulties, a least squares condition is used to establish a feasible solution scheme as

$$L = \min_{\{\sigma\}} [(u - \bar{u})^T (u - \bar{u})], \quad (20)$$

where  $\bar{u}$  is a subset of the nodal degrees of freedom  $u$ , which are provided by the experimental measurements. In the layer-by-layer cumulative identification process of the thin-film stress of the  $i$ -th thin-film layer  $\sigma_{\alpha\beta}^{f_i}$  following (13),  $\bar{u}$  is the summation of all the measured substrate deformations  $\Delta \bar{u}^{f_p}$ ,  $p = 1, \dots, i$ . The necessary condition of the minimization leads to an iterative type of equation for the determination of the thin-film stresses given as follows:

$$S_k^T S_k (\sigma_{k+1} - \sigma_k) = S_k^T (u_k - \bar{u}), \quad (21)$$

where the subscript indices  $k$  and  $k + 1$  designate the  $k$ -th and  $(k + 1)$ -th steps of the iterations, and  $S_k = (\partial u / \partial \sigma)_k$  is the sensitivity of the kinematic fields to the thin-film stresses at the  $k$ -th iteration step, which is calculated from (19). Equation (21) is a generalized linear system which may be solved by various techniques, for example, the singular-value-decomposition method, the regularization method, the iterative regularization method, etc. Here the regularization method is used. According to the Tikhonov regularization strategy [Tikhonov and Arsenin 1974], the minimization of (20) leads to

$$(S_k^T S_k + \alpha H^T H)(\sigma_{k+1} - \sigma_k) = S_k^T (u_k - \bar{u}), \quad (22)$$

where  $H$  is a regularization matrix, which is derived from regularization functions, and  $\alpha$  is a regularization parameter, which may be determined by the  $L$ -curve method, the minimum discrepancy method, etc. [Golub et al. 1979; Hansen 1992; Engl et al. 1996]. In the case of ensuring the first-order derivative smoothness of the thin-film stresses, the regularization function may be taken as the  $L^2$ -norm of the thin-film stress field:

$$\Phi = \sum_{i=1}^n \int_{\Omega^{f_i}} \left( \frac{\partial \sigma_{\alpha\beta}^{f_i}}{\partial x_\alpha} \right) d\Omega^{f_i}, \quad (23)$$

where no summation convention is used. The discretized form of (23) leads to

$$\Phi = \sigma^T H^T H \sigma, \quad (24)$$

where the corresponding regularization matrix  $H$  is defined. The regularization matrix  $H$  given by (24) is invariant in the iteration process. Other regularization functions can be used depending on the smoothness requirement of the thin-film stress fields.

In the identification process of the thin-film stresses from a set or subset of the degrees of freedom of the measured deformation  $\bar{u}$  provided by the experiments, an iterative solution procedure with the use of (18), (19), (21), (22), and (24) is given in following algorithm.

**Algorithm.** (i) Beginning of the identification process.

- (ii) Initialization of the iterations on thin-film layers:  $i = 1$  and  $\bar{u} = \Delta \bar{u}^{f_i}$ .
- (iii) Initialization of the identification iterations on  $\sigma_{\alpha\beta}^{f_i}$ :  $k = 0$ ,  $u_k = 0$ , and  $\sigma_k = 0$ .
- (iv) Computation of  $H$  according to (24).
- (v) Update of the thin-film stresses of the lower-ordered layers from  $\Delta \bar{u}^{f_i}$ .
- (vi) Computation of  $u_k$  with the use of (18) from  $\sigma_k$ .
- (vii) Computation of  $S_k$  with the use of (19) from  $u_k$ .
- (viii) Computation of  $\sigma_{k+1}$  with the use of (21) or (22) from  $u_k$ ,  $\sigma_k$ ,  $S_k$ , and  $H$ .
- (ix) Examination of convergence:
  - If  $u_k$  converges to  $\bar{u}$ , go to (x); else  $k = k + 1$ , go to (vi).
- (x) Output of  $\sigma_{k+1}$  as the identified thin-film stresses of the  $i$ -th thin-film layer; go to (xi).
- (xi) Examination of the iteration on thin-film layers:
  - If  $i = n$ , go to (xii); else  $i = i + 1$ ,  $\bar{u} = \bar{u} + \Delta \bar{u}^{f_i}$ , go to (iii).
- (xii) End of the identification process.

## 5. Numerical examples

In this section, two examples are presented to illustrate the characteristics and effectiveness of the proposed method. In the identification processes of the two examples, the least squares method with regularization is used, while the convergence criterion is set to be of a relative error of  $1.0 \times 10^{-4}$  for deflection fields.

**5.1. Nonlinear thin-film stress of a two-layered thin-film material strip.** In this example, a strip of rectangular plan shape of two-layered thin-film material is considered. The substrate material is of silicon (Si) and its original shape is flat. A uniform thin-film layer of tungsten (W) is supposed to be deposited on the substrate. The sizes and material parameters of the thin-film material are given in Table 1.

A Cartesian coordinate system is set up in such a way that the origin is taken at the centroid of the strip geometry; the longitudinal axis of symmetry of the strip, the normal of the longitudinal vertical plan of symmetry, and the upward normal of the neutral surface are taken as the  $x_1$ ,  $x_2$ , and  $x_3$ -axes, respectively. Consider that a uniformly distributed thin-film stress exists in the thin film, of which the stress components denoted by  $\sigma_{11}^f$ ,  $\sigma_{22}^f$ , and  $\sigma_{12}^f$  are 1111.0 MPa, 0.0 MPa, and 0.0 MPa, respectively.

According to the large deflection theory of thin plates, an analytic solution on the deflection of the free strip along the longitudinal direction in the above-mentioned thin-film stress state is found based on [Zhang et al. 2004] as

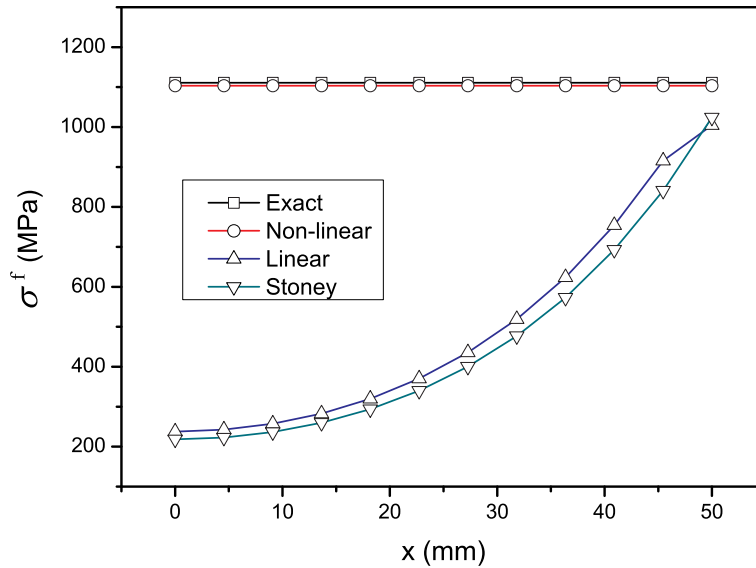
$$u_3(x_1) = \frac{h^s}{2 \cosh(L\sqrt{h^f \sigma_{11}^f/D})} [\cosh(x_1\sqrt{h^f \sigma_{11}^f/D}) - 1], \quad (25)$$

where  $h^s$ ,  $h^f$ ,  $L$ , and  $D$  are the substrate thickness, the thin-film thickness, the semilength of the strip, and the bending stiffness of the plate, respectively.

The deflection of the strip under the aforementioned state of stress calculated from (25) is used as the pseudoexperimental data for the identification of the thin-film stress. The entire strip is discretized with the use of a  $2 \times 22$  mesh of quadrilateral elements of equal size. The values of the deflection  $u_3$  at the element nodal points given by (25) are used as a subset of the nodal degrees of freedom  $\bar{u}$ , which are used in the inverse calculations of both the present nonlinear numerical approach and the linear numerical approach of [Fu 2012]. The curvature of the strip deformation derived from (25) in the same state of stress is also used for the evaluation of the thin-film stress from Stoney's formula. The results of identified thin-film stresses obtained from three different methods in comparison with the original exact stresses in half of the strip are shown in Figure 2. It is noticed that the results of the present nonlinear numerical approach accord well with the original thin-film stresses, while those given by the linear numerical approach and Stoney's formula manifest important discrepancies with respect to the exact values in the same way.

	Young's modulus (GPa)	Poisson's ratio	Thickness ( $\mu\text{m}$ )	Plan sizes ( $\text{mm}^2$ )
Substrate (Si)	130	0.28	350	$100 \times 2$
Film (W)	248	0.30	0.9	$100 \times 2$

**Table 1.** Sizes and material parameters of a thin-film material strip.



**Figure 2.** Comparison of the identified thin-film stresses from different methods.

**5.2. Nonlinear thin-film stress in a three-layered wafer.** A wafer of three-layered thin-film material is considered in this example. The material parameters and geometrical sizes of the thin-film materials are given in Table 2.

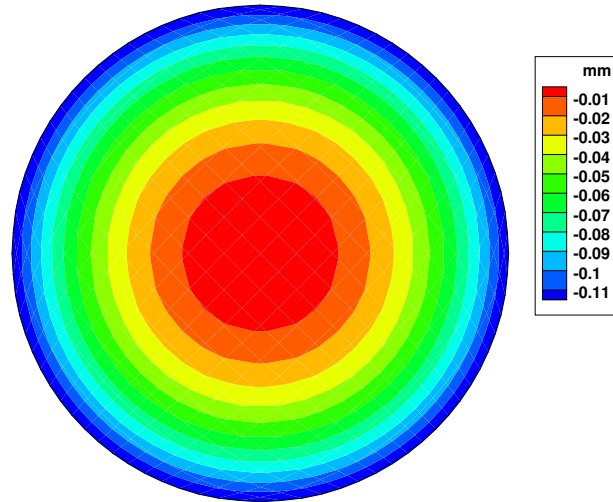
The initial shape of the substrate of silicon is flat, on which the first thin-film layer of nickel and the second thin-film layer of tungsten are supposed to be deposited uniformly. The thin-film material is discretized on its neutral surface with the use of a mesh of 400 four-node quadrilateral elements.

As an example for algorithmic test, the degrees of freedom of experimental measurement  $\bar{u}$  in this example are given by solving a direct problem of the thin-film material with the use of (18), of which the uniformly distributed thin-film stress fields of the first and second thin-film layers are prescribed such that  $\sigma_{11}^{f_1} = \sigma_{22}^{f_1} = 345.0$  MPa,  $\sigma_{11}^{f_2} = \sigma_{22}^{f_2} = 1044.7$  MPa, and  $\sigma_{12}^{f_1} = \sigma_{12}^{f_2} = 0.0$  MPa. Here, the thin-film stresses of the two thin-film layers are intentionally related by the transitional conditions of (16) and (17) in supposing that the eigenstrains of the two thin-film layers are equal, which means also that the elastic strains of the two thin-film layers are equal on the interface between the two layers.

The direct solution of the substrate deflection under only the prescribed thin-film stress of the first thin-film layer  $\Delta \bar{u}_3^{f_1}$  is shown in Figure 3, and is considered as a deflection field of the substrate configuration change due to the removal of the first thin-film layer from the substrate, and has in addition been adjusted by the action of the thin-film stress of the second layer. The direct solution of the substrate deflection

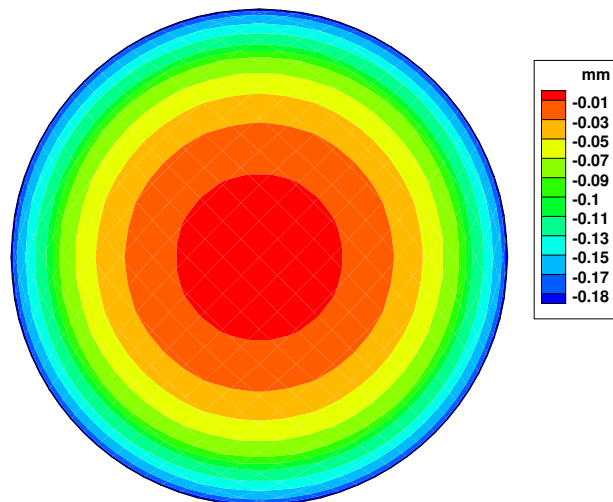
	Young's modulus (GPa)	Poisson's ratio	Thickness ( $\mu\text{m}$ )	Radius (mm)
Substrate (Si)	130	0.28	350	50
Layer 1 (Ni)	130	0.31	0.9	50
Layer 2 (W)	385	0.30	0.9	50

**Table 2.** Sizes and material parameters of a three-layered wafer.



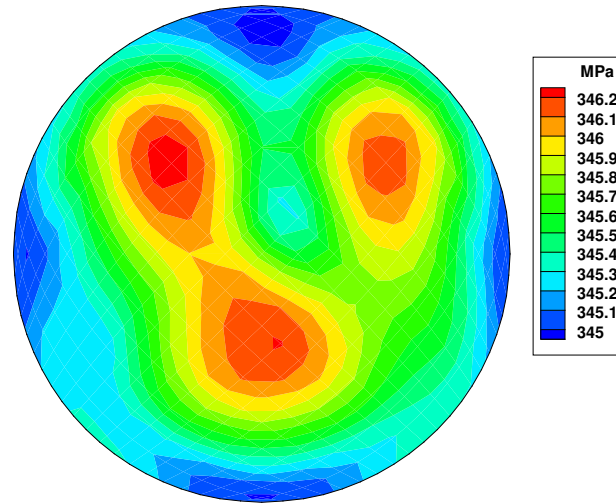
**Figure 3.** Change of substrate deflection due to the removal of the first thin-film layer.

under the prescribed thin-film stresses of the first and second thin-film layers  $\Delta\bar{u}_3 = \Delta\bar{u}_3^{f1} + \Delta\bar{u}_3^{f2}$  is given in Figure 4, which is considered as a deflection field of the total substrate configuration change due to the removal of all the two thin-film layers from a fabricated two-layered thin-film material. Since the maximum values of the deflections in Figures 3 and 4 are of the same order as the thin-film material thickness, the deformations of the thin-film material belong to large deflection problems. In this case, the geometrically nonlinear deformation is more sensitive to the loads at the primary deformation stage. As a consequence, the first deflection field  $\Delta\bar{u}_3^{f1}$  given by the first-layer thin-film stress  $\sigma_{\alpha\beta}^{f1}$  is more important than the adding-up deflection field  $\Delta\bar{u}_3^{f2} = \Delta\bar{u}_3 - \Delta\bar{u}_3^{f1}$  given by the adding-up second-layer thin-film stress  $\sigma_{\alpha\beta}^{f2}$ .

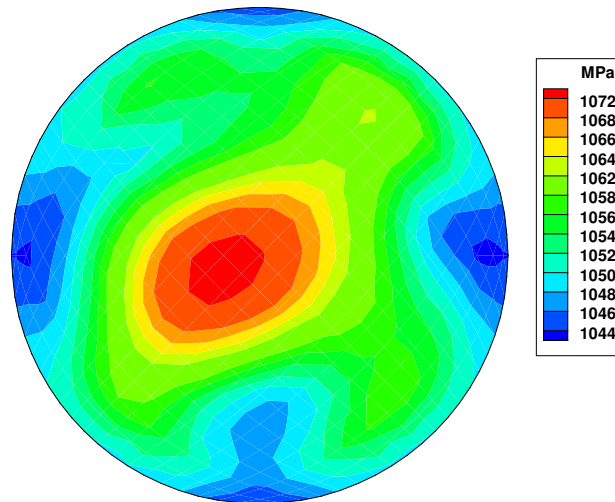


**Figure 4.** Change of substrate deflection due to the removal of the two thin-film layers.





**Figure 5.** Identified maximum principal thin-film stress of the first thin-film layer.



**Figure 6.** Identified maximum principal thin-film stress of the second thin-film layer.

In following the first approach proposed in Section 2.2, the first-layer thin-film stress  $\sigma_{\alpha\beta}^{f_1}$  is first identified with the use of the deflection field  $\Delta\bar{u}_3^{f_1}$ . Then, the second-layer thin-film stress  $\sigma_{\alpha\beta}^{f_2}$  is identified with the use of the total deflection field  $\Delta\bar{u}_3$  and the previously identified first-layer thin-film stress  $\sigma_{\alpha\beta}^{f_1}$ .

The identified maximum principal thin-film stress of the first thin-film layer and that of the second thin-film layer are shown in Figures 5 and 6, respectively. Notice from Figures 5 and 6 that the identified thin-film stresses of the two thin-film layers are in general similar to the prescribed thin-film stresses of the two thin-film layers. The maximum discrepancies on the maximum principal stresses of the first and second thin-film layers with respect to those of the prescribed thin-film stresses of the two thin-film layers

are only locally 0.3% and 2.6%, respectively. The existing discrepancies of the values and distributions compared with those of the original fields may be improved by various numerical means such as mesh refinement, regularization parameter optimization, convergence criteria enhancement, etc.

Finally, notice that the second approach can be also applied in this example. With the use of the transitional conditions for the thin-film stresses on the interface between the first and the second thin-film layers, the thin-film stress of a chosen principal layer can be first identified from the total deformation and then the thin-film stress of other layer can be determined according to the transitional conditions from the identified thin-film stress.

## 6. Conclusions

The identification of thin-film stress in nonlinear multilayered thin-film materials is studied, combining use of mechanical modeling, the finite element method, and the theory of inverse problems. From the method and numerical examples presented, the main findings of the work are summarized as follows.

- (1) When the deformation of the thin-film materials is as large as in the range of the large deflection theory of classical plates, the estimation of thin-film stress should turn to the plate models of nonlinear large deflection theories. The important differences between the geometrically nonlinear and linear theories may manifest in this range of thin-film material deformations. It is shown that the thin-film stress determined from the linear models is underestimated for a given nonlinear correlating kinematic field.
- (2) In the nonlinear identification process, the change of the configuration of thin-film materials does not linearly depend on the thin-film stress. The layer-by-layer cumulative formulation allows for identifying the thin-film stress of each layer from the cumulative deformation and the corresponding cumulative thin-film stresses.
- (3) The transitional conditions for the thin-film stresses at the interfaces between thin-film layers are proposed according to the deformation continuity requirement. In supposing that the eigenstrains as well as their differences between thin-film layers are invariant in the identification process, the stress transitional conditions can be obtained in the layer-by-layer cumulative identification process, and are used as key factors in establishing an efficient solution strategy.
- (4) The applied inverse problem method allows for using a subset of the total degrees of freedom to determine the thin-film stresses, which provides great flexibility in choosing appropriate methods or techniques for measuring correlating fields. The method itself is also adaptive to different computations.
- (5) With use of the finite element method, the nonhomogeneities of geometrical shapes, material characteristics, and mechanical fields in thin-film materials no longer constitute major hurdles in the modeling and solution of the thin-film stress problems since they can be defined element-by-element over the thin-film materials.
- (6) The structural model used in the present work is suitable for both thin and moderately thick thin-film materials. The application of a reliable finite element ensures the validity of the numerical approach in the two cases. The structural model used here can be easily enriched in kinematic fields or in thin-film stress fields whenever needed.

## References

- [Brenner and Senderoff 1949] A. Brenner and S. Senderoff, "Calculation of stress in electrodeposits from the curvature of a plated strip", *J. Res. Nat. Bur. Standards* **42**:2 (1949), 105–123. No. RP1954.
- [Brezzi et al. 1989] F. Brezzi, K.-J. Bathe, and M. Fortin, "Mixed-interpolated elements for Reissner–Mindlin plates", *Int. J. Numer. Methods Eng.* **28**:8 (1989), 1787–1801.
- [Bucalem and Bathe 1993] M. L. Bucalem and K.-J. Bathe, "Higher-order MITC general shell elements", *Int. J. Numer. Methods Eng.* **36**:21 (1993), 3729–3754.
- [Engelstad et al. 2005] R. L. Engelstad, Z. Feng, E. G. Lovell, A. R. Mikkelsen, and J. Sohn, "Evaluation of intrinsic film stress distributions from induced substrate deformation", *Microelectron. Eng.* **78–79** (2005), 404–409.
- [Engl et al. 1996] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, Mathematics and its Applications **375**, Kluwer Academic, Dordrecht, 1996.
- [Feng et al. 2006] X. Feng, Y. Huang, H. Jiang, D. Ngo, and A. J. Rosakis, "The effect of thin film/substrate radii on the Stoney formula for thin film/substrate subjected to nonuniform axisymmetric misfit strain and temperature", *J. Mech. Mater. Struct.* **1**:6 (2006), 1041–1053.
- [Finot and Suresh 1996] M. Finot and S. Suresh, "Small and large deformation of thick and thin-film multi-layers: effects of layer geometry, plasticity and compositional gradients", *J. Mech. Phys. Solids* **44**:5 (1996), 683–721.
- [Finot et al. 1997] M. Finot, I. A. Blech, S. Suresh, and H. Fujimoto, "Large deformation and geometric instability of substrates with thin-film deposits", *J. Appl. Phys.* **81**:8 (1997), 3457–3464.
- [Flinn 1989] P. A. Flinn, "Principles and applications of wafer curvature techniques for stress measurements in thin films", *Mater. Res. Soc. Proc.* **130** (1989), 41–51.
- [Freund 2000] L. B. Freund, "Substrate curvature due to thin film mismatch strain in the nonlinear deformation range", *J. Mech. Phys. Solids* **48**:6-7 (2000), 1159–1174.
- [Freund et al. 1999] L. B. Freund, J. A. Floro, and E. Chason, "Extensions of the Stoney formula for substrate curvature to configurations with thin substrates or large deformations", *Appl. Phys. Lett.* **74**:14 (1999), 1987–1989.
- [Fu 2012] K. Fu, "Identification of thin film stress based on Mindlin–Reissner plate finite element method", *Mech. Adv. Mater. Struct.* **19**:7 (2012), 503–512.
- [Giannakopoulos et al. 2001] A. E. Giannakopoulos, I. A. Blech, and S. Suresh, "Large deformation of thin films and layered flat panels: effects of gravity", *Acta Mater.* **49**:18 (2001), 3671–3688.
- [Glang et al. 1965] R. Glang, R. A. Holmwood, and R. L. Rosenfeld, "Determination of stress in films on single crystalline silicon substrates", *Rev. Sci. Instrum.* **36**:1 (1965), 7–10.
- [Golub et al. 1979] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter", *Technometr.* **21**:2 (1979), 215–223.
- [Hansen 1992] P. C. Hansen, "Analysis of discrete ill-posed problems by means of the  $L$ -curve", *SIAM Rev.* **34**:4 (1992), 561–580.
- [Harper and Wu 1990] B. D. Harper and C.-P. Wu, "A geometrically nonlinear model for predicting the intrinsic film stress by the bending-plate method", *Int. J. Solids Struct.* **26**:5–6 (1990), 511–525.
- [Huang and Rosakis 2006] Y. Huang and A. J. Rosakis, "Extension of Stoney's formula to arbitrary temperature distributions in thin film/substrate systems", *J. Appl. Mech. (ASME)* **74**:6 (2006), 1225–1233.
- [von Kármán 1910] T. von Kármán, "Festigkeitsprobleme im Maschinenbau", pp. 311–385 in *Encyklopädie der mathematischen Wissenschaften*, vol. IV/4, edited by F. Klein and C. Müller, Teubner, Leipzig, 1910.
- [Masters and Salamon 1993] C. B. Masters and N. J. Salamon, "Geometrically nonlinear stress-deflection relations for thin film/substrate systems", *Int. J. Eng. Sci.* **31**:6 (1993), 915–925.
- [Nix 1989] W. D. Nix, "Mechanical properties of thin films", *Metall. Trans. A* **20**:11 (1989), 2217–2245.
- [Reddy 2003] J. N. Reddy, *Mechanics of laminated composite plates and shells: theory and analysis*, 2nd ed., CRC Press, Boca Raton, 2003.

- [Roccabianca et al. 2010] S. Roccabianca, D. Bigoni, and M. Gei, “Plane strain bifurcations of elastic layered structures subject to finite bending: theory versus experiments”, *IMA J. Appl. Math.* **75**:4 (2010), 525–548.
- [Roccabianca et al. 2011] S. Roccabianca, D. Bigoni, and M. Gei, “Long wavelength bifurcations and multiple neutral axes of elastic layered structures subject to finite bending”, *J. Mech. Mater. Struct.* **6**:1–4 (2011), 511–527.
- [Salamon and Masters 1995] N. J. Salamon and C. B. Masters, “Bifurcation in isotropic thin film/substrate plates”, *Int. J. Solids Struct.* **32**:3–4 (1995), 473–481.
- [Stoney 1909] G. G. Stoney, “The tension of metallic films deposited by electrolysis”, *Proc. R. Soc. Lond. A* **82**:553 (1909), 172–175.
- [Tikhonov and Arsenin 1974] A. N. Tikhonov and V. Y. Arsenin, *Методы решения некорректных задач*, Nauka, Moscow, 1974. Translated as *Solutions of ill-posed problems*, Winston, Washington, DC, 1977.
- [Timoshenko and Woinowsky-Krieger 1959] S. Timoshenko and S. Woinowsky-Krieger, *Theory of plates and shells*, 2nd ed., McGraw-Hill, New York, 1959.
- [Zhang et al. 2004] Y. Zhang, Q. Ren, and Y.-p. Zhao, “Modelling analysis of surface stress on a rectangular cantilever beam”, *J. Phys. D Appl. Phys.* **37**:15 (2004), 2140–2145.

Received 1 Apr 2013. Revised 22 May 2013. Accepted 3 Jun 2013.

KANG FU: fukang@dlut.edu.cn

Department of Engineering Mechanics, Dalian University of Technology, Dalian 116023, China

## SUBMISSION GUIDELINES

### ORIGINALITY

Authors may submit manuscripts in PDF format online at the Submissions page. Submission of a manuscript acknowledges that the manuscript is original and has neither previously, nor simultaneously, in whole or in part, been submitted elsewhere. Information regarding the preparation of manuscripts is provided below. Correspondence by email is requested for convenience and speed. For further information, write to [contact@msp.org](mailto:contact@msp.org).

### LANGUAGE

Manuscripts must be in English. A brief abstract of about 150 words or less must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate), and email address if available. A home-page URL is optional.

### FORMAT

Authors can use their preferred manuscript-preparation software, including for example Microsoft Word or any variant of  $\text{T}_{\text{E}}\text{X}$ . The journal itself is produced in  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ , so accepted articles prepared using other software will be converted to  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  at production time. Authors wishing to prepare their document in  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  can follow the example file at [www.jomms.net](http://www.jomms.net) (but the use of other class files is acceptable). At submission time only a PDF file is required. After acceptance, authors must submit all source material (see especially Figures below).

### REFERENCES

Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of  $\text{BibT}_{\text{E}}\text{X}$  is preferred but not required. Tags will be converted to the house format (see a current issue for examples); however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations; authors can supply their own links in addition to those provided by the editorial process.

### FIGURES

Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector format for all diagrams and graphs in your manuscript: vector EPS or vector PDF files are the most useful. (EPS stands for Encapsulated PostScript.)

Most drawing and graphing packages—Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.—allow the user to save files in one of these formats. Make sure that what you're saving is vector graphics and not a bitmap. If you need help, please write to [graphics@msp.org](mailto:graphics@msp.org) with as many details as you can about how your graphics were generated.

Please also include the original data for any plots. This is particularly important if you are unable to save Excel-generated plots in vector format. Saving them as bitmaps is not useful; please send the Excel (.xls) spreadsheets instead. Bundle your figure files into a single archive (using zip, tar, rar or other format of your choice) and upload on the link you been given at acceptance time.

Each figure should be captioned and numbered so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text (“the curve looks like this:”). It is acceptable to submit a manuscript with all figures at the end, if their placement is specified in the text by means of comments such as “Place Figure 1 here”. The same considerations apply to tables.

### WHITE SPACE

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

### PROOFS

Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

# Journal of Mechanics of Materials and Structures

Volume 8, No. 5-7

July–September 2013

---

- Efficiencies of algorithms for vibration-based delamination detection: A comparative study**      **OBINNA K. IHESIULOR, KRISHNA SHANKAR, ZHIFANG ZHANG and TAPABRATA RAY**      **247**
- Evaluation of the effective elastic moduli of particulate composites based on Maxwell's concept of equivalent inhomogeneity: microstructure-induced anisotropy**      **VOLODYMYR I. KUSHCH, SOFIA G. MOGILEVSKAYA, HENRYK K. STOLARSKI and STEVEN L. CROUCH**      **283**
- On successive differentiations of the rotation tensor: An application to nonlinear beam elements**      **TEODORO MERLINI and MARCO MORANDINI**      **305**
- Predicting the effective stiffness of cellular and composite materials with self-similar hierarchical microstructures**      **YI MIN XIE, ZHI HAO ZUO, XIAODONG HUANG and XIAOYING YANG**      **341**
- On acoustoelasticity and the elastic constants of soft biological tissues**      **PHAM CHI VINH and JOSE MERODIO**      **359**
- Identification of multilayered thin-film stress from nonlinear deformation of substrate**      **KANG FU**      **369**



1559-3959(2013)8:5;1-7