


*Journal of*  
***Mechanics of***  
***Materials and Structures***

*Volume 2, N° 2*

*February 2007*

 mathematical sciences publishers

# JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES

<http://www.jomms.org>

EDITOR-IN-CHIEF Charles R. Steele  
ASSOCIATE EDITOR Marie-Louise Steele  
Division of Mechanics and Computation  
Stanford University  
Stanford, CA 94305  
USA

## BOARD OF EDITORS

D. BIGONI University of Trento, Italy  
H. D. BUI École Polytechnique, France  
J. P. CARTER University of Sydney, Australia  
R. M. CHRISTENSEN Stanford University, U.S.A.  
G. M. L. GLADWELL University of Waterloo, Canada  
D. H. HODGES Georgia Institute of Technology, U.S.A.  
J. HUTCHINSON Harvard University, U.S.A.  
C. HWU National Cheng Kung University, R.O. China  
IWONA JASIUK University of Illinois at Urbana-Champaign  
B. L. KARIHALOO University of Wales, U.K.  
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  
Y. SHINDO Tohoku University, Japan  
A. N. SHUPIKOV Ukrainian Academy of Sciences, Ukraine  
T. TARNAI University Budapest, Hungary  
F. Y. M. WAN University of California, Irvine, U.S.A.  
P. WRIGGERS Universität Hannover, Germany  
W. YANG Tsinghua University, P.R. China  
F. ZIEGLER Technische Universität Wien, Austria

## PRODUCTION

PAULO NEY DE SOUZA Production Manager  
SHEILA NEWBERY Senior Production Editor  
SILVIO LEVY Scientific Editor

---

See inside back cover or <http://www.jomms.org> for submission guidelines.


---

Regular subscription rate: \$500 a year.

Subscriptions, requests for back issues, and changes of address should be sent to Mathematical Sciences Publishers, 798 Evans Hall, Department of Mathematics, University of California, Berkeley, CA 94720-3840.

---

©Copyright 2007. Journal of Mechanics of Materials and Structures. All rights reserved.

 mathematical sciences publishers

## **BEHAVIOR OF PAPER ON A HIGH SPEED CONVEYOR SUBJECTED TO AIR JET IMPINGEMENT: A METHOD FOR ESTIMATING BENDING STIFFNESS**

M. K. RAMASUBRAMANIAN, RICHARD A. VENDITTI AND KALYAN C. KATURI

Stiffness based sorting of mixed paper waste at high speed is desirable for sorting board, cardstock, and paper from each other for efficient recycling. We propose a quick, non-contact method for measuring the relative stiffness of paper samples moving on a conveyor at high speeds. The method consists of impinging an air jet on a moving paper sample while it is crossing a gap during transfer from one conveyor belt to another. The deflection caused by the impinging air jet can be measured and used to categorize different grades of paper. In this paper, we report results from a simulation of the proposed method using a finite element model. Influence of conveyor speed, sample orientation and nozzle pressure is studied using the model. Results indicate that the stiffness estimation has sufficient resolution to distinguish between board and paper grades at high conveyor speeds.

### **1. Introduction**

The paper recycling industry is important world-wide, providing both economic and environmental benefits. Paper recycling significantly decreases landfill utilization rates and avoids landfilling costs and in many cases reduces the energy requirements to produce papermaking fibers relative to pulping wood. According to the AFPA Recovered Paper Statistical Highlights (2005 edition), in the year 2004 approximately 50 million tons of paper were recycled out of the 102 million tons available in the United States. This is a 49.5% recovery rate. Paper recycling thereby increases the efficient usage of natural resources. Optimized paper recycling requires a homogenous recovered paper feed stream [Pryor 2000]. Often, co-mingled paper is recovered as a convenient method to encourage higher recovery rates of paper despite this. The co-mingled paper of different grades can be sorted either manually or through automatic sorting processes [Arzoumanian 2000] so as to improve the subsequent economics of paper recycling. The quality of the recycled paper depends greatly on the effectiveness of the sorting procedure.

Typically, paper products are recycled into the same type of products. For instance recovered newspapers are recycled into newsprint, corrugated boxes are recycled into new boxes, sorted office waste is recycled into printing and writing grades. The reason is that each of these grades utilizes fibers of different chemistry and quality. Newsprint paper utilizes fibers that are short and essentially have the same composition as wood. These inexpensive fibers have high contents of lignin which cause the paper to yellow over time. Corrugated containers use longer fibers that have been chemically pulped. Products like copy paper must use highly bleached, high brightness fibers. Comingled waste has low value and can

---

*Keywords:* paper recycling, bending stiffness, paper sorting, copy paper, finite element model, cardboard.

Research supported by the U.S. Department of Energy under the Industries for the Future Program, Forest Products Industry Agenda 2020; project number DE-FC07-00ID13880.

only be utilized for the most inexpensive products with low quality demand such as tube stock or as board. The implementation of improved and more economical sorting processes will allow the conversion of co-mingled waste to homogenous streams that can be diverted to the most appropriate target product recycling process.

Sorting can also potentially be used to distinguish and separate common types of papers that need to have special consideration when recycled. These could include coated papers for publication, colored papers, wax coated papers and many other types. Currently sensors are available for sorting glossy paper, office paper, newsprint, and colored paper from the input feed. The glossy paper is sensed by using a sensor which measures the intensity of the glare from the surface of the paper sample. Colored paper can be identified by using a high speed camera or matrices of discrete color sensors. A lignin sensor which can distinguish between office paper and newsprint was described in [Ramasubramanian et al. 2005; Mallapragada et al. 2004; Nilsson et al. 2001; Roman and Jeffers 1992]. Large amounts of lignin, a fluorescing chemical compound, is present in newsprint fibers and not in copy paper fibers. The lignin sensor measures the amount of fluorescent radiation from the sample by first exciting the sample using a laser source and measuring the fluorescence intensity using a photo multiplier detection system.

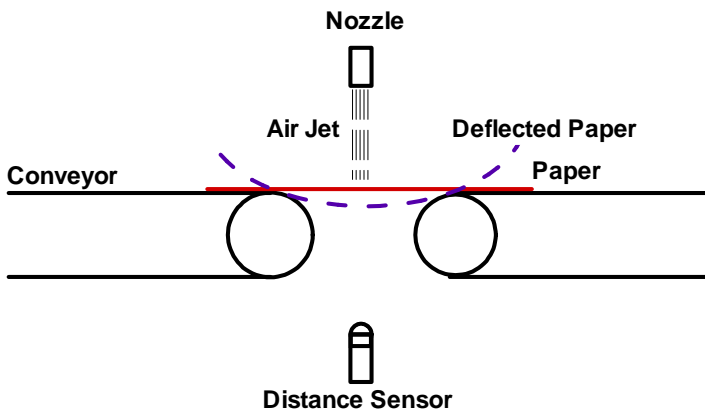
The efficiency of the sorting procedure will be maximum when all the above mentioned sensors along with stiffness sensor work to provide input data to a central decision making algorithm residing in a micro-controller. The decision making algorithm can then be used for sample identification and for the actuation of the mechanisms to move the paper to different collection bins based on paper mill requirements [Singhal et al. 2001].

**1.1. Stiffness sensing techniques.** There has been a lot of development in the field of on-machine paper stiffness sensing. Most of the developed techniques are meant for testing paper webs in the paper manufacturing industry. The goal is to determine paper strength properties since they directly influence the quality of the paper and process parameters. The elastic stiffness properties of paper correlates well with paper strength properties. The stiffness properties of paper can be found by studying the dispersion curves of Lamb waves. Lamb waves are surface waves in plates. Dispersion characteristics of lamb wave modes in machine made paper were discussed by [Habeger et al. 1979] . Air coupled transducers can be used to generate Lamb waves on paper surface. However these transducers are not well suited for on-line implementation due to air turbulence at the surface of the fast-moving web. The paper thickness should also be high to excite lamb waves using these transducers [Habeger et al. 1979]. When a laser pulse is focused on the top surface of the paper, a transient displacement field results due to localized thermal expansion. Thermoelastic excitation of Lamb waves along the principal directions in machine made paper has been discussed in [Cheng and Berthelot 1996; Brodeur et al. 1997; Johnson 1996]. Ridgway et al. [1999] have compared the performance of three different interferometry techniques in measuring laser generated elastic wave phase velocities on paper moving at production speeds. Slight discrepancies were found in the non-contact measurements due to humidity variations. The ultrasonic wave velocities were also found to be very sensitive to water content and temperature of the paper. Walter et al. [2000] have used a Fabry–Perot detector and a tracking mirror to increase the optical efficiency of the detection system. This setup has avoided the ablation of paper material by not using a high intensity laser to generate elastic waves on paper surface. The determination of bending stiffness of paper using laser ultrasonics was demonstrated in [Cornwell et al. 2001]. A pulsed laser was used to excite lamb

waves and fiber optic Mach–Zehnder interferometer was used to detect the normal component of lamb waves. The dispersion curve obtained from the wavelet transform of the signal was compared to the material dispersion curve calculated using the material constants and both the curves were found to be in good agreement. Ridgway et al. [2003] estimated the flexural rigidity of paper web running at 5000 ft/min by measuring the frequency dependence of the phase velocity of  $A_0$  mode lamb waves. While this method worked for fine paper grades, the low frequency signal amplitude from paperboard was found to be insufficient for making accurate measurements.

All the above mentioned techniques are meant to be used for testing the strength properties of paper webs of almost constant thickness values, that is, the thickness of the samples does not change drastically between two consecutive measurements. This is not the case in paper sorting. Unlike paper webs, the thickness of paper on a sorting conveyor varies widely from one sample to another. One sample might have high thickness like cardboard and the next one can be a finer grade paper such as copy paper. Therefore, the transducer input energy needs to be varied depending on the sample for getting better signal to noise ratio for each sample, which is difficult to achieve. In order to sort recovered paper, there is no need to find the exact elastic constants of each and every sample. The relative stiffness values should be sufficient to sort the paper into different bins and hence eliminates the need for complex and expensive equipment.

**1.2. Proposed method.** The relative bending stiffness sensor setup is shown in Figure 1. The idea is to apply a constant pneumatic pressure load on top of the paper sample spanning two supports that are a known distance apart. The two supports are two belt conveyors which are positioned one after another with a small clearance between them. The deflection of the sample to the load is inversely proportional to the stiffness of the sample. A distance sensor can be used to measure the deflection of paper due air jet impingement. Any of the distance sensing techniques such as optical, ultrasonic or RF could be used for measuring the paper deflection. The proposed method will be useful in distinguishing the samples based on their relative bending stiffnesses. While a static load-deflection curve can be generated and modeled using elementary mechanics, studying the method under dynamic conditions is the focus of this study.

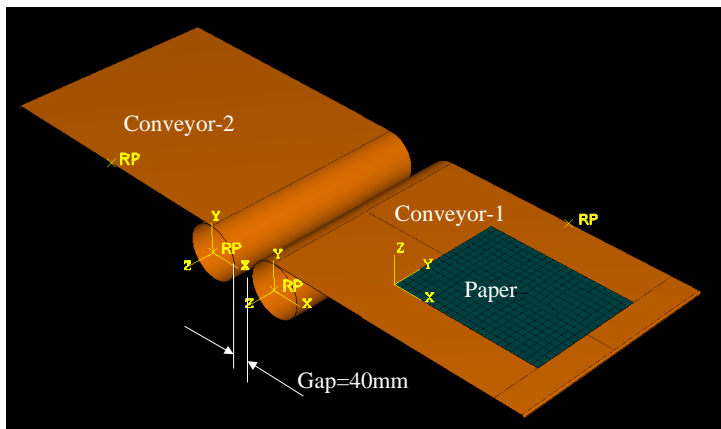


**Figure 1.** Bending stiffness sensing setup.

## 2. Finite element model

A finite element model of the paper-conveyor system was built to provide a detailed understanding of the sensor concept. A parametric study using the finite element model was carried out to provide a clear insight into the process parameters and their influence on the deflection of the sample. A macroscopic modeling approach was adopted for modeling the paper. This presupposes a homogeneous deformation field and does not account for local phenomenon [Ramasubramanian and Muthuraman 2003]. The model was built in ABAQUS/Explicit environment which is suitable for transient dynamic analysis problems such as this one. Direct-integration dynamic analysis procedure was used for finding the deflection of paper samples over the conveyor gap for various conveyor speeds and air jet loadings as it is computationally efficient for the analysis of large models with relatively short dynamic response times [ABAQUS 2003]. It is also suitable for the analysis of discontinuous events such as the current model, which contains several contact interactions that change rapidly during the analysis [Sun et al. 2000; Zhong 1993; Hughes 1987].

**2.1. Model geometry.** The paper-conveyor system was modeled in 3D space. The high speed conveyor is modeled using analytical rigid elements. Rectangular shaped analytical rigid surface is used for modeling the conveyor belt surface. The conveyor roller surfaces are modeled separately from the conveyor belt surface by using cylindrical analytical rigid surfaces. This discretization of the conveyor surface allows the specification of separate boundary conditions for flat and cylindrical surfaces of the belt conveyor. A clearance of 40 mm is specified between the roller surfaces of the two conveyors. The flat conveyor surface is constrained in all directions whereas an angular velocity boundary condition is specified for both the conveyor rollers. This boundary condition simulates the rotating roller surface of the actual conveyor. The angular velocity of the rollers is computed from the speed of the conveyor. For a given value of conveyor speed, the angular velocity of the conveyor is found from the expression  $\omega = V/R$ , where  $\omega$  is the angular velocity of the roller,  $V$  is the tangential velocity of the conveyor belt and  $R = 4''$  is the roller radius. Paper is modeled as a thin rectangular plate sitting flat on the right conveyor belt



**Figure 2.** Assembled view of the model.

surface in the 1-3 plane. The assembled model consisting of paper and conveyor surfaces is shown in Figure 2.

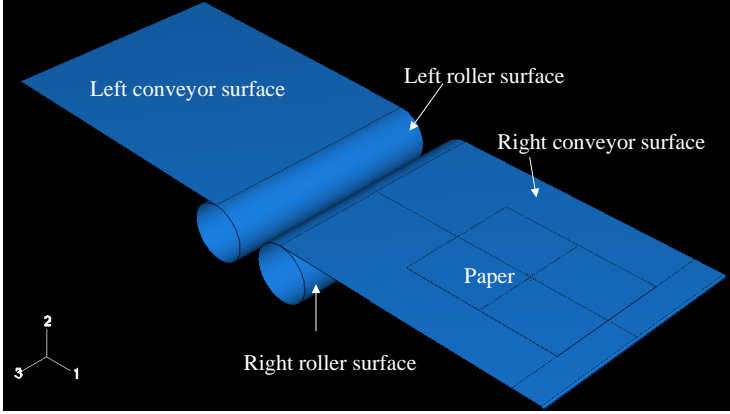
To simulate the high-speed moving conveyor, a uniform initial velocity field is specified on the nodes of the paper sample. The magnitude of the velocity is specified to be equal to the tangential velocity of the conveyor belt. But in the real paper-conveyor system, the paper sample moves along with the conveyor belt and there will not be any relative motion between the two until the paper reaches the end of the conveyor. To meet this requirement, the interaction between the flat conveyor surface and the paper surface is modeled as frictionless.

**2.2. Contact model.** The contact behavior between the paper and the conveyor surfaces is modeled by specifying contact interactions between the surfaces that might come into contact during the course of the analysis. The Contact modeling involved two steps. In the first step, various surfaces that might come into contact during the analysis procedure are identified. Then these surfaces are coupled together by specifying them as contact pairs. In the second step, mechanical property models are assigned to the contact pairs. The contact property model specifies the normal and tangential behavior of the surfaces when they come into contact. The normal behavior between the paper surface and the conveyor surface is specified as hard. The “Hard” contact relationship minimizes the penetration of the paper element nodes into the conveyor surface and does not allow transfer of tensile stress across the interface. When the surfaces are in contact, any contact pressure can be transmitted between them. The surfaces separate if the contact pressure reduces to zero. Separated surfaces come into contact when the clearance between them reduces to zero. Kinematic predictor/corrector contact algorithm is used to enforce contact constraints. Finite sliding approach which allows arbitrary motion of the surfaces forming the contact pair is used for specifying the sliding between any two contact surfaces in the model.

Classical isotropic Coulomb friction model is used to specify the tangential behavior of the contacting surfaces. The standard Coulomb friction model assumes that no relative motion occurs if the equivalent frictional stress is less than the critical stress,  $\tau_{critical}$  which is proportional to the contact pressure,  $p$ , in the form  $\tau_{critical} = \mu p$ , where  $\mu$  is the friction coefficient that is defined as a function of the contact pressure  $p$ . The tangential behavior of the surface is defined by specifying the coefficient of friction between the contacting surfaces. For determining the coefficients of friction between the paper samples and the conveyor, the surface roughness of the conveyor and the paper needs to be known. The surface properties of the paper samples vary widely during the sorting procedure and it is hard to obtain the coefficients of friction for all possible surface combinations. Therefore, simplifying assumptions were made by considering the friction to be isotropic and to be independent of the contact pressure and other field variables. The surface roughness of the paper surface is less than the surface roughness of the conveyor roller. The coefficient of friction of driving pulley of steel for industrial conveyor is found to be equal to 0.15 from the specifications of the conveyor. The static coefficient of friction between the conveyor and the paper surface is assumed to be equal to half of that, 0.07, and the kinetic friction coefficient was set to be equal to 0.04. These friction coefficients are not the exact values. The actual values can only be found by first choosing the conveyor and measuring the friction coefficient for the conveyor and different paper grades that might be found in the input feed. The effective friction coefficient between the contacting surfaces is assumed to satisfy the equation described by [Oden and Martins 1985]:

$$\mu = \mu_k + (\mu_s - \mu_k)e^{-d_c \dot{\gamma}_{eq}}, \quad (1)$$

where  $\mu_k$  is the kinetic coefficient of friction,  $\mu_s$  is the static coefficient of friction,  $d_c$  is the decay coefficient and  $\dot{\gamma}_{eq} = (\dot{\gamma}_1^2 + \dot{\gamma}_2^2)^{-1/2}$  is the slip rate with  $\dot{\gamma}_i$  is the slip rate in the direction  $i$ . The decay coefficient specifies the rate at which the effective friction coefficient value reaches the value of kinetic coefficient of friction once the slip starts to occur. Totally four surfaces are identified in the Paper-conveyor model that can come into contact with the paper sample during the analysis, see Figure 3.



**Figure 3.** Potential contact surfaces for the paper sample.

Two mechanical contact interaction properties were created. The first property has hard normal behavior and frictionless tangential behavior. For the second property, the tangential behavior is specified to follow Equation (1). Four contact pairs are created and the interaction properties are assigned, see Table 1.

The tangential behavior between the paper surface and the conveyor surface is modeled as frictionless because there is no relative motion between the paper and the conveyor surface. Both of them will be moving at the same speed. There will be some slip between the paper surface and the left conveyor surface because the velocity of the paper will be reduced due to the applied pneumatic load. But that slip is neglected because its effect on the paper deflection is insignificant.

Contact Pair	Interaction Property	
	Tangential Behavior	Normal Behavior
Paper-Right conveyor surface	Frictionless	Hard
Paper-Right roller surface	Friction specified	Hard
Paper-Left roller surface	Friction specified	Hard
Paper-Left conveyor surface	Frictionless	Hard

**Table 1.** Contact interactions.



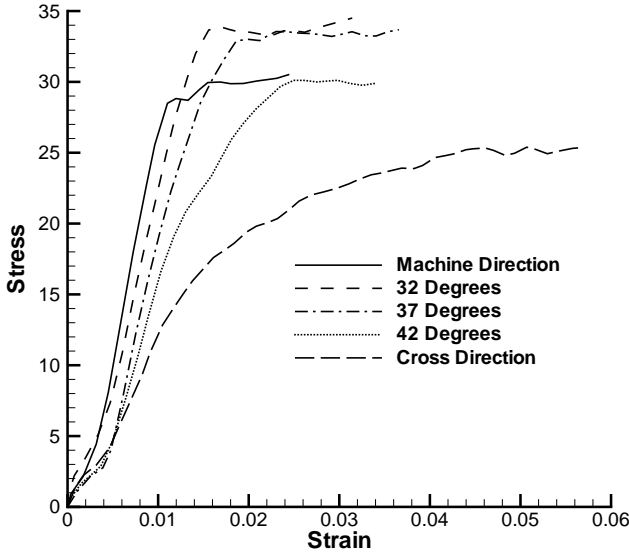
**2.3. Material model.** Even though paper is highly anisotropic at the microscopic level, it is treated as orthotropic at the macroscopic level [Makela and Ostlund 2003; Leppänen et al. 2005]. During the paper manufacturing process, most of the fibers are oriented in a direction parallel to the length of the paper web which is the Machine Direction (MD). The Cross Direction is perpendicular to the MD. The MD and CD of the paper are the in-plane principal material directions of the paper. The Young's modulus in MD is significantly higher than that of CD because of preferential alignment of the fibers. The constitutive relation for paper is

$$\begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{23} \end{bmatrix} = \begin{bmatrix} 1/E_{MD} & -\nu_{21}/E_{CD} & -\nu_{31}/E_{ZD} & 0 & 0 & 0 \\ -\nu_{12}/E_{MD} & 1/E_{CD} & -\nu_{32}/E_{ZD} & 0 & 0 & 0 \\ -\nu_{13}/E_{MD} & -\nu_{23}/E_{CD} & 1/E_{ZD} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/G_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/G_{13} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/G_{23} \end{bmatrix} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{13} \\ \sigma_{23} \end{bmatrix},$$

where the subscripts 1, 2 indicate machine and cross-machine directions.  $E_{MD}$ ,  $E_{CD}$  and  $E_{ZD}$  are the Young's moduli in the machine, cross-machine and thickness directions, respectively.  $G_{12}$  is the in-plane shear modulus and  $G_{13}$ ,  $G_{23}$  are the out-of-plane shear moduli.

The properties of the paper samples are obtained by mechanical testing. Three grades of paper that are commonly found in the recovered paper were tested using the Minimat tensile tester (Rheometrics, Inc). Thin strips of paper of size 1 cm  $\times$  5 cm were cut both in machine and cross machine directions and tested to get the Young's moduli. Flat faced screw clamps were used to hold the paper strips. The tests were conducted in a non-conditioned environment at a cross-head speed of 1 inch/min. The thicknesses of paper grades were measured with a hard platen. Engineering estimates were made for the Poisson's ratio and the out of plane shear moduli of the three paper grades based on the results obtained by [Mann et al. 1980]. Since the effect of transverse shear deformation is negligible for very thin samples such as paper, the out of plane shear moduli is assumed to be same for all the three grades. An optimization technique as discussed by [Seo et al. 1992] was used for finding the in plane shear modulus. The goal of this method is to perform the measurements of the Young's moduli in the optimum angular region (the so-called error boundary) to achieve the best accuracy in the determination of  $G_{12}$ . According to this method, the anisotropy of the Young's moduli was found by finding the ratio of  $E_{MD}$  to  $E_{CD}$  and it was used for finding the minimum error boundary region. The stress-strain curves were obtained at an angle  $\theta$  (minimum error angle),  $\theta - 5^\circ$  and  $\theta + 5^\circ$  in addition to machine and cross-machine directions; see Figure 4. A least-squares algorithm was used to estimate the shear modulus best fitted to the experimental data: see Table 2.

Since the lateral (in-plane) dimensions of the paper are much larger than the thickness of the paper, plane stress shell elements are used for modeling the paper surface. These elements also allow transverse shear deformation. Kirchoff's thin shell theory is used for the 3D paper model, implying that a material line that is originally normal to the mid-surface of the shell elements remains so throughout the deformation. Transverse shear stress acts as penalty function to impose Kirchoff's constraints. The shell elements also account for finite membrane strains and arbitrarily large rotations. Each element has four nodes at which element variables are calculated and use linear interpolation in each direction to find the displacements at any other point in the element. Each node has six degrees of freedom which include



**Figure 4.** Light-weight paper stress strain curves in different directions relative to the machine direction.

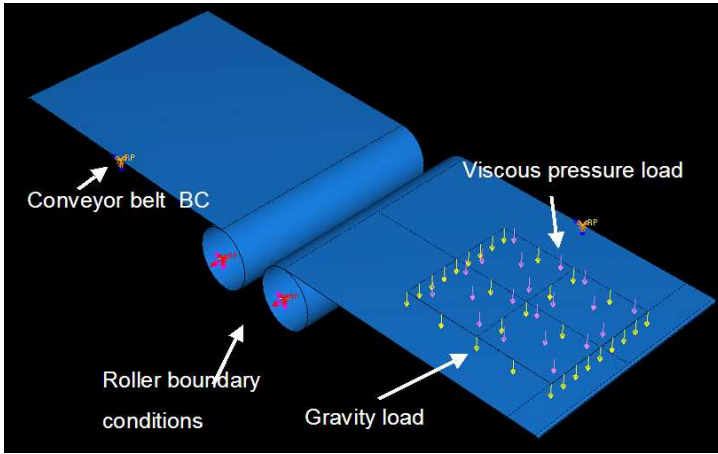
both the rotational and translational degrees of freedom. Even though the shell elements allow transverse shear deformation, it becomes very small as the shell thickness decreases. Therefore, the error due to the estimated out of plane shear moduli is negligible. The in-plane dimensions of the paper sample are specified to be equal to the standard letter paper size:  $11'' \times 8.5''$ . The machine direction of the paper is specified to be parallel to the length of the paper and cross-machine direction is specified to be parallel to the width of the paper.

**2.4. Loading conditions.** The external loads acting on the paper sample are: the load due to surrounding air, weight of the paper and the pneumatic load that acts on the paper when it passes over the gap. The

	Light-weight paper	Medium card stock	Heavy card stock	Method
Thickness, $\mu m$	105	206	229	Measured
Basis weight, $g/m^2$	75	145	200	Measured
$E_{MD}$ , GPa	3.98	1.69	1.79	Measured
$E_{CD}$ , GPa	1.27	1.11	1.09	Measured
$\nu_{12}$	0.15	0.15	0.15	Estimated
$G_{12}$ , GPa	0.93	0.71	0.63	Computed
$G_{13}$ , GPa	0.35	0.35	0.35	Estimated
$G_{23}$ , GPa	0.35	0.35	0.35	Estimated

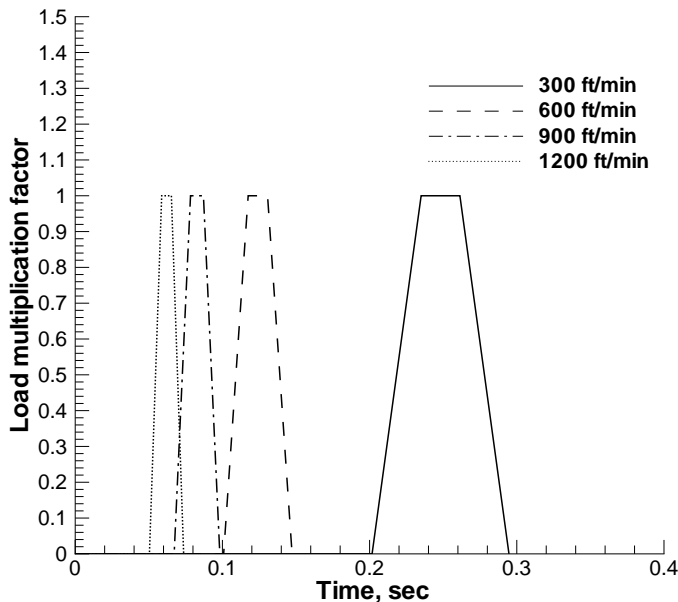
**Table 2.** Input parameters.

load due to the air surrounding the paper sample is modeled as a viscous pressure load whose magnitude is equal to the atmospheric pressure; see Figure 5.



**Figure 5.** Specified loading and boundary conditions.

The pneumatic load due to the nozzle does not act continuously on the paper sample but acts when the sample jumps over the clearance between the two conveyors. To simulate this loading behavior, a set of elements was created by specifying circular partition on the top surface of the paper and the time history of the load on the element set was prescribed by using the amplitude curves. The amplitude curve specifies the load on a set of elements as a function of time. By initially positioning the paper sample at a known distance from the center of the gap, the position of the paper sample as a function of the analysis



**Figure 6.** Amplitude curves for different conveyor speeds.

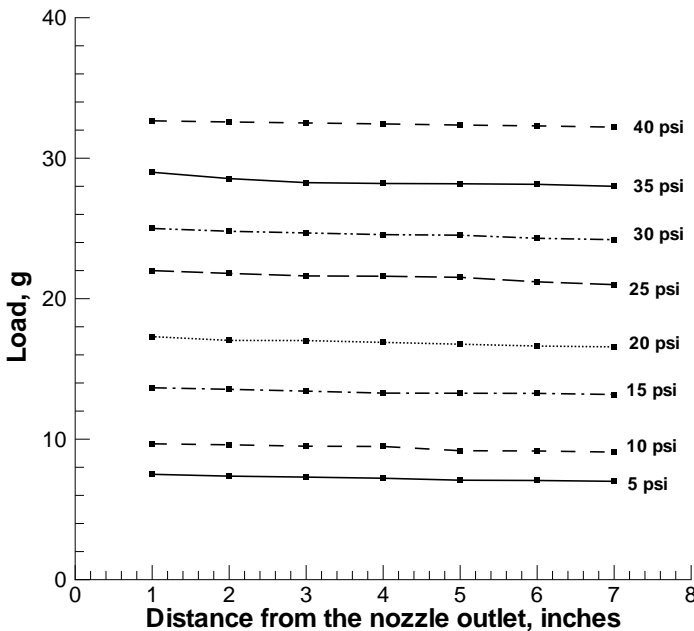
Nozzle inlet pressure, psi	Load intensity N/m <sup>2</sup>
10	1293.91
20	2391.52
25	3952.28
30	4641.43

Nozzle height = 1"  
Spread area = 76 ± 0.73 mm<sup>2</sup>

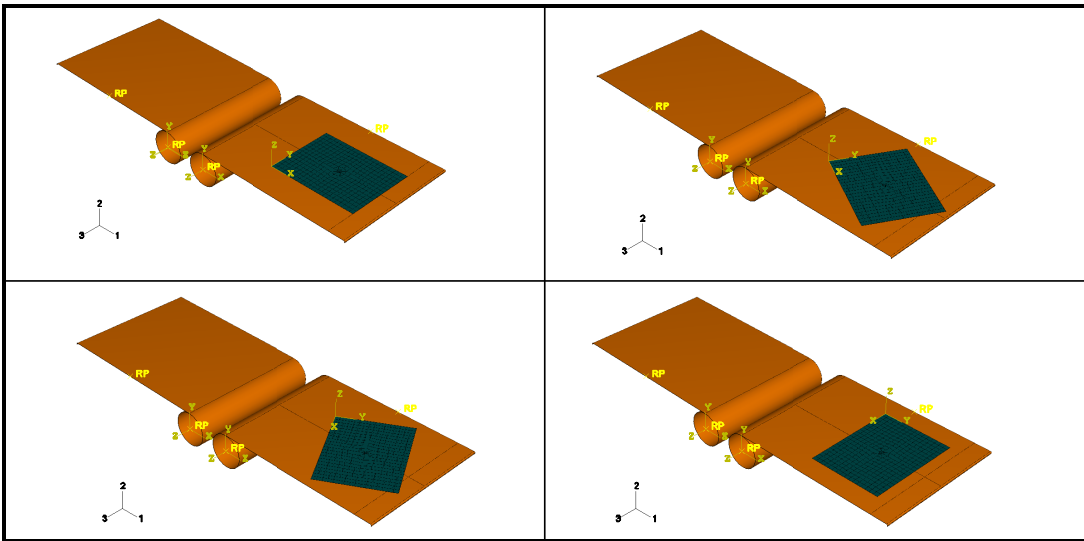
**Table 3.** Cylindrical spread nozzle data.

time can be roughly known from the initial velocity of the paper sample. This information was used to construct the amplitude curves for different conveyor speeds; see Figure 6.

At any time during the analysis, the actual load applied on the element set will be equal to the product of the maximum load value given on the input data line and the load multiplication factor at that instant of time, which was obtained from the amplitude curve. In this way, the paper samples will be loaded only when they reach the gap between the two conveyors. To obtain the load intensity and spread area data, several nozzles with different profiles were investigated to find effective nozzle for the sorting procedure. A cylindrical stream nozzle (Series 544 with orifice diameter  $d = 0.032$  inches from Lechler, Inc.) which produces a concentrated direct stream of air was selected and the load distribution data of



**Figure 7.** Cylindrical nozzle load profile.



**Figure 8.** Paper sample with its machine direction oriented at  $0^\circ$ ,  $30^\circ$ ,  $60^\circ$ ,  $90^\circ$  to the conveyor travel direction.

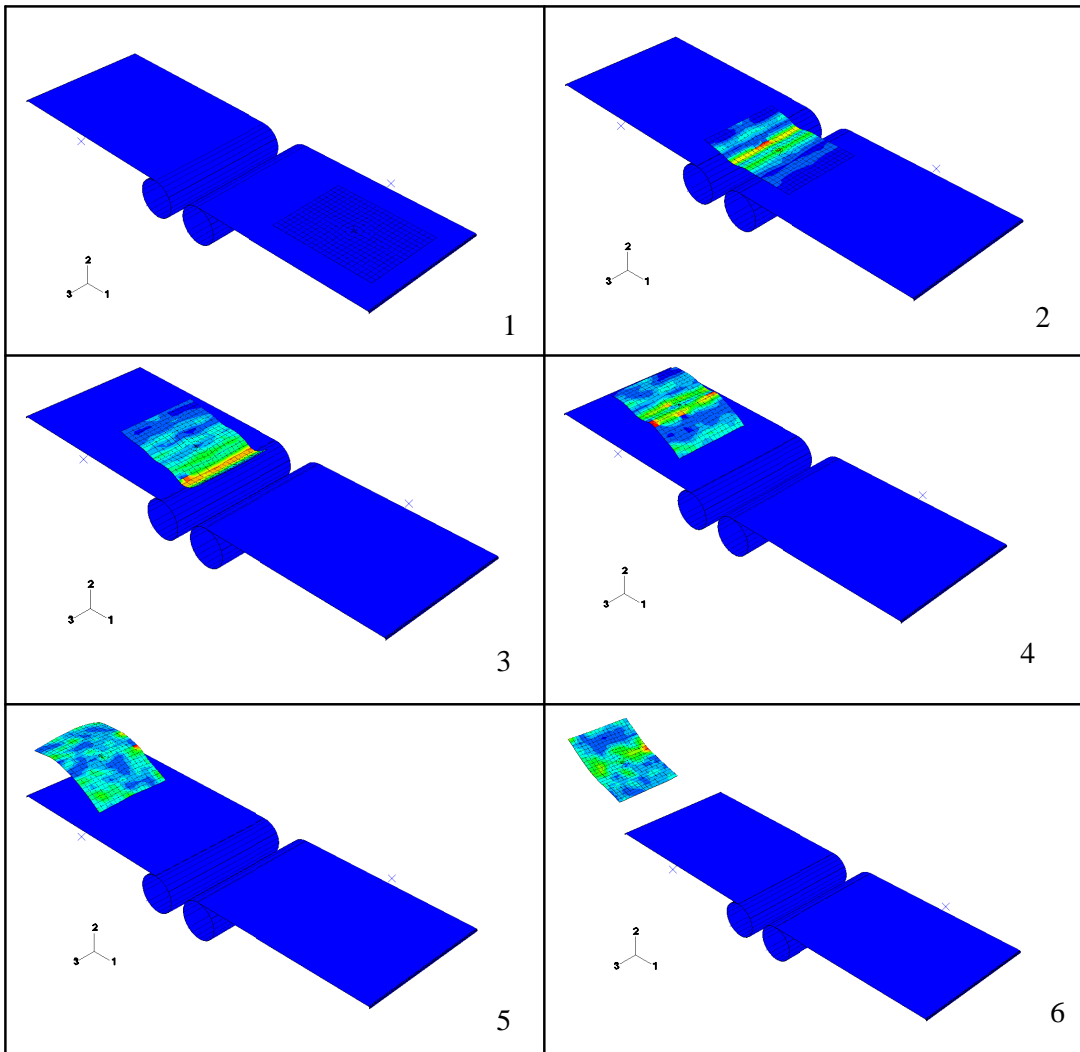
the nozzle was obtained experimentally. This data was used for specifying the circular partition diameter and the maximum load intensity on the elements in the set, see Table 3 and Figure 7.

**2.5. Simulation.** The deflection values were obtained for three paper grades as a function of load, conveyor speed and the orientation of the samples on the conveyor belt. Deflection values of the samples were found at four nozzle loads. All the nozzle loads used in the simulation were obtained from the load intensity data of the cylindrical profile nozzle when it was operating at different inlet pressures and held at a distance of one inch above the conveyor surface. The samples were loaded when the machine direction was oriented parallel to conveyor travel direction, and when machine direction was at  $30^\circ$ ,  $60^\circ$ ,  $90^\circ$  to the conveyor travel direction; see Figure 8. The deflection behavior of the paper for all the four loading conditions was observed at the conveyor speeds of 300 ft/min and 1200 ft/min.

The deflection values are measured relative to the flat conveyor surface. The deflections are considered positive when the paper rises above the flat conveyor surface and negative when the paper deflects downwards below the conveyor surface due to the applied load. This happens when the paper spans the gap between the two conveyors.

### 3. Results

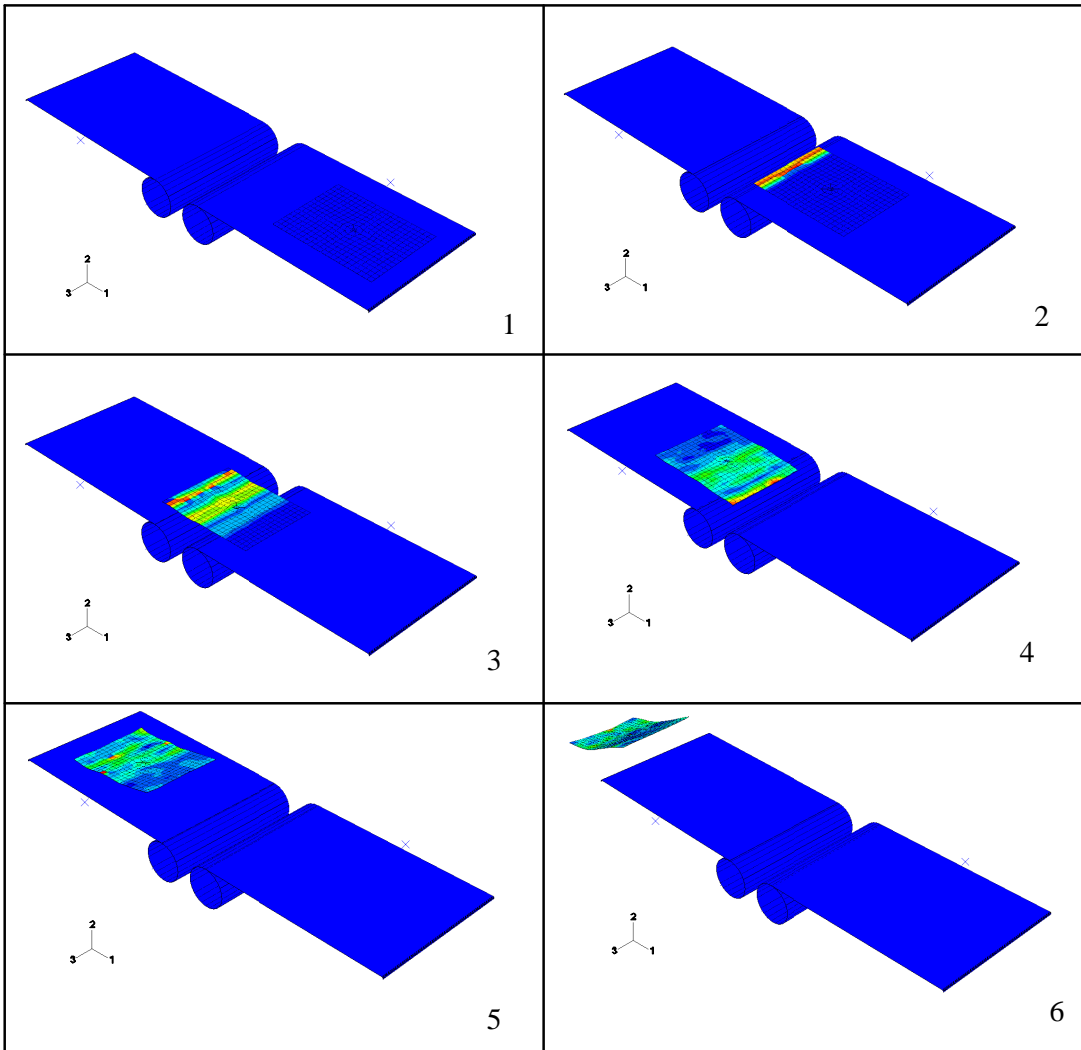
In order to identify the effects of atmospheric pressure on the paper sample, a finite element model was created with all the loading and boundary conditions as mentioned earlier except for the atmospheric pressure and the external pneumatic load. Figure 9 shows the results: (1) shows the state of the paper at the start of the simulation; (2) shows the state of the paper after it has crossed the gap and has touched the second conveyor surface; (3), (4) and (5) show the sheet leaving the surface of the conveyor due to lack of damping force; (6) shows the sheet moving downwards due to gravity. In the next simulation the ambient air pressure was specified on the top surface of the paper sample. All the remaining boundary



**Figure 9.** Response of the paper sample moving at 300 ft/min in the absence of viscous pressure.

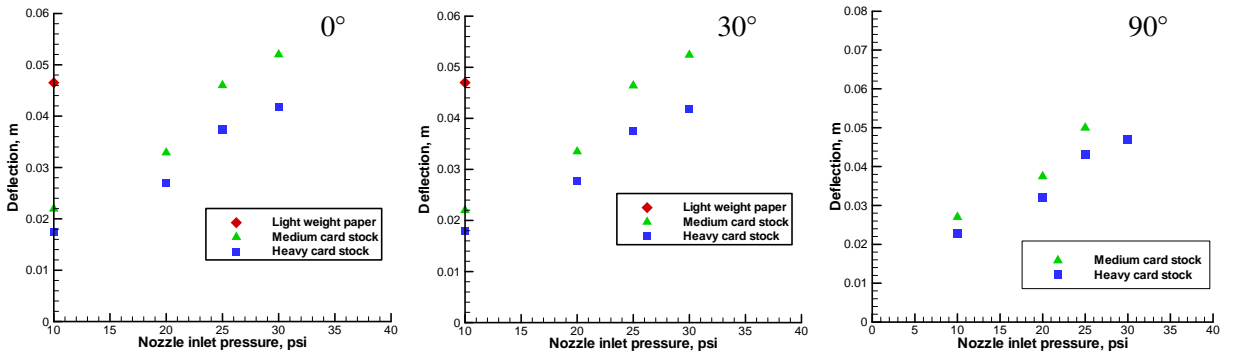
and loading conditions were kept the same as the earlier simulation. In this case (Figure 10), the paper does not leave the conveyor surface as in the earlier case; the pressure keeps the sheet on the conveyor after crossing over. This indicates the importance of the air pressure acting as a damping medium on the paper surface.

Next, two sets of simulations were run by specifying two different conveyor speeds. For each speed, the sample deflections are found at each of the four nozzle inlet pressures and at different orientations of the paper on the conveyor. The deflection of the paper due to applied pneumatic load is found by measuring the maximum displacement of the center node of the circular partitioned element set in the downward direction. For the first set of simulations, the conveyor speed is set to be equal to 300 ft/min and deflections were measured for the three paper grades due to the pneumatic loads acting at the nozzle

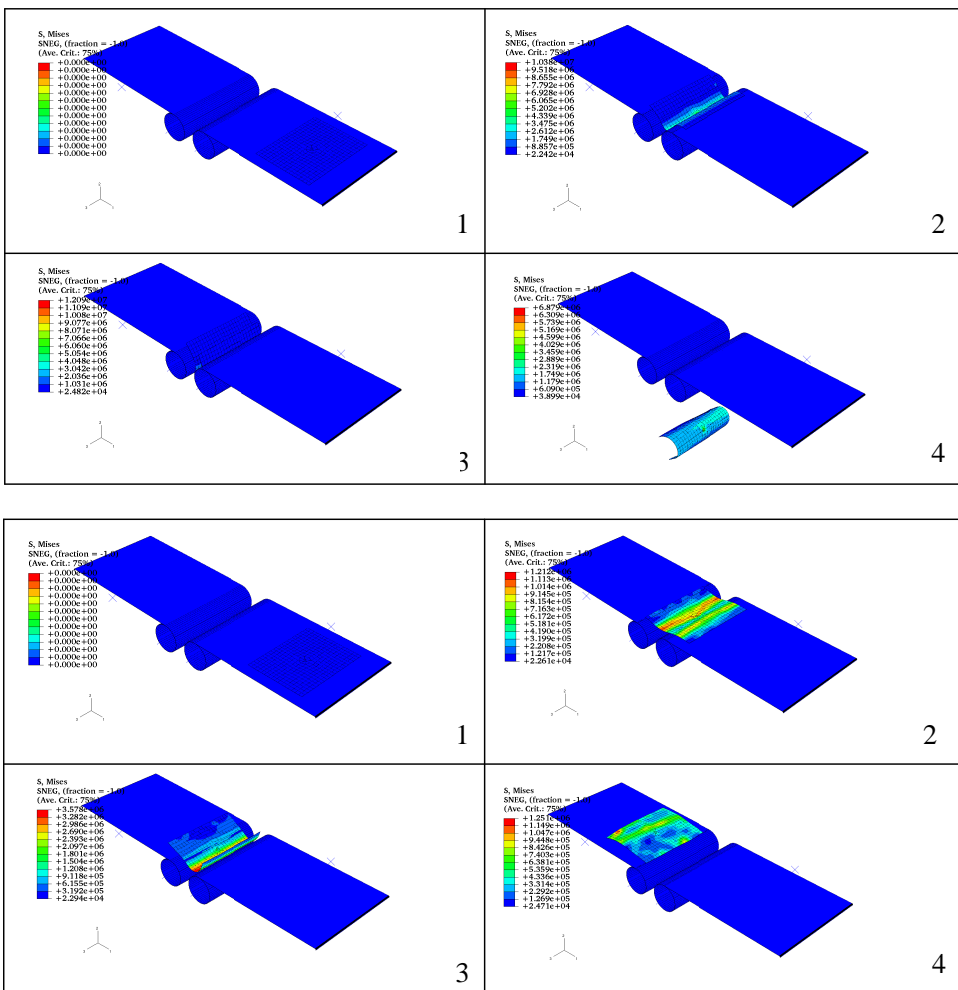


**Figure 10.** Response of the paper sample moving at 300 ft/min in the presence of viscous pressure.

inlet pressures of 10, 20, 25 and 30 psi, respectively. The light-weight paper was able to clear the gap only when its MD was less than  $30^\circ$  to the conveyor length direction and the nozzle input pressure was 10 psi. In all other cases, it was unable to clear the gap; see Figure 11, which shows paper deflection values for samples that were able to withstand the pneumatic load. The heavy card stock had no difficulty in clearing the gap when the load was applied, as was expected since its flexural rigidity is higher than that of the light-weight paper. The response of the light-weight paper and the heavy card stock when they were loaded in CD is shown in Figure 12. The medium card stock also failed to clear the gap at the nozzle inlet pressure of 30 psi; see Figure 11, right. The deflection values increase as the orientation of the paper changes from MD to CD.



**Figure 11.** Deflection of the paper samples moving at 300 ft/min. Angles indicate paper orientation (see Figure 8 for convention).



**Figure 12.** Response of paper loaded in CD and moving at 300 ft/min. Top four panes: light-weight paper; bottom: heavy card stock.

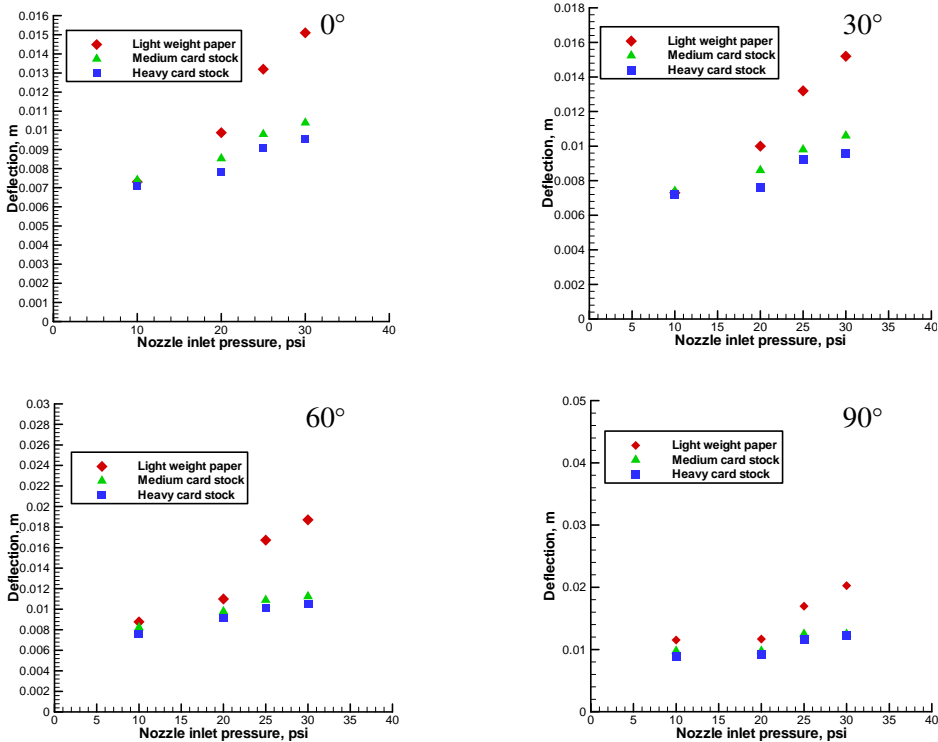


During the second set of the simulations, the speed of the conveyor was increased from 300 ft/min to the expected operating speed of 1200 ft/min. The loading and boundary conditions were kept the same. In this case all paper samples cleared the gap for all four nozzle; see Figure 13.

The deflection, however, decreased significantly compared to that of the samples moving at 300 ft/min. This is expected because the samples moving at 1200 ft/min are subjected to air jet load for a shorter amount of time than the samples moving at 300 ft/min. The deflections increased as the orientation of the machine direction of the paper sample increase from  $0^\circ$  to  $90^\circ$  with respect to the conveyor travel direction. This is because of the lower values of the Young's modulus of the sample in cross-machine direction when compared to the machine direction.

After the paper moves past the gap, the paper does not immediately land completely on the other conveyor surface because of the disturbance caused by the momentary application of the pneumatic load. This is illustrated in the Figure 14. The heavy card stock does not closely follow the conveyor surface at 1200 ft/min when compared to 300 ft/min.

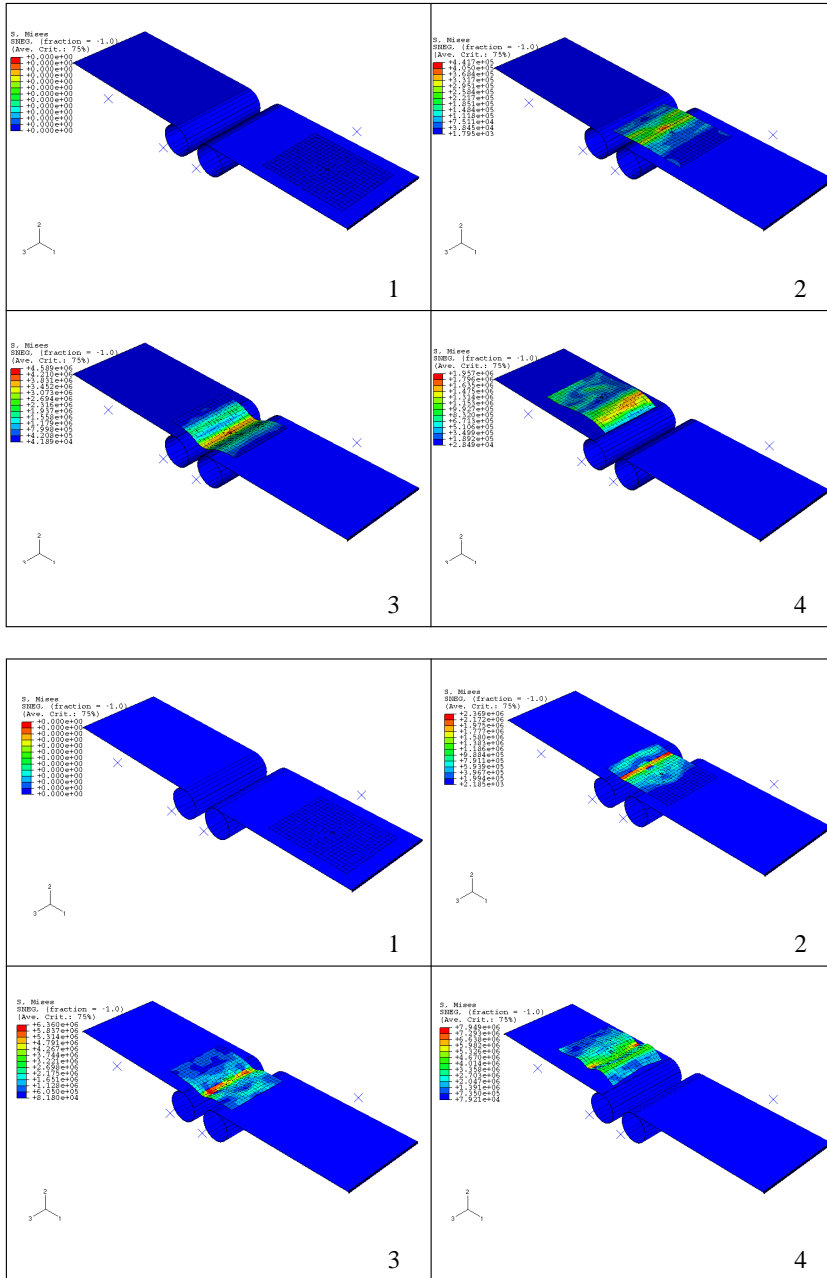
Another important aspect is the settling time which the samples take to settle completely on the left conveyor surface after clearing the gap. The samples moving at 300 ft/min took less time to settle on the conveyor surface than those moving at 1200 ft/min. The time histories of the three paper grades during the analysis at the two conveyor speeds are shown in Figure 15. It is evident from the response curves that the paper samples exhibit more positive displacement at 1200 ft/min as compared to 300 ft/min.



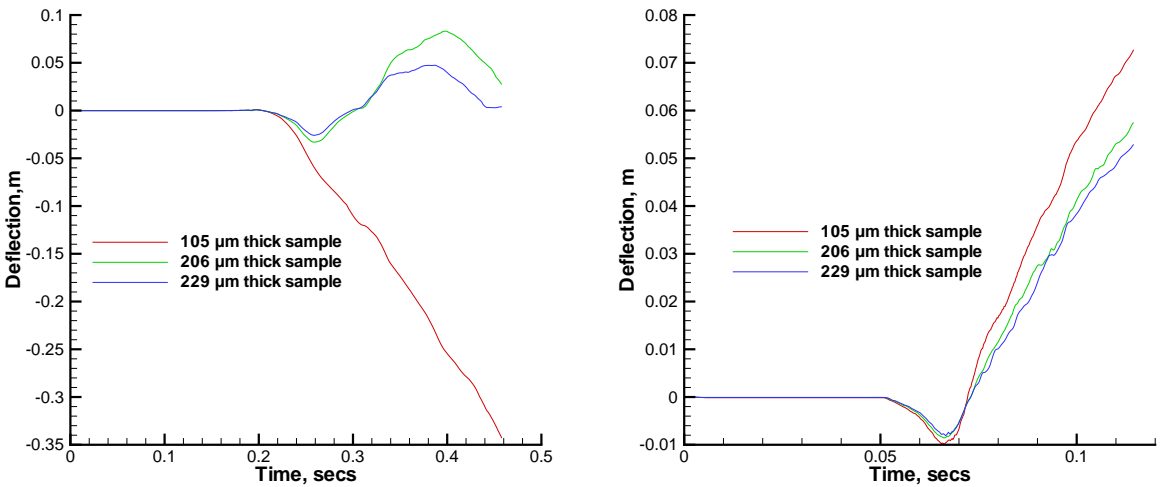
**Figure 13.** Deflection of the paper samples moving at 1200 ft/min. Angles indicate paper orientation.

The light-weight paper has higher negative deflection values since it failed to withstand the load when moving at 300 ft/min; see Figure 15.

The bending stiffnesses and deflections of the three paper grades in the machine direction were computed and normalized with respect to the corresponding values of light-weight paper at the conveyor



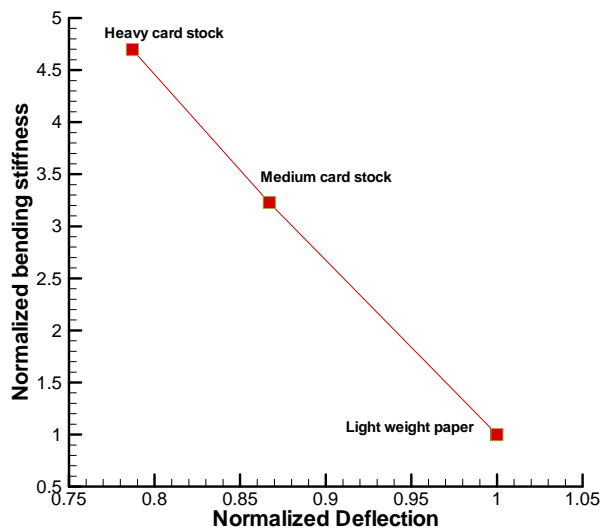
**Figure 14.** Response of heavy card stock when nozzle inlet pressure equals 20 psi. Top four panes: speed of 300 ft/min; bottom: 1200 ft/min.



**Figure 15.** Time history of paper grades at a nozzle inlet pressure of 20 psi. Left: conveyor speed of 300 ft/min; right: 1200 ft/min.

speed of 1200 ft/min and at the nozzle inlet pressure of 20 psi. As expected there was an inverse relation between the bending stiffness and deflections of three samples; see Figure 16.

The thickness of the paper has the most influence in determining the bending stiffness of the paper. The basis weight is directly proportional to the thickness of the paper. Therefore the higher the basis weight, the lower the deflection. The basis weight of the sample is also related to the density of the sample. Paper such as heavy card stock, with higher value of density, when moving on high speed conveyor, will have more kinetic energy when compared to finer grades of paper. The kinetic energy of the sample influences the deflection and hence heavy card stock deflects less when compared to light-weight paper.



**Figure 16.** Variation of the sample deflection with respect to the bending stiffness.

#### 4. Conclusions

A stiffness estimation method has been proposed for moving discrete paper samples on a high-speed conveyor. A finite element model of the sensing system, namely, the paper-conveyor system, with air impingement while the paper crosses over a gap, has been developed. The response of three dissimilar paper grades has been examined under different loading and boundary conditions. The influence of viscous pressure is shown to ensure that the sample will stick to the conveyor and successfully move with it and deflect under the influence of the impinging air jet. The deflection values estimated for the three grades at different speeds show that they could be ranked in terms of stiffness. Low stiffness samples, which fail to clear the gap at low conveyor speeds when the pneumatic load is applied, do clear the gap at high conveyor speeds. The orientation of the paper sample can be a confounding factor in decision making if the stiffness in MD and CD overlap between different grades, limiting the resolution that can be achieved in sorting. However, the method has sufficient resolution at operating speeds to distinguish between board and paper, which itself is a significant step in sorting mixed waste at high speeds.

#### References

- [ABAQUS 2003] *ABAQUS analysis: user's manual*, Version 6.4, Hibbit, Karlsson and Sorenson, Pawtucket, RI, 2003.
- [Arzoumanian 2000] M. Arzoumanian, "Sorting the scientific way", *Official Board Markets* **76**:29 (2000).
- [Brodeur et al. 1997] P. H. Brodeur, M. A. Johnson, Y. H. Berthelot, and J. P. Gerhardstein, "Noncontact laser generation and detection of Lamb waves in paper", *J. Pulp. Pap. Sci.* **23**:5 (1997), 238–243.
- [Cheng and Berthelot 1996] J.-C. Cheng and Y. H. Berthelot, "Theory of laser-generated transient Lamb waves in orthotropic plates", *J. Phys. D Appl. Phys.* **29**:7 (1996), 1857–1867.
- [Cornwell et al. 2001] M. Cornwell, Y. H. Berthelot, D. Griggs, and C. Habeger, "Noncontact determination of the bending stiffness of paper using laser ultrasonics and wavelet analysis: effects of moisture and temperature", pp. 747–750 in *IEEE Ultrasonics Symposium* (Atlanta, GA), vol. 1, edited by D. E. Yuhas and S. C. Schneider, IEEE, Piscataway, NJ, 2001. IEEE Catalog Number 01CH37263.
- [Habeger et al. 1979] C. C. Habeger, R. W. Mann, and G. A. Baum, "Ultrasonic plate waves in paper", *Ultrasonics* **17**:2 (1979), 57–62.
- [Hughes 1987] T. J. R. Hughes, *The finite element method: linear static and dynamic finite element analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1987.
- [Johnson 1996] M. A. Johnson, *Investigation of the mechanical properties of copy paper using laser generated and detected Lamb waves*, Ph.D. thesis, Georgia Institute of Technology, 1996.
- [Leppänen et al. 2005] T. Leppänen, J. Sorvari, A. Erkkilä, and J. Hämäläinen, "Mathematical modelling of moisture induced out-of-plane deformation of a paper sheet", *Model. Simul. Mater. Sci. Eng.* **13**:6 (2005), 841–850.
- [Makela and Ostlund 2003] P. Makela and S. Ostlund, "Orthotropic elastic-plastic material model for paper materials", *Int. J. Solids Struct.* **40**:21 (2003), 5599–5620.
- [Mallapragada et al. 2004] V. Mallapragada, M. K. Ramasubramanian, and R. Venditti, "A compact, high speed lignin sensor for the automated sorting of newsprint from mixed waste", pp. 1137–1147 in *TAPPI Paper Summit: Spring Technical and International Environmental Conference* (Atlanta, GA), 2004.
- [Mann et al. 1980] R. W. Mann, G. A. Baum, and C. C. Habeger, "Determination of all nine orthotropic constants for machine-made paper", *TAPPI J.* **63**:2 (1980), 163–166.
- [Nilsson et al. 2001] C. M. Nilsson, L. Malmqvist, and J. Carlsson, "Red fluorescence sensor for noncontact on-line measurements in paper production", *Opt. Eng.* **40**:8 (2001), 1674–1681.

- [Oden and Martins 1985] J. T. Oden and J. A. C. Martins, “Models and computational methods for dynamic friction phenomena”, *Comput. Methods Appl. Mech. Eng.* **52**:1–3 (1985), 527–634.
- [Pryor 2000] A. B. Pryor, “A paper sorting dream”, *Scrap* **57**:5 (2000), 77–85.
- [Ramasubramanian and Muthuraman 2003] M. K. Ramasubramanian and K. Muthuraman, “A computational mechanics model for the brim forming process in paperboard container manufacturing”, *J. Manuf. Sci. Eng. (ASME)* **125**:3 (2003), 476–483.
- [Ramasubramanian et al. 2005] M. K. Ramasubramanian, R. A. Venditti, C. Ammineni, and V. Mallapragada, “Optical sensor for non-contact measurement of lignin content in high speed moving paper surfaces”, *IEEE Sens. J.* **5**:5 (2005), 1132–1139.
- [Ridgway et al. 1999] P. Ridgway, A. Hunt, M. Quinby-Hunt, and R. Russo, “Laser ultrasonics on moving paper”, *Ultrasonics* **37**:6 (1999), 395–403.
- [Ridgway et al. 2003] P. Ridgway, R. Russo, E. Lafond, C. Habeger, and T. Jackson, “Laser ultrasonic system for on-line measurement of elastic properties of paper”, *J. Pulp Pap. Sci.* **29**:9 (2003), 289–293.
- [Roman and Jeffers 1992] G. Roman and L. Jeffers, “Development of a prototype lignin concentration sensor”, pp. 44 in *DOE/Industry Advanced Sensors Technical Conference*, 1992.
- [Seo et al. 1992] Y. B. Seo, B. Castagnede, and R. E. Mark, “An optimization approach for the determination of in-plane elastic constants of paper”, *TAPPI J.* **75**:11 (1992), 209–214.
- [Singhal et al. 2001] D. K. Singhal, C. Dushyant, and J. Singh, “Fuzzy logic and sorting of waste paper”, *IPPTA J.* **13**:1 (2001), 25–27.
- [Sun et al. 2000] J. S. Sun, K. H. Lee, and H. P. Lee, “Comparison of implicit and explicit finite element methods for dynamic problems”, *J. Mater. Process. Technol.* **105**:1–2 (2000), 110–118.
- [Walter et al. 2000] J. B. Walter, T. L., J. P. Gerhardstein, B. M. Pufahl, C. C. Habeger, M. Lafonde, and P. H. Brodeur, “Fabry–Perot laser ultrasonic elastic anisotropy measurements on a moving paper web”, pp. 247–254 in *Review of progress in quantitative nondestructive evaluation: proceedings of the 26th annual Symposium on Qualitative Nondestructive Evaluation* (Montréal, 1999), vol. 19A, edited by S. Kallsen et al., AIP Conference Proceedings **509**, AIP, Melville, NY, 2000.
- [Zhong 1993] Z. H. Zhong, *Finite element procedures for contact-impact problems*, Oxford University Press, New York, 1993.

Received 6 Jun 2006. Accepted 28 Aug 2006.

M. K. RAMASUBRAMANIAN: [rammk@ncsu.edu](mailto:rammk@ncsu.edu)

*Department of Mechanical and Aerospace Engineering, 1217 Broughton Hall, Campus Box 7910, North Carolina State University, Raleigh, NC 27695-7910, United States*

<http://www.mae.ncsu.edu/homepages/ram/index.html>

RICHARD A. VENDITTI: [richard\\_venditti@ncsu.edu](mailto:richard_venditti@ncsu.edu)

*Department of Wood and Paper Science, North Carolina State University, Raleigh, NC 27695-7910, United States*

<http://www4.ncsu.edu/unity/users/r/richardv/www>

KALYAN C. KATURI: [kckaturi@ncsu.edu](mailto:kckaturi@ncsu.edu)

*Department of Mechanical and Aerospace Engineering, North Carolina State University, Raleigh, NC 27695-7910, United States*

## STATIC ANALYSIS OF MASONRY VAULTS, CONSTITUTIVE MODEL AND NUMERICAL ANALYSIS

MASSIMILIANO LUCCHESI, CRISTINA PADOVANI, GIUSEPPE PASQUINELLI AND NICOLA ZANI

The paper deals with some of the explicit and numerical methods used for modeling the mechanical behavior of masonry vaults and domes. After a brief description of the constitutive equation of masonry-like materials, a numerical method for the structural analysis of masonry vaults is presented. Then, the concept of maximum modulus eccentricity surface for masonry vaults is recalled. Subsequently, the collapse load and the corresponding mechanism are explicitly calculated for two particular cases: a circular plate subjected to a permanent load acting on the lateral surface and a variable vertical load applied to the extrados, and a spherical dome subjected to its own weight and variable point load applied to the keystone. The exact solutions are compared to the numerical results obtained via finite element analysis. Lastly, a study of the dome of the church of Santa Maria Maddalena in Morano Calabro is described. The method has enabled the stress field and the fractures distribution in the dome to be determined and the maximum modulus eccentricity surface to be evaluated.

### 1. Introduction

Many historically and artistically important masonry buildings of the world's architectural heritage are in dire need of maintenance and restoration. In order to optimize such operations in terms of cost-effectiveness, architectural impact and static effectiveness, accurate models of the structural behavior of masonry constructions are invaluable. The ultimate aim of such modeling is to obtain important information, such as the stress field, and to estimate the extent of cracking and its evolution when the structure is subjected to variations in loading conditions.

Although masonry has been used in building for centuries, it is only recently that constitutive models and calculation techniques have been available that enable realistic description of the static behavior of structures made of this heterogeneous material, whose response to tension is fundamentally different from the response to compression. Thus, with the aim of determining the collapse load for masonry structures, many authors have proposed rigid blocks models with different kind of interfaces [Livesley 1978; 1992b], and apply these models to the study of vaults as well [Orduña and Lourenço 2003]. The elastic-plastic model is widely adopted; see, for example, [Lourenço and Rots 1997; Lourenço et al. 1998]. A comparison among different models of this type can be found in [Genna et al. 1998], and an application to the study of vaults is presented in [Theodossopoulos et al. 2002]. Many models have been proposed that, by using homogenization techniques, allow one to take into account the masonry texture, their application is in general limited to the study of panels [Luciano and Sacco 1997; Maier et al. 1992; Trovalusci and Masiani 2005].

---

*Keywords:* masonry vaults, nonlinear elasticity, finite-element method, static analysis, limit analysis.

A constitutive equation widely adopted to model the behavior of masonry materials considers them as nonlinear elastic materials with zero tensile strength and infinite compressive strength [Signorini 1925a; 1925b; Heyman 1966; 1982; Di Pasquale 1992; Del Piero 1989; Panzeca and Polizzotto 1988; Como 1992]. This equation, known as the *masonry-like* or *no-tension* model, can, at least in certain aspects, realistically describe the mechanical behavior of masonry [Bernardeschi et al. 2004; Lucchesi et al. 1996].

Despite the relative simplicity of the constitutive equation of masonry-like materials, explicit solution of equilibrium problems of any practical interest is nonetheless very difficult [Bennati and Padovani 1997; Lucchesi and Zani 2003]. Therefore, in order to study real problems, it is necessary to resort to numerical methods. To this end, suitable numerical techniques have been developed [Lucchesi et al. 1994]. The constitutive models and the numerical method studied have therefore been implemented into the finite element code NOSA [Lucchesi et al. 2000] developed at the Institute of Information Science and Technologies (ISTI) in Pisa. NOSA enables determination of the stress state and the presence of any cracking, as well as the modeling of any potential strengthening and restoration work, such as, for example, the fitting of rods and reinforcement rings.

The code has been successfully applied to the analysis of some buildings of historical and architectural interest, including S. Nicolò's Mother-House in Noto, the Goldoni Theatre in Livorno, the Medici Arsenal in Pisa, the Buti bell tower, and the Pisa church of San Pietro in Vinculis. These studies proved that even though the masonry-like model does not take into account, at least in its original formulation, some features such as material anisotropy, it allows one to conduct realistic analyses of complex structures with large dimensions as frequently encountered in the applications.

In this paper we propose a method for studying masonry vaults. The aim is to provide a computational tool that can be used both to evaluate the safety of masonry monuments, as well as to guide the choice of strengthening operations.

In Section 2, we describe the constitutive equation adopted to study the equilibrium problem of masonry vaults.

Section 3 deals with the numerical methods used for modeling the mechanical behavior of masonry vaults and domes implemented in the NOSA code. A shell element based on the Love-Kirchhoff shell theory [Nagtegaal and Slater 1981] employed for discretizing vaulted structures is described.

In Section 4, the concept of maximum modulus eccentricity surface (m.m.e.s.) is formulated with the aim of allowing concise, effective rendering of the results of finite-element analyses, as well as evaluation of the safety of vaults. In dealing with masonry vaults the m.m.e.s. plays a role analogous to that of line of thrust for arches. In addition, this surface can provide useful information about the possible collapse mechanism when the numerical analysis is performed by progressively increasing the load, until it is no longer possible to determine an admissible equilibrated solution. Then, it allows one to define for the vaults a geometric safety factor similar to that proposed for the arches by Heyman [1982; 1966]. The numerical method proposed in this paper with the consequent calculation of the m.m.e.s., allows one to analyze any kind of vault (masonry spire, fan vault, incomplete dome, et cetera), subjected to any type of static load — a problem that can be hard to solve by classical methods [Heyman 1977]. Moreover, for the different types of load, it is possible to determine the corresponding collapse multiplier. An interactive graphic code has moreover been developed for calculation and visualization of the m.m.e.s.

Section 5 provides some example applications. In particular, we explicitly solve the collapse problem for two distinct structures: a circular plate subjected to a permanent load acting on its lateral surface

and a variable vertical load acting on the extrados, and a spherical dome subjected to its own weight and a variable point load at the keystone. The explicit solutions are compared with the numerical results obtained with the NOSA code. Lastly, we describe a study of the dome of the Church of Santa Maria Maddalena in Morano Calabro. This study has been commissioned by the Office for the Arts and Environment of the Region of Calabria, which is conducting restoration interventions on the church. Finite element analysis has allowed determination of both the stress field and the distribution of fractures within the dome.

## 2. The constitutive equation

To model the mechanical behavior of the material, we use the constitutive equation of *masonry-like* or *no-tension* materials [Heyman 1982; Del Piero 1989; Lucchesi et al. 1994] which are assumed to be nonlinear hyperelastic, with zero tensile strength and infinite compressive strength.

The infinitesimal strain  $\mathbf{E}$  is the sum of an elastic part and a positive semidefinite fracture part: in symbols,  $\mathbf{E} = \mathbf{E}^e + \mathbf{E}^f$ . Moreover, the Cauchy stress  $\mathbf{T}$ , negative semidefinite and orthogonal to  $\mathbf{E}^f$ , depends linearly and isotropically on  $\mathbf{E}^e$  as

$$\mathbf{T} = 2\mu\mathbf{E}^e + \lambda \operatorname{tr}(\mathbf{E}^e)\mathbf{I},$$

with  $\mu > 0$  and  $\lambda \geq 0$  the Lamé moduli of the material. (We assume throughout that Poisson's ratio is nonnegative.)

Consider the vault element with thickness  $h$  shown in Figure 1. Let  $\eta_1$  and  $\eta_2$  be an orthogonal coordinate system, not necessarily the principal one, defined on the mean surface  $\mathfrak{S}$ , with  $\zeta \in [-h/2, h/2]$  as the coordinate in the normal direction  $\mathbf{n}$ . Denote by  $\mathbf{g}_1$  and  $\mathbf{g}_2$  the unit tangent vectors to  $\eta_1$  and  $\eta_2$  axis, respectively.

We assume that for each  $\mathbf{p} \in \mathfrak{S}$  and  $\zeta \in [-h/2, h/2]$ , the stress tensor  $\mathbf{T}$  satisfies the condition

$$\mathbf{T}(\mathbf{p}, \zeta)\mathbf{n}(\mathbf{p}) = \mathbf{0}, \quad (1)$$

$\mathbf{T}$  can be expressed as a nonlinear function of the total strain  $\mathbf{E}$ ,  $\mathbf{T} = \widehat{\mathbf{T}}(\mathbf{E})$ , by using the coaxiality of  $\mathbf{E}$ ,  $\mathbf{T}$  and  $\mathbf{E}^f$ , and taking into account assumption (1); see [Lucchesi et al. 1994]. To this purpose, we indicate with the same symbols the restrictions of  $\mathbf{E}$ ,  $\mathbf{T}$  and  $\mathbf{E}^f$  to the two-dimensional linear space generated by  $\mathbf{g}_1$  and  $\mathbf{g}_2$ .

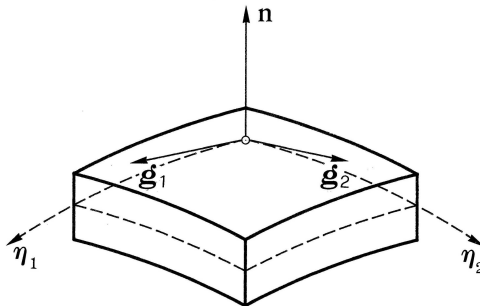


Figure 1. Vault element.



Let  $e_1 \leq e_2$  be the eigenvalues of  $\mathbf{E}$ , and  $\mathbf{q}_1$  and  $\mathbf{q}_2$  the corresponding orthonormal eigenvectors. We then define the tensors

$$\mathbf{O}_{11} = \mathbf{q}_1 \otimes \mathbf{q}_1, \quad \mathbf{O}_{22} = \mathbf{q}_2 \otimes \mathbf{q}_2, \quad \mathbf{O}_{12} = \frac{1}{\sqrt{2}}(\mathbf{q}_1 \otimes \mathbf{q}_2 + \mathbf{q}_2 \otimes \mathbf{q}_1),$$

with  $\otimes$  the tensor product, and we introduce the following subsets of the space  $\text{Sym}$  of all symmetric tensors:

$$\begin{aligned} \mathcal{R}_1 &= \{\mathbf{E} \in \text{Sym} \mid \alpha e_1 + 2(1 + \alpha)e_2 < 0\}, \\ \mathcal{R}_2 &= \{\mathbf{E} \in \text{Sym} \mid e_1 > 0\}, \\ \mathcal{R}_3 &= \{\mathbf{E} \in \text{Sym} \mid \alpha e_1 + 2(1 + \alpha)e_2 > 0, e_1 < 0\}, \end{aligned}$$

where  $\alpha = \lambda/\mu$ . Moreover, we define the interfaces between regions  $\mathcal{R}_1, \mathcal{R}_3$  and  $\mathcal{R}_2, \mathcal{R}_3$  as

$$\mathcal{I}_{13} = \{\mathbf{E} \in \text{Sym} \mid \alpha e_1 + (2 + \alpha)e_2 = 0\}, \quad \mathcal{I}_{23} = \{\mathbf{E} \in \text{Sym} \mid e_1 = 0\}.$$

As proved in [Lucchesi et al. 1994], if  $\mathbf{E} \in \mathcal{R}_1 \cup \mathcal{I}_{13}$ , then

$$\mathbf{E}^f = \mathbf{0}, \quad \mathbf{T} = \frac{2\mu}{2 + \alpha} \{ [2(1 + \alpha)e_1 + \alpha e_2] \mathbf{O}_1 + [\alpha e_1 + 2(1 + \alpha)e_2] \mathbf{O}_2 \};$$

if  $\mathbf{E} \in \mathcal{R}_2 \cup \mathcal{I}_{23}$ , then

$$\mathbf{E}^f = \mathbf{E}, \quad \mathbf{T} = \mathbf{0};$$

and if  $\mathbf{E} \in \mathcal{R}_3$ , then

$$\mathbf{E}^f = \left[ e_2 + \frac{\alpha}{2(1 + \alpha)} e_1 \right] \mathbf{O}_2, \quad \mathbf{T} = \mu \frac{2 + 3\alpha}{1 + \alpha} e_1 \mathbf{O}_1.$$

The derivative  $D_E \widehat{\mathbf{T}}(\mathbf{E})$  of  $\widehat{\mathbf{T}}$  with respect to  $\mathbf{E}$ , which is used for numerical solution of the equilibrium problem (see Section 4), has been explicitly calculated in Lucchesi et al. [1994]. Here, we limit ourselves to recalling its expression. Specifically,

$$D_E \widehat{\mathbf{T}}(\mathbf{E}) = \begin{cases} \mathbb{S}_1 = 2\mu \mathbb{I} + \frac{2\mu\lambda}{2\mu + \lambda} \mathbf{I} \otimes \mathbf{I} & \text{if } \mathbf{E} \in \mathcal{R}_1, \\ \mathbb{S}_2 = \mathbb{O} & \text{if } \mathbf{E} \in \mathcal{R}_2, \\ \mathbb{S}_3 = \frac{\mu(2\mu + 3\lambda)}{\mu + \lambda} (\mathbf{O}_{11} \otimes \mathbf{O}_{11} + \frac{e_1}{e_1 - e_2} \mathbf{O}_{12} \otimes \mathbf{O}_{12}) & \text{if } \mathbf{E} \in \mathcal{R}_3, \end{cases}$$

where  $\mathbb{I}$  and  $\mathbb{O}$  are the identity and the null fourth-order tensors.

On the interfaces  $\mathcal{I}_{13}$  and  $\mathcal{I}_{23}$ , the fourth-order tensor  $D_E \widehat{\mathbf{T}}(\mathbf{E})$  does not exist. Nevertheless, it can be replaced by suitable convex combinations of  $\mathbb{S}_3$  and  $\mathbb{S}_1$ , and  $\mathbb{S}_3$  and  $\mathbb{S}_2$ , respectively.

$\mathbf{E}^f$  is called the fracture strain because if it is nonnull in any region of the structure, we can expect fractures to be present in that region. Knowing  $\mathbf{E}^f$  allows one to get information on the crack distribution. To this end, the vault can be considered constituted by the layers

$$\mathfrak{S}_\zeta = \{\mathbf{p}' \mid \mathbf{p}' = \mathbf{p} + \zeta \mathbf{n}, \mathbf{p} \in \mathfrak{S}, \mathbf{n} = \mathbf{n}(\mathbf{p}), \zeta \in [-h/2, h/2]\}.$$

Assume that in a fixed point  $\mathbf{p}'$  belonging to the layer  $\mathfrak{S}_\zeta$ , one (and only one) eigenvalue of  $\mathbf{T}$  is zero, or equivalently, that  $\mathbf{E}(\mathbf{p}')$  belongs to the region  $\mathcal{R}_3$ . In this case,  $\mathbf{E}^f$  has one eigenvalue that equals zero and one that is positive, and the characteristic direction of the latter coincides with the characteristic

direction corresponding to the null eigenvalue of  $\mathbf{T}$ . Then the fractures can open in this direction. Let  $\varepsilon_{11}^f$ ,  $\varepsilon_{22}^f$  and  $\varepsilon_{12}^f$  be the components of the fracture strain  $\mathbf{E}^f$  with respect to the orthonormal basis  $\{\mathbf{g}_1, \mathbf{g}_2\}$ . The eigenvalues  $e_1^f$  and  $e_2^f$  of  $\mathbf{E}^f$  have the expressions

$$e_1^f = \frac{\varepsilon_{11}^f + \varepsilon_{22}^f - \sqrt{(\varepsilon_{11}^f - \varepsilon_{22}^f)^2 + 4(\varepsilon_{12}^f)^2}}{2}, \quad e_2^f = \frac{\varepsilon_{11}^f + \varepsilon_{22}^f + \sqrt{(\varepsilon_{11}^f - \varepsilon_{22}^f)^2 + 4(\varepsilon_{12}^f)^2}}{2}.$$

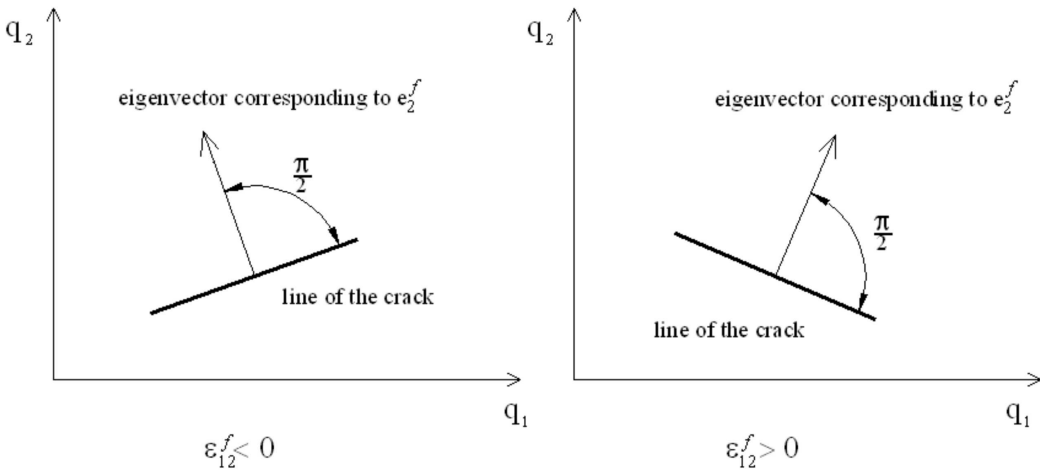
Now assume that  $0 = e_1^f < e_2^f$ . Since we can limit ourselves to the case of  $\varepsilon_{12}^f \neq 0$ , the components  $q_1, q_2$  of the eigenvector corresponding to  $e_2^f$  satisfy the condition

$$q_2 = (\varepsilon_{22}^f / \varepsilon_{12}^f) q_1.$$

Because  $\varepsilon_{22}^f > 0$ , we can conclude that as the sign of  $\varepsilon_{12}^f$  varies, the fractures exhibit the behavior shown in Figure 2. By applying this criterion for each point of the layer  $\mathfrak{S}_\zeta$  it is possible to draw the corresponding fracture curve.

### 3. The numerical method

In this section we describe the finite element procedure implemented into the NOSA code for static analysis of masonry vaults and domes. The equilibrium problem of masonry vaults and domes is solved by using quadrilateral eight-node shell elements based on the Love–Kirchhoff hypothesis [Nagtegaal and Slater 1981]. These elements have three degrees of freedom per corner node. The four mid-side nodes have one degree of freedom, namely the rotation around the side. This rotation is independent of the displacements of the corner nodes, and then the elements are nonconforming.



**Figure 2.** Behavior of the cracks as the sign of  $\varepsilon_{12}^f$  varies.

The displacements within the element and the rotations are given by

$$\mathbf{u} = \sum_{i=1}^4 \varphi_i \mathbf{u}_i, \quad \theta = \sum_{j=5}^9 \psi_j \theta_j,$$

where  $\mathbf{u}_i$  is the displacement vector of the  $i$ -th node,  $\varphi_i$ ,  $i = 1, \dots, 4$  are bilinear shape functions,  $\theta_j$ ,  $j = 5, \dots, 8$  are the rotations of the mid-side nodes,  $\theta_9$  is the rotation of the centroid, and  $\psi_j$  are biquadratic shape functions. We denote by  $\varepsilon_{ij}$  the component of the total strain with respect to the local basis  $\{\mathbf{g}_1, \mathbf{g}_2\}$ . The strain vector  $\varepsilon = (\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{12})^T$  can be expressed as a function of the generalized displacement vector  $\tilde{\mathbf{u}} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \theta_5^s, \theta_6^s, \theta_7^s, \theta_8^s)^T$ , where  $\theta_5^s, \theta_6^s, \theta_7^s$  and  $\theta_8^s$  are the rotations of the mid-side nodes around the side itself, via the matrix  $\mathbf{B}$  containing the derivatives of the shape functions,  $\varepsilon = \mathbf{B}\tilde{\mathbf{u}}$ . In particular, matrix  $\mathbf{B}$  can be written as

$$\mathbf{B} = [\mathbf{B}^m + \zeta \mathbf{B}^c \quad \zeta \mathbf{B}^\theta] \tag{2}$$

where  $\mathbf{B}^m$  accounts for membrane strains, and  $\mathbf{B}^c$  and  $\mathbf{B}^\theta$  account for the curvature changes.

Suitable numerical techniques based on the Newton–Raphson method for solving the nonlinear system obtained by discretizing the structure into finite elements have been developed. To this end, we use the tangent stiffness matrix, whose expression is

$$\mathbf{K}_T = \int_A \int_{-h/2}^{h/2} \mathbf{B}^T \mathbf{D} \mathbf{B} \, d\zeta \, dA, \tag{3}$$

where  $A$  is the area of the element and  $\mathbf{D}$  is the matrix of the components of the derivative  $D_E \widehat{\mathbf{T}}$  of the stress  $\widehat{\mathbf{T}}(\mathbf{E})$  with respect to the strain  $\mathbf{E}$ , given in Section 2. In view of Equation (2), by means of simple calculations, Equation (3) becomes

$$\mathbf{K}_T = \int_A \begin{bmatrix} \mathbf{B}^{mT} & \mathbf{B}^{cT} \\ \mathbf{0} & \mathbf{B}^{\theta T} \end{bmatrix} \begin{bmatrix} \mathbf{D}_0 & \mathbf{D}_1 \\ \mathbf{D}_1 & \mathbf{D}_2 \end{bmatrix} \begin{bmatrix} \mathbf{B}^m & \mathbf{0} \\ \mathbf{B}^f & \mathbf{B}^\theta \end{bmatrix} dA,$$

where

$$\mathbf{D}_i = \int_{-h/2}^{h/2} \zeta^i \mathbf{D} \, d\zeta, \quad i = 0, 1, 2.$$

#### 4. The maximum modulus eccentricity surface

For each point  $\mathbf{p}' = \mathbf{p} + \zeta \mathbf{n}$ , with  $\mathbf{p} = (\eta_1, \eta_2)$  belonging to the mean surface  $\mathfrak{S}$ , we denote by  $\mathbf{g}(\gamma) = (\cos \gamma, \sin \gamma)$  the unit tangent vector to  $\mathfrak{S}$  at  $\mathbf{p}$ , where  $\gamma \in [-\pi/2, \pi/2)$  is the angle formed by  $\mathbf{g}(\gamma)$  with the direction of  $\eta_1$ . We set  $\sigma(\mathbf{p}, \zeta, \gamma) = \mathbf{g}(\gamma) \cdot \mathbf{T}(\mathbf{p}, \zeta) \mathbf{g}(\gamma)$ . Thus, the relations

$$N(\mathbf{p}, \gamma) = \int_{-h/2}^{h/2} \sigma(\mathbf{p}, \zeta, \gamma) \, d\zeta \quad \text{and} \quad M(\mathbf{p}, \gamma) = \int_{-h/2}^{h/2} \sigma(\mathbf{p}, \zeta, \gamma) \zeta \, d\zeta$$

define the normal force and the bending moment per unit length corresponding to  $\mathbf{g}(\gamma)$ , respectively. In view of the fact that  $\mathbf{T}(\mathbf{p}, \zeta)$  is negative semidefinite,  $N(\mathbf{p}, \gamma)$  is nonpositive for every  $\gamma$ . Moreover,

$$|M(\mathbf{p}, \gamma)| = \left| \int_{-h/2}^{h/2} \sigma(\mathbf{p}, \zeta, \gamma) \zeta \, d\zeta \right| \leq \int_{-h/2}^{h/2} |\sigma(\mathbf{p}, \zeta, \gamma) \zeta| \, d\zeta \leq \frac{h}{2} |N(\mathbf{p}, \gamma)|, \quad (4)$$

where the next to last step is justified by the fact that  $\sigma(\mathbf{p}, \zeta, \gamma) \leq 0$  for each  $\gamma$  and  $\zeta$ .

We denote by  $\mathbf{M}(\mathbf{p})$  and  $\mathbf{N}(\mathbf{p})$  the tensors defined by

$$\mathbf{M}(\mathbf{p}) = \int_{-h/2}^{h/2} \mathbf{T}(\mathbf{p}, \zeta) \zeta \, d\zeta, \quad \mathbf{N}(\mathbf{p}) = \int_{-h/2}^{h/2} \mathbf{T}(\mathbf{p}, \zeta) \, d\zeta.$$

Since

$$M(\mathbf{p}, \gamma) = \mathbf{g}(\gamma) \cdot \mathbf{M}(\mathbf{p}) \mathbf{g}(\gamma), \quad N(\mathbf{p}, \gamma) = \mathbf{g}(\gamma) \cdot \mathbf{N}(\mathbf{p}) \mathbf{g}(\gamma), \quad (5)$$

it is a simple matter to show that inequality in Equation (4) is equivalent to the conditions

$$\mathbf{g}(\gamma) \cdot (\mathbf{M}(\mathbf{p}) - \frac{1}{2}h\mathbf{N}(\mathbf{p})) \mathbf{g}(\gamma) \geq 0, \quad \mathbf{g}(\gamma) \cdot (\mathbf{M}(\mathbf{p}) + \frac{1}{2}h\mathbf{N}(\mathbf{p})) \mathbf{g}(\gamma) \leq 0, \quad (6)$$

which in turn, express the positive and negative semidefiniteness of  $\mathbf{M}(\mathbf{p}) - \frac{1}{2}h\mathbf{N}(\mathbf{p})$  and  $\mathbf{M}(\mathbf{p}) + \frac{1}{2}h\mathbf{N}(\mathbf{p})$ . Now let us exclude that the normal force may be zero in all directions at any point of the vault. For each  $\gamma \in [-\pi/2, \pi/2)$ , let us define the corresponding eccentricity

$$e(\mathbf{p}, \gamma) = \frac{M(\mathbf{p}, \gamma)}{N(\mathbf{p}, \gamma)}. \quad (7)$$

The function  $e(\mathbf{p}, \gamma)$  is well-defined and continuous whenever  $N(\mathbf{p}, \gamma) \neq 0$ . Moreover, if there exists  $\gamma_1$  such that  $N(\mathbf{p}, \gamma_1) = 0$ , then  $e(\mathbf{p}, \gamma)$  is a constant function of  $\gamma$ . In fact, we start by observing that if  $N(\mathbf{p}, \gamma_1) = 0$ , then, from the negative semidefiniteness of  $\mathbf{N}(\mathbf{p})$ , the relation  $\mathbf{N}(\mathbf{p}) \mathbf{g}(\gamma_1) = \mathbf{0}$  follows. Moreover, using the preceding relation in Equation (6), we arrive at  $\mathbf{M}(\mathbf{p}) \mathbf{g}(\gamma_1) = \mathbf{0}$ . Then, a vector  $\mathbf{g}(\gamma_2)$  orthogonal to  $\mathbf{g}(\gamma_1)$  is a common eigenvector of  $\mathbf{M}(\mathbf{p})$  and  $\mathbf{N}(\mathbf{p})$ . Denote by  $m$  and  $n < 0$  the corresponding eigenvalues. Setting  $\mathbf{g}(\gamma) = \beta_1 \mathbf{g}(\gamma_1) + \beta_2 \mathbf{g}(\gamma_2)$ , with  $\beta_1^2 + \beta_2^2 = 1$ , we obtain

$$e(\mathbf{p}, \gamma) = \frac{(\beta_1 \mathbf{g}(\gamma_1) + \beta_2 \mathbf{g}(\gamma_2)) \cdot \mathbf{M}(\mathbf{p})(\beta_2 \mathbf{g}(\gamma_2))}{(\beta_1 \mathbf{g}(\gamma_1) + \beta_2 \mathbf{g}(\gamma_2)) \cdot \mathbf{N}(\mathbf{p})(\beta_2 \mathbf{g}(\gamma_2))} = \frac{m}{n}, \quad (8)$$

and thus  $e(\mathbf{p}, \gamma)$  does not depend on  $\gamma$ .

For each point  $\mathbf{p}$  of the mean surface, let  $\gamma_0 \in [-\pi/2, \pi/2)$  be the value of  $\gamma$  (not necessarily unique) for which the function  $|e(\mathbf{p}, \gamma)| = |M(\mathbf{p}, \gamma)/N(\mathbf{p}, \gamma)|$  reaches its maximum value. The quantity

$$\tilde{e}(\mathbf{p}) = \frac{M(\mathbf{p}, \gamma_0)}{N(\mathbf{p}, \gamma_0)}$$

is the maximum modulus eccentricity at point  $\mathbf{p}$ . The surface

$$\mathfrak{M} = \{\mathbf{p}' \mid \mathbf{p}' = \mathbf{p} + \tilde{e}(\mathbf{p}) \mathbf{n}, \mathbf{p} \in \mathfrak{S}, \mathbf{n} = \mathbf{n}(\mathbf{p})\}, \quad (9)$$

obtained by translating points  $\mathbf{p}$  belonging to  $\mathfrak{S}$  by the value  $\tilde{e}(\mathbf{p})$  along the unit normal vector  $\mathbf{n}$  to  $\mathfrak{S}$ , is the maximum modulus eccentricity surface (m.m.e.s.). The only points  $\mathbf{p}$  where  $\mathfrak{M}$  is not defined are those for which two values  $\gamma_0$  and  $\gamma_1$  exist that maximize the function  $|e(\mathbf{p}, \gamma)|$  with  $e(\mathbf{p}, \gamma_0) = -e(\mathbf{p}, \gamma_1)$ .

In the particular case in which the vault’s geometry and loads possess axial symmetry, then the eccentricity can attain its maximum modulus only in the direction of parallels or meridians [Lucchesi et al. 1999]. Note that, in view of inequality in Equation (4), the m.m.e.s. corresponding to a negative semidefinite stress field is entirely contained within the vault.

With the aim of determining the maximum and the minimum of function  $e(\mathbf{p}, \gamma)$ , as  $\gamma$  varies in  $[-\pi/2, \pi/2)$ , we can limit ourselves to the case in which  $\det \mathbf{N}(\mathbf{p}) \neq 0$  (see Equation (8)). In fact, as already pointed out, if  $\det \mathbf{N}(\mathbf{p}) = 0$ , then  $e(\mathbf{p}, \gamma)$  is the constant function  $m/n$ . Due to Equations (5) and (7) the quantity  $e(\mathbf{p}, \gamma)$  is the Rayleigh quotient corresponding to the generalized eigenvalue problem  $\mathbf{M}(\mathbf{p})\mathbf{f} = \omega\mathbf{N}(\mathbf{p})\mathbf{f}$ , whose eigenvalues are<sup>1</sup>

$$\omega_1 = \frac{\text{tr}(\mathbf{N}^{-1}\mathbf{M})(\det \mathbf{N}) - \sqrt{(\text{tr}(\mathbf{N}^{-1}\mathbf{M}))^2(\det \mathbf{N})^2 - 4 \det \mathbf{N} \det \mathbf{M}}}{2 \det \mathbf{N}},$$

$$\omega_2 = \frac{\text{tr}(\mathbf{N}^{-1}\mathbf{M})(\det \mathbf{N}) + \sqrt{(\text{tr}(\mathbf{N}^{-1}\mathbf{M}))^2(\det \mathbf{N})^2 - 4 \det \mathbf{N} \det \mathbf{M}}}{2 \det \mathbf{N}},$$

with eigenvectors  $\mathbf{f}_1$  and  $\mathbf{f}_2$ . It is well known that the Rayleigh quotient  $e(\mathbf{p}, \gamma)$  belongs to the interval  $[\omega_1, \omega_2]$ ; thus, we can conclude that the minimum and the maximum of  $e(\mathbf{p}, \gamma)$  are  $\omega_1$  and  $\omega_2$ . Finally,

$$\tilde{e}(\mathbf{p}) = \begin{cases} \omega_1 & \text{if } \det \mathbf{N}(\mathbf{p}) \neq 0, |\omega_1| \geq |\omega_2|, \\ \omega_2 & \text{if } \det \mathbf{N}(\mathbf{p}) \neq 0, |\omega_1| \leq |\omega_2|, \\ m/n & \text{if } \det \mathbf{N}(\mathbf{p}) = 0. \end{cases}$$

If the m.m.e.s. is tangent to the extrados or intrados along a path, such a path can be considered the site of cylindrical hinges. The corresponding rotational axis coincides with the direction orthogonal to the direction for which the absolute value of the eccentricity is maximum. Therefore, under the hypothesis that the masonry has infinite compressive strength and sliding failure cannot occur, as the load increases, vault collapse occurs when the m.m.e.s. is tangent to the intrados or extrados along paths, in such a way as to determine a hinge distribution sufficient to render the vault a kinematically undetermined structure. It should be pointed out that in the model presented herein, collapse of a masonry vault is due to the formation of hinges on its extrados and intrados. However, other collapse mechanisms can be considered, for example by removing the “non-sliding” assumption [Livesley 1978] or assuming a bounded compressive strength for the masonry [Livesley 1992a].

Knowing  $\tilde{e}$  at each point of the mean surface allows for defining some “geometrical” factors of safety. By following the suggestion of Heyman [1982], for each  $\mathbf{p} \in \mathfrak{S}$  we put

$$\bar{h}(\mathbf{p}) = \begin{cases} \omega_2 - \omega_1 & \text{if } \det \mathbf{N}(\mathbf{p}) \neq 0, \\ 2|m/n| & \text{if } \det \mathbf{N}(\mathbf{p}) = 0. \end{cases}$$

<sup>1</sup>The expression for these same quantities presented in [Lucchesi et al. 1999] contained errors which are corrected here.

We determine  $v = \int_{\mathfrak{S}} \bar{h}(\mathbf{p}) \, da$ , which is the volume of the smallest vault able to contain the m.m.e.s. Then we determine the volume  $V$  of the vault and set

$$\varphi_v = \frac{V - v}{V} = 1 - \frac{v}{V}. \quad (10)$$

Alternatively, we can put

$$\bar{e} = \sup_{\mathbf{p} \in \mathfrak{S}} \frac{2 |\tilde{e}(\mathbf{p})|}{h(\mathbf{p})}, \quad (11)$$

where  $h(\mathbf{p})$  is the thickness of the vault, and consider

$$\varphi_s = 1 - \bar{e}. \quad (12)$$

The coefficient  $\varphi_s$  can turn out to be very conservative. In fact, in view of the definition in Equation (11), the presence of a single point at which the m.m.e.s. comes close to the boundary of the vault causes  $\varphi_s$  to approach zero.

## 5. Examples

This section is aimed at assessing the effectiveness of the method proposed in the paper and implemented in the finite element code NOSA. Three example applications are addressed. The first two applications deal with the limit analysis of simple masonry structures. We determine the collapse load and the collapse mechanism for a circular plate subjected to a lateral pressure and an increasing vertical force. Then, we study a spherical vault subjected to its own weight and a point load applied at the keystone. For both cases, we determine the exact solution and then, for the sake of comparison, conduct a finite element analysis by increasing the variable load until it is no longer possible to find an admissible (that is, negative semidefinite at each point of the structure) equilibrated stress field. The reader is referred to [Del Piero 1998] for a general treatment of the limit analysis of masonry-like materials and to [Lucchesi et al. 1999] for the definitions and the results concerning the limit analysis of masonry vaults. Last, we describe the study conducted via the NOSA code on the dome and the drum of the Church of Santa Maria Maddalena in Morano Calabro. We analyzed the behavior of the structure subjected to its own weight as well as in the presence of settlements of the base of the drum.

**5.1. The circular plate.** In the cylindrical reference system  $\{\mathbf{O}, r, \theta, z\}$ , consider the circular plate  $\mathcal{P}$  with radius  $R$  and thickness  $h$ , constrained along the lateral surface  $\{r = R\}$  in such a way that rotation and vertical displacements are prevented (Figure 3). The plate is subjected to a horizontal pressure  $\mathbf{c}_0$ , uniformly distributed on the lateral surface, and a vertical pressure  $\lambda \mathbf{c}_1$ , uniformly distributed on the circle of the extrados  $0 \leq r \leq b$ , with  $b \leq R$ . For the sake of simplicity, we have ignored weight, but the corresponding generalization presents no difficulties.

We aim to determine the collapse multiplier  $\lambda_c$ . To this end, for  $\lambda > 0$  we determine an admissible internal forces field equilibrated with the load  $\mathbf{c}_0 + \lambda \mathbf{c}_1$  and then find the multiplier  $\lambda = \lambda_c$ , such that there is a collapse mechanism corresponding to the internal forces equilibrated with  $\mathbf{c}_0 + \lambda_c \mathbf{c}_1$  [Lucchesi et al. 1999]. Let  $Q$ ,  $N_r$ ,  $N_\theta$ ,  $M_r$  and  $M_\theta$  be respectively the shear, radial and circumferential normal forces, and the radial and circumferential bending moment per unit length. They must satisfy the equilibrium

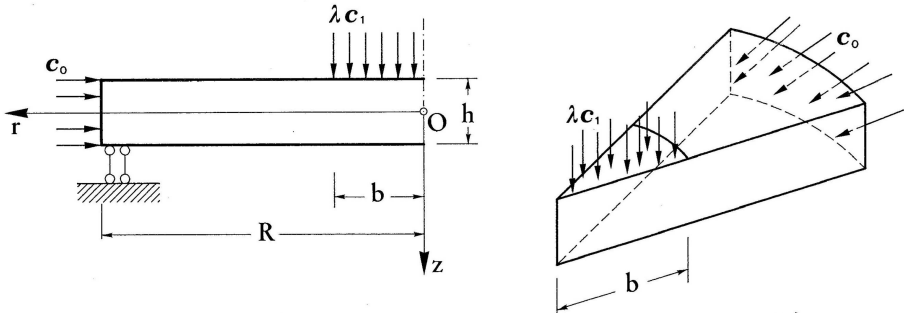


Figure 3. The circular plate.

equations [Timoshenko and Woinowsky-Krieger 1987]

$$\frac{dQ}{dr} + \frac{1}{r}Q = \lambda c_1, \tag{13}$$

$$\frac{dN_r}{dr} + \frac{1}{r}N_r - \frac{1}{r}N_\theta = 0, \tag{14}$$

$$\frac{dM_r}{dr} + \frac{1}{r}M_r - \frac{1}{r}M_\theta + Q = 0, \tag{15}$$

where

$$c_1(r) = \begin{cases} 1 & \text{if } 0 \leq r \leq b, \\ 0 & \text{if } b < r \leq R, \end{cases} \tag{16}$$

and the boundary conditions

$$Q(R) = \frac{\lambda b^2}{2R}, \quad N_r(R) = -c_0 h, \quad c_0 = |c_0|. \tag{17}$$

Since the internal forces to be determined must be collapse internal forces, by setting  $e_r = M_r/N_r$ ,  $e_\theta = M_\theta/N_\theta$ , we require that they satisfy the further boundary conditions

$$e_r(0) = -\frac{h}{2}, \quad e_r(R) = \frac{h}{2}, \tag{18}$$

and that the equality

$$e_\theta(r) = -\frac{h}{2} \tag{19}$$

holds for  $0 \leq r \leq R$ . Finally, in order for internal forces to be admissible, we require that the inequalities

$$N_r(r) \leq 0, \quad N_\theta(r) \leq 0, \quad -\frac{h}{2} \leq e_r(r) \leq \frac{h}{2} \tag{20}$$

be satisfied for  $0 \leq r \leq R$ .

From Equations (13), (16) and (17) we can immediately deduce the shear:

$$Q(r) = \begin{cases} \lambda r/2 & \text{if } 0 \leq r \leq b, \\ \lambda b^2/(2r) & \text{if } b < r \leq R. \end{cases}$$

Equations (14), (15) and (19) are insufficient to determine the remaining four internal forces. A further condition is needed. This condition is obtained by assuming the circumferential normal force to be constant,  $N_\theta(r) = \kappa$  for  $0 \leq r \leq R$ , with  $\kappa \leq 0$ . From Equation (19) we then deduce  $M_\theta(r) = -\frac{1}{2}h\kappa$  for  $0 \leq r \leq R$ . By integrating Equation (14) with the help of Equation (17)<sub>2</sub>, we obtain

$$N_r(r) = -\frac{R}{r}(c_0h + \kappa) + \kappa,$$

and from the admissibility conditions in (20)<sub>1</sub> and (20)<sub>2</sub>, we obtain  $-c_0h \leq \kappa \leq 0$ . Lastly, by integrating Equation (15) with boundary condition (18)<sub>1</sub>, we arrive at the expression of  $M_r$  which depends on  $\lambda$ . The boundary condition (18)<sub>2</sub> is satisfied for  $\lambda = \lambda_c$  with

$$\lambda_c = \frac{6c_0Rh^2}{b^2(3R - 2b)}. \quad (21)$$

It, in turn, allows for determining the final expression for  $M_r$ ,

$$M_r(r) = \begin{cases} \frac{1}{2r} \left( hR(c_0h + \kappa) - h\kappa r - \frac{2c_0h^2R}{b^2(3R - 2b)}r^3 \right) & \text{if } 0 < r \leq b, \\ \frac{1}{2r} \left( hR(c_0h + \kappa) - h\kappa r - \frac{2c_0h^2b^2R}{(3R - 2b)^3}(3r - 2b) \right) & \text{if } b \leq r \leq R. \end{cases}$$

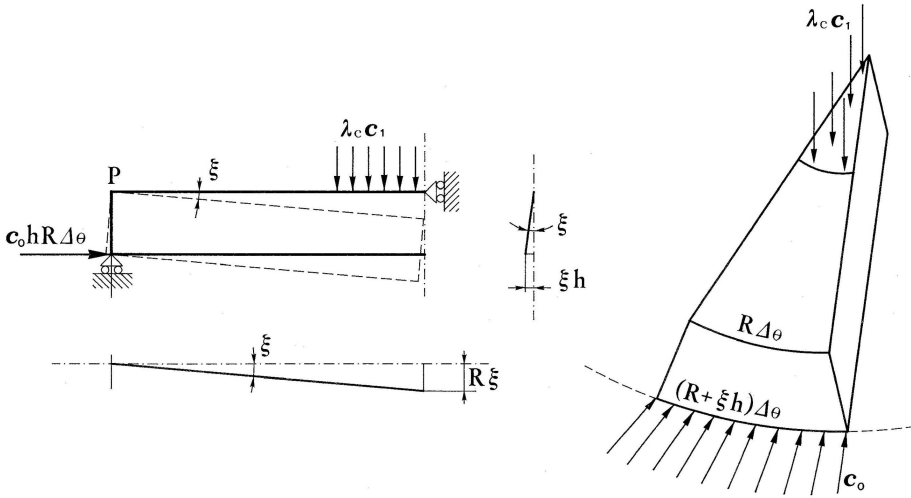
It is a simple matter to verify that the admissibility condition in Equation (20)<sub>3</sub> is satisfied.

In view of Equations (18) and (19), the m.m.e.s. corresponding to the determined internal forces is a plane region coinciding with the extrados for  $0 \leq r < R$  and undefined for  $r = R$ . In fact, for  $0 \leq r \leq R$  the eccentricity  $e_\theta$  is  $-\frac{1}{2}h$ , and the absolute value of the radial eccentricity  $e_r$  is less than  $\frac{1}{2}h$ . For  $r = R$ , we have  $e_r = \frac{1}{2}h$ .

The value of  $\lambda_c$  given by Equation (21) is the collapse multiplier. In fact, the displacement field suggested by the shape of the m.m.e.s. and shown in Figure 4 is a collapse mechanism. Each section  $\theta = \text{constant}$  rotates around the absolute center of rotation  $P$  of the extrados. This rotation is possible because, in view of Equation (18)<sub>2</sub>, the sliding clamp shown in Figure 3 has released as depicted in Figure 4, left. From Equation (19) it follows that each radius of the extrados is the axis of a cylindrical hinge. These hinges allow the surface of the extrados to form the lateral surface of a cone; see Figure 4, right.

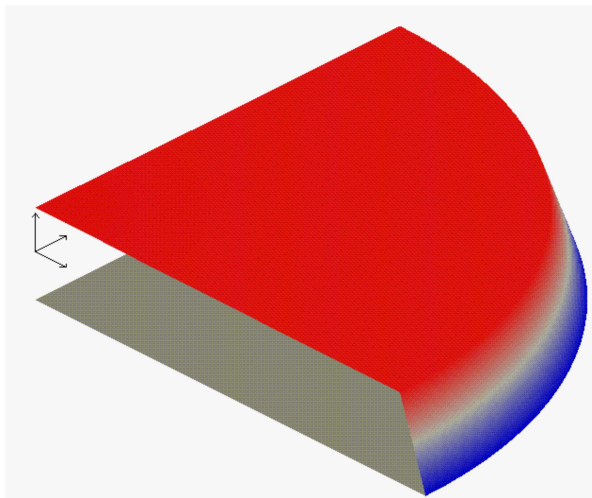
One quarter of the plate, with  $R = 0.5$  m and  $h = 0.01$  m has been discretized and analyzed via the finite element code NOSA. After applying the horizontal pressure  $c_0 = 5 \cdot 10^4$  Pa, the vertical pressure  $\lambda$ , uniformly distributed throughout the extrados ( $b = R$ ), is progressively increased by assigning successive load increments until convergence can be reached. By using the stress field corresponding to the last load increment, the maximum modulus eccentricity surface has been determined. (In Figure 5 the thickness of the plate has been magnified ten times in order to better illustrate the surface.)





**Figure 4.** The plate collapse mechanism.

For the numerical example presented here, relation (21) with  $b = R$  yields  $\lambda_c = 6c_0h^2/R^2 = 120$  Pa. Naturally, the value of  $\lambda_s$  of the pressure beyond which it is impossible to obtain a numerical solution to the plate equilibrium problem depends on the refinement of the mesh and the entity of the load increments. More precisely, we have verified that, for an assigned load increment, using a less refined mesh yields a higher value of  $\lambda_s$ . For example, with 900 and 4000 elements, we get  $\lambda_s = 125.8$  Pa and  $\lambda_s = 120$  Pa, respectively.



**Figure 5.** The m.m.e.s. at the instant of collapse.

**5.2. The spherical dome.** In the spherical reference system  $\{O, r, \theta, \varphi\}$ , consider the spherical vault  $\mathcal{D}$ , with mean radius  $R$  and thickness  $h$ . The vault is clamped at the springings and subjected to its own weight  $\mathbf{c}_0$  and a point load  $\lambda \mathbf{c}_1$  applied at the keystone. Our aim is to determine the value of the collapse multiplier  $\lambda_c$ .

We begin by writing the equilibrium equations for the dome under the assumption that the circumferential normal force and bending moment vanish. Denoting  $q$  as the weight of the vault per unit area,  $N$  and  $M$  as the meridional normal force and bending moment per unit length, respectively, and  $Q$  as the shear per unit length, we can write (see [Timoshenko and Woinowsky-Krieger 1987])

$$\frac{d\bar{N}(\varphi)}{d\varphi} - \bar{Q}(\varphi) + Rq \sin^2 \varphi = 0, \quad (22)$$

$$\frac{d\bar{Q}(\varphi)}{d\varphi} + \bar{N}(\varphi) + Rq \sin \varphi \cos \varphi = 0, \quad (23)$$

$$\frac{d\bar{M}(\varphi)}{d\varphi} - R\bar{Q}(\varphi) = 0, \quad (24)$$

where  $\bar{N}(\varphi) = N(\varphi) \sin \varphi$ ,  $\bar{Q}(\varphi) = Q(\varphi) \sin \varphi$ ,  $\bar{M}(\varphi) = M(\varphi) \sin \varphi$ . Moreover, by imposing vertical force equilibrium of the spherical bowl with amplitude  $\varphi$  we obtain

$$2\pi R\bar{N}(\varphi) + 2\pi R\bar{Q}(\varphi) \cos \varphi + T(\varphi) = 0, \quad (25)$$

where  $T(\varphi) = 2\pi R^2 q(1 - \cos \varphi) + \lambda$  is the total load acting on the spherical bowl. By taking  $\bar{Q}(\varphi)$  from Equation (25) and substituting into Equation (22), we get

$$\frac{d\bar{N}(\varphi)}{d\varphi} + \bar{N}(\varphi) \tan \varphi + \frac{Rq(p - \cos^3 \varphi)}{\cos \varphi} = 0, \quad (26)$$

where we have put

$$p = 1 + \frac{\lambda}{2\pi R^2 q}, \quad p \geq 1. \quad (27)$$

The solution to Equation (26) in the interval  $[0, \pi/2]$  is  $\bar{N}(\varphi) = \bar{N}(0) \cos \varphi + Rq \sin \varphi (\cos \varphi - p)$ . On the other hand, by imposing horizontal force equilibrium of a slice of amplitude  $\Delta\theta$ , we obtain  $R\Delta\theta[\bar{N}(0) + \bar{Q}(\pi/2)] = 0$ , from which we deduce  $\bar{N}(\varphi) = -\bar{Q}(\pi/2) \cos \varphi + Rq \sin \varphi (\cos \varphi - p)$ . From the above relation, with the help of Equation (25) we can easily derive the expression for shear,  $\bar{Q}(\varphi) = -\bar{Q}(\pi/2) - Rq \cos \varphi (p - \cos \varphi)$ . As  $\bar{Q}(\varphi)$  is now known, we can integrate Equation (24), obtaining

$$\bar{M}(\varphi) = \frac{\pi}{2} (\bar{M} - R\bar{Q} \cos \varphi) + qR^2 \left( \frac{2\varphi - \pi + \sin 2\varphi}{4} + p(1 - \sin \varphi) \right). \quad (28)$$

Since we are interested in determining collapse internal forces, we assume that two hinges form at the extrados, one at the crown and the other at the springing,

$$e(0) = \frac{\bar{M}(0)}{\bar{N}(0)} = -\frac{h}{2}, \quad e(\pi/2) = \frac{\bar{M}(\pi/2)}{\bar{N}(\pi/2)} = -\frac{h}{2}. \quad (29)$$

Using these relations and setting  $t = h/(2R)$ , we obtain

$$\bar{Q}(\pi/2) = Rq \left( p - \frac{\pi}{4(1+t)} \right), \quad \bar{M}(\pi/2) = R^2 q p t, \tag{30}$$

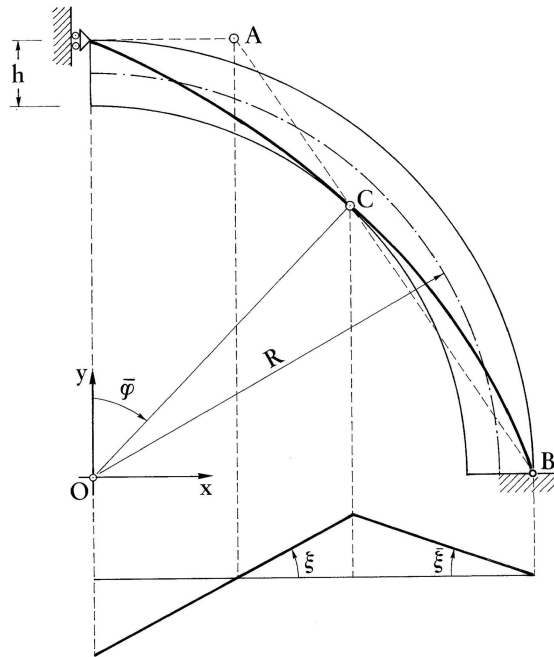
which, in turn, enables us to arrive at the expressions for the normal force and bending moment. In fact, from Equations (28)–(30), we obtain

$$\begin{aligned} \bar{N}(\varphi) &= -Rq \left( \left( p - \frac{\pi}{4(1+t)} \right) \cos \varphi + \sin \varphi (p - \cos \varphi) \right), \\ \bar{M}(\varphi) &= R^2 q \left( - \left( p - \frac{\pi}{4(1+t)} \right) \cos \varphi + \frac{2\varphi - \pi + \sin 2\varphi}{4} + p(1+t - \sin \varphi) \right). \end{aligned}$$

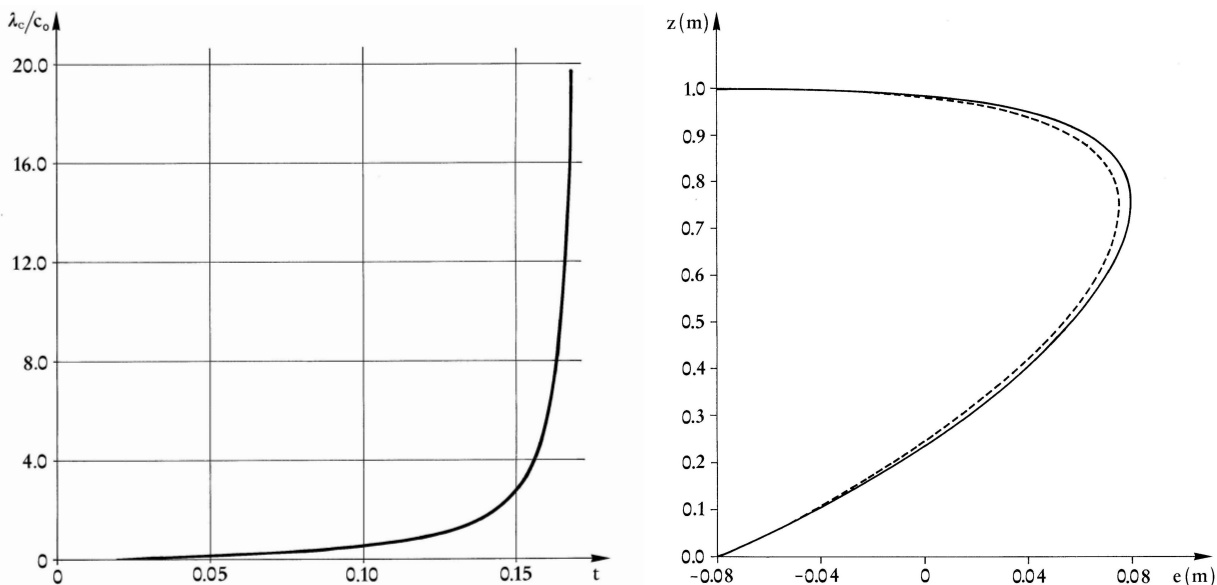
Therefore, the eccentricity, which depends on angle  $\varphi$  and load parameter  $p$ , is

$$e(\varphi, p) = -R \frac{p(1+t - \sin \varphi) - \left( p - \frac{\pi}{4(1+t)} \right) \cos \varphi + \frac{1}{4}(2\varphi - \pi + \sin 2\varphi)}{\left( p - \frac{\pi}{4(1+t)} \right) \cos \varphi + \sin \varphi (p - \cos \varphi)}. \tag{31}$$

Since the circumferential normal force is zero, the eccentricity is independent of the direction in the tangent plane. Thus, the maximum modulus eccentricity surface is the surface whose distance from the mean surface of the vault, measured along the radial direction, is  $e(\varphi, p)$ . Simple calculations reveal that  $\partial e/\partial p > 0$  for each  $\varphi \in [0, \pi/2)$  and  $p \geq 1$ . Therefore, there is a unique  $p_c$ , such that a unique  $\bar{\varphi} \in (0, \pi/2)$  exists, for which the maximum modulus eccentricity surface meets the vault intrados (see Figure 6),  $e(\bar{\varphi}, p) = \frac{1}{2}h$ . To verify that  $\lambda_c = 2\pi R^2 q (p_c - 1)$  is the collapse multiplier (compare Equation



**Figure 6.** The m.m.e.s. corresponding to collapse multiplier and associated collapse mechanism.



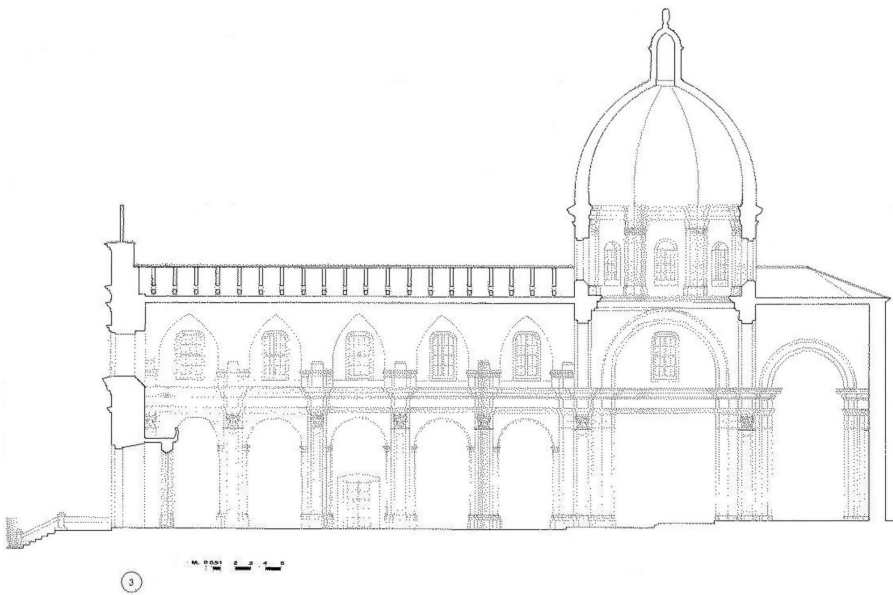
**Figure 7.** Left:  $\lambda_c/c_0$  as a function of  $t$ . Right: Eccentricity  $e$  as a function of  $z = R \cos \varphi$  at collapse.

(27)), it is sufficient to determine a corresponding mechanism. Figure 6 suggests considering the vault to be a kinematically undetermined structure made up of slices of infinitesimal amplitude, each in turn constituted by two bodies  $\{0 \leq \varphi \leq \bar{\varphi}\}$  and  $\{\bar{\varphi} \leq \varphi \leq \pi/2\}$  linked by the hinge at point  $C$ . The absolute centers of rotation of the first and second body are the points  $A$  and  $B$ , respectively. Figure 7, left, shows the behavior of the ratio  $\lambda_c/c_0$  as a function of  $t = h/(2R)$ . In particular,  $h \approx 0.0377 R$  is the minimum thickness necessary for the vault to be in equilibrium with its own weight. Moreover, as  $t$  approaches the value of about 0.17, the collapse load grows infinitely.

One quarter-vault with  $R = 1$  m,  $h = 0.16$  m and specific weight  $20000 \text{ N/m}^3$  has been discretized with 3200 shell elements and analyzed with the NOSA code. The load multiplier  $\lambda$  is increased incrementally up to the value  $\lambda_s$  beyond which it is no longer possible to obtain convergence. For the case at hand, the collapse multiplier is  $\lambda_c = 6230 \text{ N}$  and the collapse load calculated by NOSA is  $\lambda_s = 6200 \text{ N}$ . Figure 7, right, shows a comparison of the eccentricity derived from the stress field calculated numerically at the instant of collapse (dotted line) with the eccentricity obtained from Equation (31) (continuous line).

**5.3. The dome of the Church of S. Maria Maddalena in Morano Calabro.** Historical information on the church and the operations to which it has been subjected over the centuries is quite scarce. It was probably built during the Middle Ages on top of an earlier chapel situated outside the town. It was enlarged and restored in the sixteenth and eighteenth centuries.

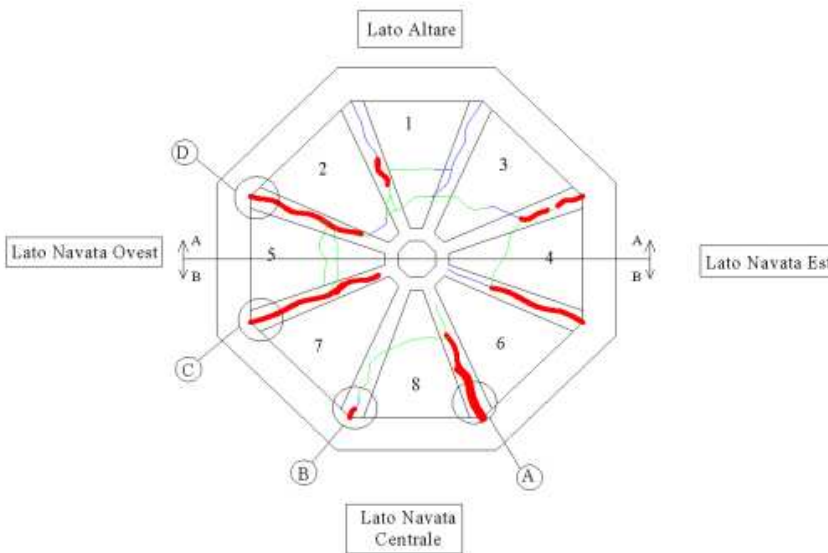
As for the structural layout, the church is a load bearing masonry building in a Latin cross plan, with a central nave and two side aisles. The church is about 50 m long, 24 m wide and 33 m high; see Figure 8. The presbytery zone is delimited by an octagonal cross-section drum supporting a dome and a lantern. The drum has a diameter of about 11 m, height of 6.10 m and constant thickness of 110 cm. The 8 m-high dome has a diameter equal to that of the drum and thickness varying from 110 cm at the springing to



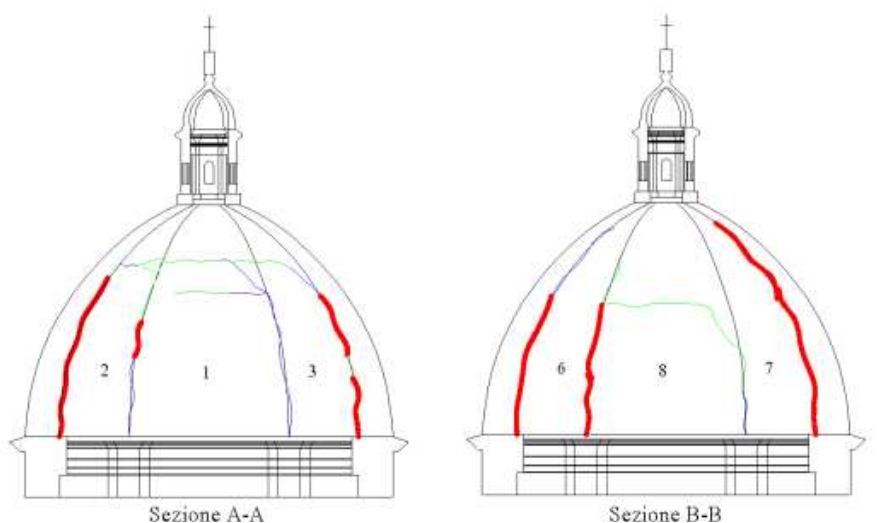
**Figure 8.** Longitudinal section of the church by the nave.

70 cm at its top. Here, there is a lantern with diameter of 2 m, height of 3.4 m and thickness of about 40 cm.

A survey of the cracking and its distribution in both the drum and dome was carried out in October 2004. In the dome, the crack distribution is known only for the intrados, because the extrados is completely covered with majolica tiles. The dome shows widespread cracking, as well as evident signs of water seepage. The intrados bears two sets of fractures, different in both dimensions and behavior. The first



**Figure 9.** Distribution of cracks in the intrados of the dome.



**Figure 10.** Sections A-A and B-B of the dome: crack distribution in the intrados.

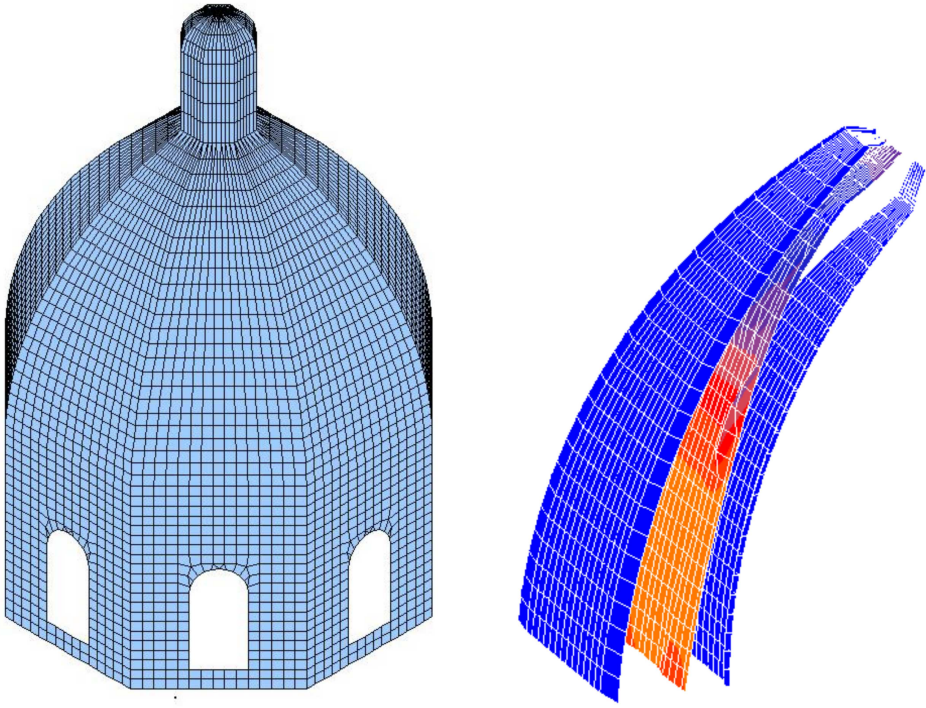
set, which contains the larger cracks, involves the ribs connecting the webs of the dome. These cracks extend from the top of the drum up to well beyond the mid-height of the dome (Figures 9 and 10). The width of the cracks is maximum in correspondence to the dome base and decreases towards its top, where the cracks close and do not affect the last part of the ribs. The second set of cracks is horizontal and located in the webs, in the upper half of the dome. They extend from one rib to the next, inscribing a closed circular curve which involves nearly the entire dome. The width of these openings is much smaller than the first, vertical set of cracks. Rather, the drum is in relatively good condition; the survey revealed the presence of a few cracks with moderate extension. In particular, the cracks are located at the top of the drum and are continuations of those on the dome. They extend onto the drum for a short stretch and stop above the windows.

The crack survey of the structure is shown in Figures 9 and 10, where different colors are used to represent fractures with different width: red, blue and green represent cracks of major, medium and minor width, respectively.

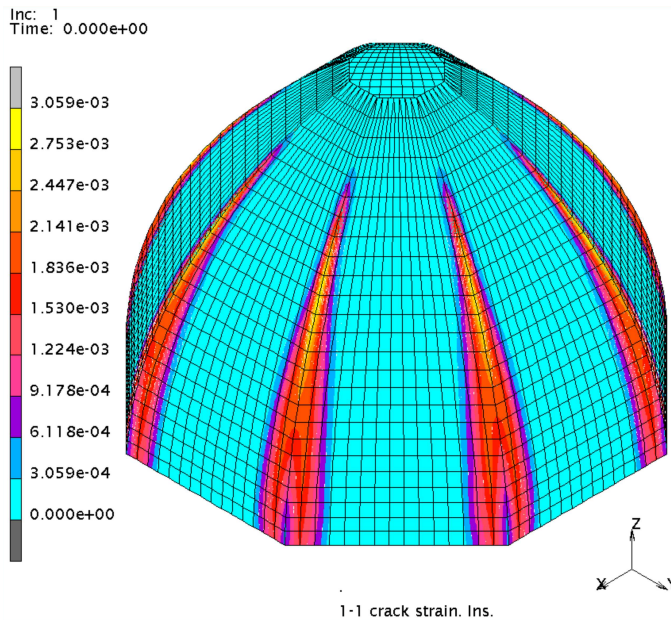
Finite-element structural analyses of the dome as well as the drum have been performed with NOSA. The structure was discretized with 5496 shell elements (Figure 11, left). At first, the base of the drum was considered to be clamped and the effects of the other parts of the church thereby neglected. Further analyses were then carried out with the aim of assessing the effects of vertical settlements of the drum. Since no experimental data are available for the masonry material of the structure, we assumed  $\gamma = 20000 \text{ N/m}^3$  for the specific weight,  $E = 3.0 \times 10^9 \text{ Pa}$  for the Young's modulus and  $\nu = 0.1$  for the Poisson's ratio.

In the following we describe the results of the analysis of the structure subjected to its own weight. Maximum horizontal displacement values of about 1.9 mm are reached at the base of the dome. The maximum vertical displacements, of about 1.6 mm, occur in the lantern and on the top of the dome.

A global analysis of the entire structure allows for concluding that compressive stresses are less than  $13 \times 10^5 \text{ Pa}$ , a value considered acceptable. Figure 11, right, illustrates the m.m.e.s. for a web together



**Figure 11.** Left: Finite element discretization of the dome and the drum. Right: Maximum modulus eccentricity surface in a web.



**Figure 12.** Component  $\varepsilon_{11}^f$  of the fracture strain at the intrados of the dome subjected to its own weight.

with the extrados and the intrados (in blue). As for the intrados, it is in agreement with the crack distribution shown in Figures 9 and 10. Moreover, as for the factors defined in Equations (10) and (12), we have  $\varphi_v = 0.71$  and  $\varphi_s = 0.17$ .

With the aim of determining the direction of any eventual cracking, the components of the fracture strain in correspondence to both the intrados and the extrados were analyzed. Let

$$\varepsilon_{11}^f(\mathbf{p}, \zeta) = \mathbf{g}_1 \mathbf{E}^f(\mathbf{p}, \zeta) \mathbf{g}_1 \quad \text{and} \quad \varepsilon_{22}^f(\mathbf{p}, \zeta) = \mathbf{g}_2 \mathbf{E}^f(\mathbf{p}, \zeta) \mathbf{g}_2$$

be the components of the fracture strain along parallels and meridians, respectively, and

$$\varepsilon_{12}^f(\mathbf{p}, \zeta) = \mathbf{g}_1 \mathbf{E}^f(\mathbf{p}, \zeta) \mathbf{g}_2$$

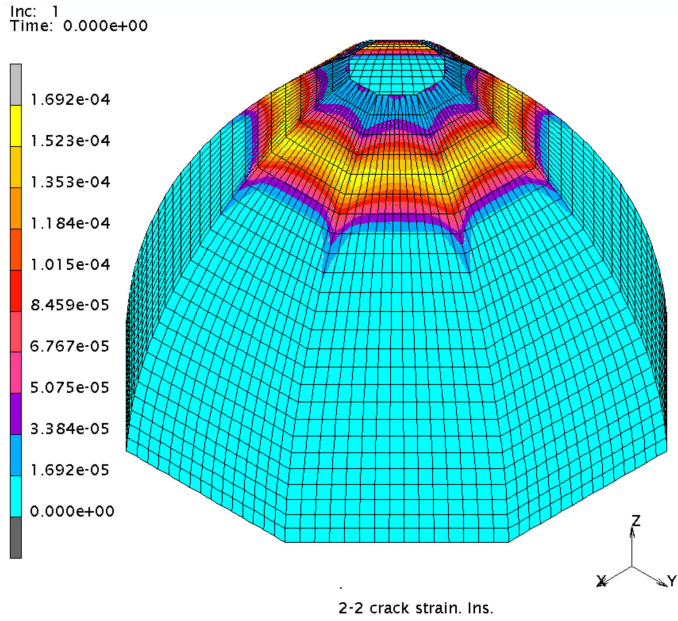
be the shear component. The behavior of  $\varepsilon_{11}^f$ ,  $\varepsilon_{22}^f$ ,  $\varepsilon_{12}^f$  is shown in Figures 12–14 for the dome. For  $\varepsilon_{11}^f$ , the highest values are reached in the ribs starting at the top of the drum and continuing up to a certain height. In these regions  $\varepsilon_{22}^f$  and  $\varepsilon_{12}^f$  are equal to zero, which leads to the conclusion that there are fractures along the meridians, concentrated in the ribs. These fractures start at the base of the dome, reach a certain height and then close. On the lower part, they continue downward, past the dome onto the upper portion of the drum. The  $\varepsilon_{22}^f$  component reaches its maximum value in the upper half of the dome, at about 6.90 m from the springing, in the center of the webs, where  $\varepsilon_{11}^f$  and  $\varepsilon_{12}^f$  are negligible. We deduce that in these regions the fractures have the same directions as parallels of latitude. Moreover, near the ribs,  $\varepsilon_{11}^f$ ,  $\varepsilon_{22}^f$  and  $\varepsilon_{12}^f$  are nonzero and the parallels are no longer principal directions for the fracture strain. By applying the criterion described in Section 2 (see Figure 2) we deduce that here the fractures join the ribs, arranged so as to form an approximate upside down U shape (see Figure 9).

As for the extrados, the numerical results suggest the presence of fracturing along the ribs, higher up than that on the intrados. Over the base of the dome, at about 2 m from the springing, there is another series of fractures similar to those on the intrados: they follow the parallel in the central part of each web and join the ribs, with a similar upside down U trend.

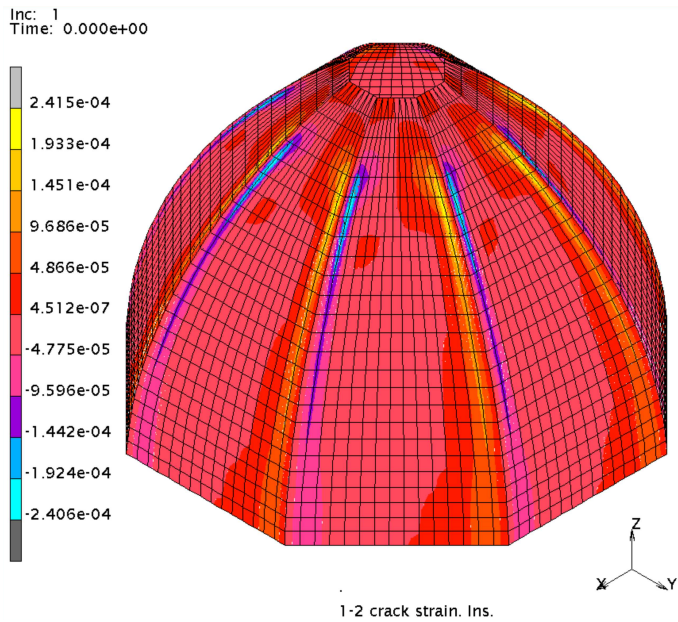
The agreement between the crack distribution actually observed in the intrados and the numerical results obtained confirms the effectiveness of the finite-element code NOSA and represents substantial grounds for considering the results obtained for the dome extrados as realistic and reliable as well.

Further numerical analyses were conducted with the aim of assessing the behavior of the structure in the presence of vertical settlements of the base of the drum. These investigations can be useful to explain the origin of the asymmetries of the crack distribution seen in Figures 9 and 10. To model sinking of the foundations, three adjacent sides of the base of the drum are subjected to a vertical piecewise linear displacement. In particular, we imposed a constant displacement of 1 cm to the central side, and a linear displacement varying from 1 to 0 to the other two sides. A displacement of the same kind has been assigned to the other three opposite sides, with a maximum displacement of 0.5 cm. In this case compressive stresses are less than  $18 \times 10^5$  Pa. As for the factors of safety, the worst condition occurs in webs 3 and 7, where we have  $\varphi_v = 0.46$  and  $\varphi_s = 0$ . Figures 15–17 show the components  $\varepsilon_{11}^f$ ,  $\varepsilon_{22}^f$  and  $\varepsilon_{12}^f$  in the intrados of the dome. We point out that due to the assigned vertical displacement, no new fractures form in addition to those corresponding to the weight, and they rearrange in an asymmetric way, in agreement with Figures 9 and 10.

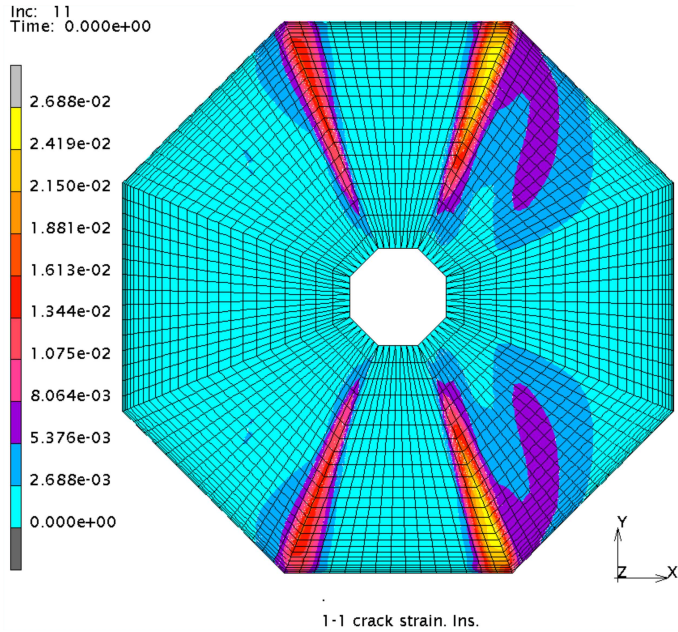




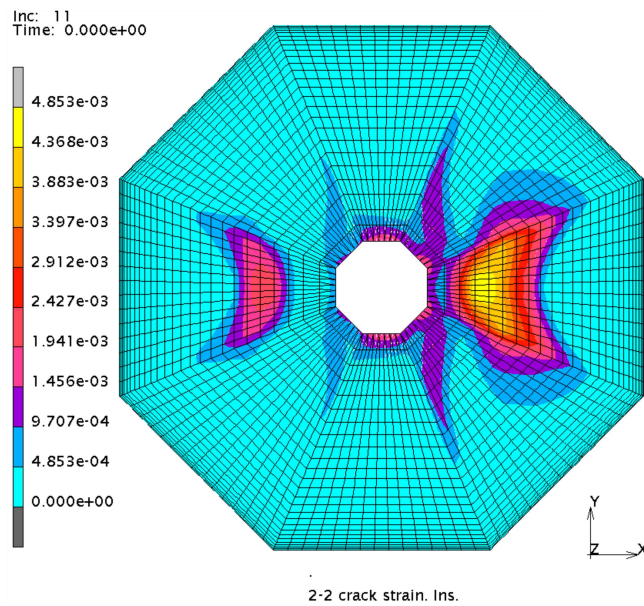
**Figure 13.** Component  $\varepsilon_{22}^f$  of the fracture strain at the intrados of the dome subjected to its own weight.



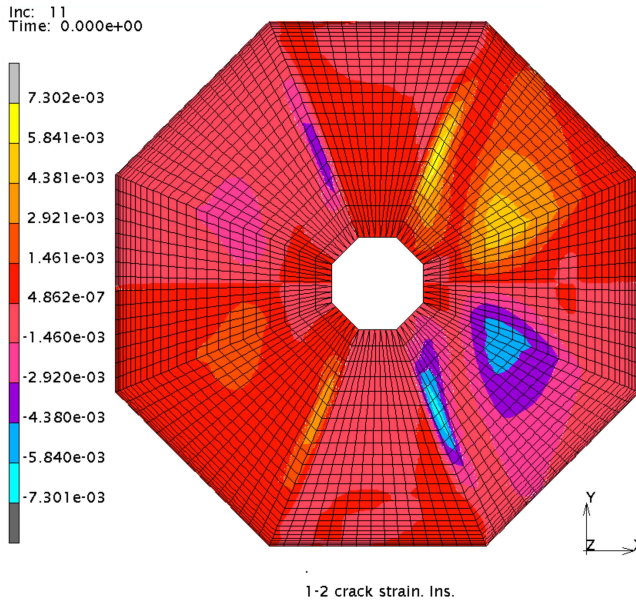
**Figure 14.** Component  $\varepsilon_{12}^f$  of the fracture strain in the intrados of the dome subjected to its own weight.



**Figure 15.** Component  $\epsilon_{11}^f$  of the fracture strain at the intrados of the dome subjected to its own weight and a vertical displacement of the drum.



**Figure 16.** Component  $\epsilon_{22}^f$  of the fracture strain at the intrados of the dome subjected to its own weight and a vertical displacement of the drum.



**Figure 17.** Component  $\varepsilon_{12}^f$  of the fracture strain at the intrados of the dome subjected to its own weight and a vertical displacement of the drum.

## 6. Conclusions

The methods proposed in this paper have proved to be useful in studying the static behavior of any kind of masonry vault subjected to very general load conditions, both for conducting collapse analyses and solving equilibrium problems. In particular, the study on the dome presented in Section 5.3 has shown that using suitable constitutive equations and appropriate numerical techniques to solve nonlinear problems enables determining the stress field, assessing vault safety on the basis of the derived maximum modulus eccentricity surface and, lastly, predicting the distribution and depth of eventual fracturing with good accuracy. Moreover, the ability to model variations in boundary conditions allows one to assess the behavior of a structure in the presence of settlement of the foundations, by determining the influence of such settlement on the stress field and the crack distribution.

## References

- [Bennati and Padovani 1997] S. Bennati and C. Padovani, “Some non-linear elastic solutions for masonry solids”, *Mech. Struct. Mach.* **25**:2 (1997), 243–266.
- [Bernardeschi et al. 2004] K. Bernardeschi, C. Padovani, and G. Pasquinelli, “Numerical modelling of the structural behaviour of Buti’s bell tower”, *J. Cult. Herit.* **5**:4 (2004), 371–378.
- [Como 1992] M. Como, “Equilibrium and collapse analysis of masonry bodies”, *Meccanica (Milano)* **27**:3 (1992), 185–194.
- [Del Piero 1989] G. Del Piero, “Constitutive equation and compatibility of the external loads for linear elastic masonry-like materials”, *Meccanica (Milano)* **24**:3 (1989), 150–162.
- [Del Piero 1998] G. Del Piero, “Limit analysis and no-tension materials”, *Int. J. Plast.* **14**:1–3 (1998), 259–271.
- [Di Pasquale 1992] S. Di Pasquale, “New trends in the analysis of masonry structures”, *Meccanica (Milano)* **27**:3 (1992), 173–184.

- [Genna et al. 1998] F. Genna, M. Di Pasqua, M. Veroli, and P. Ronca, “Numerical analysis of old masonry buildings: a comparison among constitutive models”, *Eng. Struct.* **20**:1–2 (1998), 37–53.
- [Heyman 1966] J. Heyman, “The stone skeleton”, *Int. J. Solids Struct.* **2**:2 (1966), 249–256.
- [Heyman 1977] J. Heyman, *Equilibrium of shell structures*, Clarendon Press, Oxford, 1977.
- [Heyman 1982] J. Heyman, *The masonry arch*, Wiley, New York, 1982.
- [Livesley 1978] R. K. Livesley, “Limit analysis of structures formed from rigid blocks”, *Int. J. Numer. Methods Eng.* **12**:12 (1978), 1853–1871.
- [Livesley 1992a] R. K. Livesley, “The collapse analysis of masonry arch bridges”, pp. 261–274 in *Proceedings of the 4th Conference on Applied Solid Mechanics*, Elsevier, 1992.
- [Livesley 1992b] R. K. Livesley, “A computational model for the limit analysis of three-dimensional masonry structures”, *Meccanica (Milano)* **27**:3 (1992), 161–172.
- [Lourenço and Rots 1997] P. B. Lourenço and J. G. Rots, “Multisurface interface model for analysis of masonry structures”, *J. Eng. Mech. (ASCE)* **123**:7 (1997), 660–668.
- [Lourenço et al. 1998] P. B. Lourenço, J. G. Rots, and J. Blaauwendraad, “Continuum model for masonry: parameter estimation and validation”, *J. Struct. Eng. (ASCE)* **124**:6 (1998), 642–652.
- [Lucchesi and Zani 2003] M. Lucchesi and N. Zani, “Some explicit solutions to plane equilibrium problem for no-tension bodies”, *Struct. Eng. Mech.* **16**:3 (2003), 295–516.
- [Lucchesi et al. 1994] M. Lucchesi, C. Padovani, and A. Pagni, “A numerical method for solving equilibrium problems of masonry-like solids”, *Meccanica (Milano)* **29**:2 (1994), 175–193.
- [Lucchesi et al. 1996] M. Lucchesi, C. Padovani, and N. Zani, “Masonry-like solids with bounded compressive strength”, *Int. J. Solids Struct.* **33**:14 (1996), 1961–1994.
- [Lucchesi et al. 1999] M. Lucchesi, C. Padovani, G. Pasquinelli, and N. Zani, “The maximum modulus eccentricity surface for masonry vaults and limit analysis”, *Math. Mech. Solids* **4**:1 (1999), 71–87.
- [Lucchesi et al. 2000] M. Lucchesi, C. Padovani, A. Pagni, G. Pasquinelli, and N. Zani, “COMES-NOSA: A finite element code for non-linear structural analysis”, Technical report B4-2000-003, CNUCE, 2000.
- [Luciano and Sacco 1997] R. Luciano and E. Sacco, “Homogenization technique and damage model for old masonry material”, *Int. J. Solids Struct.* **34**:24 (1997), 3191–3208.
- [Maier et al. 1992] G. Maier, A. Nappi, and E. Papa, “Constitutive laws for engineering materials”, pp. 427–432 in *Damage models for masonry as a composite material: a numerical and experimental analysis*, edited by C. S. Desai, ASME Press, New York, 1992.
- [Nagtegaal and Slater 1981] J. C. Nagtegaal and I. G. Slater, “A simple nonconforming thin shell element based on discrete Kirchhoff theory”, pp. 167–192 in *Nonlinear finite element analysis of plates and shells*, edited by T. J. R. Hughes et al., AMD **48**, ASME, New York, 1981. Zbl 0511.00033
- [Orduña and Lourenço 2003] A. Orduña and P. B. Lourenço, “Cap model for limit analysis and strengthening of masonry structures”, *J. Struct. Eng. (ASCE)* **129**:10 (2003), 1367–1375.
- [Panzeca and Polizzotto 1988] T. Panzeca and C. Polizzotto, “Constitutive equations for no-tension materials”, *Meccanica (Milano)* **23**:2 (1988), 88–93.
- [Signorini 1925a] A. Signorini, “Un teorema di esistenza e unicità nello studio dei materiali poco resistenti a trazione”, *Rend. Accad. Naz. Lincei* **2** (1925), 401–406.
- [Signorini 1925b] A. Signorini, “Sulla pressoflessione delle murature”, *Rend. Accad. Naz. Lincei* **2** (1925), 484–489.
- [Theodossopoulos et al. 2002] D. Theodossopoulos, B. P. Sinha, A. S. Usmani, and A. S. J. Macdonald, “Assessment of the structural response of masonry cross vaults”, *Strain* **38**:3 (2002), 119–127.
- [Timoshenko and Woinowsky-Krieger 1987] S. P. Timoshenko and S. Woinowsky-Krieger, *Theory of plates and shells*, 2nd ed., McGraw-Hill, New York, 1987.
- [Trovalusci and Masiani 2005] P. Trovalusci and R. Masiani, “A multifield model for blocky materials based on multiscale description”, *Int. J. Solids Struct.* **42**:21–22 (2005), 5778–5794.

Received 8 May 2006. Accepted 18 Sep 2006.

MASSIMILIANO LUCCHESI: [massimiliano.lucchesi@unifi.it](mailto:massimiliano.lucchesi@unifi.it)

*Dipartimento di Costruzioni, Università di Firenze, Piazza Brunelleschi 6, 50121 Firenze, Italia*

CRISTINA PADOVANI: [cristina.padovani@isti.cnr.it](mailto:cristina.padovani@isti.cnr.it)

*Istituto di Scienza e Tecnologie dell'Informazione Alessandro Faedo, ISTI-CNR, Via G. Moruzzi 1, 56124 Pisa, Italia*

GIUSEPPE PASQUINELLI: [giuseppe.pasquinelli@isti.cnr.it](mailto:giuseppe.pasquinelli@isti.cnr.it)

*Istituto di Scienza e Tecnologie dell'Informazione Alessandro Faedo, ISTI-CNR, Via G. Moruzzi 1, 56124 Pisa, Italia*

NICOLA ZANI: [nicola.zani@unifi.it](mailto:nicola.zani@unifi.it)

*Dipartimento di Costruzioni, Università di Firenze, Piazza Brunelleschi 6, 50121 Firenze, Italia*

## WORST CASE PLASTIC LIMIT ANALYSIS OF TRUSSES UNDER UNCERTAIN LOADS VIA MIXED 0-1 PROGRAMMING

YOSHIHIRO KANNO AND IZURU TAKEWAKI

The paper presents a global optimization method to compute the minimum limit load factor of trusses subjected to unknown but bounded loads. We assume that the external forces consist of a part proportional to a load factor and a part that is uncertain around its nominal value. The worst-case limit load factor is introduced as the smallest limit load factor realized with some uncertain parameters. In order to detect the worst case, we have to find the global optimal solution of a nonconvex optimization problem, which is the major difficulty of the worst-case limit analysis. By reformulating the worst-case determination problem as a mixed 0-1 programming problem, we propose a global optimization algorithm as a combination of a branch-and-bound method based on the linear programming relaxations and a cutting plane method based on the disjunctive or lift-and-project cuts. The worst-case limit loads, as well as the corresponding critical loading patterns, are computed to demonstrate that our method converges to the global optimal solutions successfully.

### 1. Introduction

In designing civil, mechanical and aerospace structures, plastic limit analysis has been used widely for decades as a means of estimating the ultimate strength of structures. While dead and live loads are uncertain around their nominal values, the disturbance load is applied proportionally with a load factor. This paper discusses a global optimization technique for computing the smallest limit load factor of truss structures, where the applied dead and live loads are imprecisely known.

Limit analysis still receives much attention by numerous researchers with regard to algorithms [Muraidhar and Jagannatha Rao 1997; Andersen et al. 1998; Cocchetti and Maier 2003; Krabbenhoft and Damkilde 2003] and issues relevant to the finite element method [Tin-Loi and Ngo 2003; Lyamin et al. 2005]. Based on the probabilistic uncertainty models of structural systems, various approaches to stochastic limit analysis have also been proposed [Lloyd Smith et al. 1990; Rocho and Sonnenberg 2003; Staat and Heitzer 2003; Marti and Stoeckel 2004]. In the framework of probabilistic uncertainties, reliability-based structural design methods have been investigated extensively [Zang et al. 2005; Kharmanda et al. 2004].

Besides these probabilistic uncertainty models, nonprobabilistic uncertainty models have also been developed, where a mechanical system is assumed to contain uncertain parameters which are unknown but bounded. [Ben-Haim and Elishakoff 1990] developed the well known convex model approach, with which [Ganzerli and Pantelides 1999] proposed a robust truss optimization method. Interval linear algebra has been well developed for uncertain linear equations [Alefeld and Mayer 2000], and has been employed in structural analyses with uncertainties [Chen et al. 2002]. In contrast to probabilistic models, these

---

*Keywords:* data uncertainty, limit analysis, integer programming, cutting plane, branch-and-bound, global optimization.

nonprobabilistic uncertainty models require only upper bounds on the magnitude of uncertain parameters, and engineers need not estimate the probabilistic density distributions of uncertain parameters.

[Elishakoff et al. 1994] proposed a structural optimization scheme under unknown-but-bounded uncertainty by using antioptimization. The bilevel optimization problems were formulated and solved numerically for robust structural design against the worst case [Craig et al. 2003; Cheng et al. 2004]. [Gu et al. 2000] proposed an estimation method for the worst case of propagated uncertainty in a multidisciplinary system. A unified methodology which is a robust counterpart of various convex optimization problems was developed by [Ben-Tal and Nemirovski 2002], and was applied to robust compliance minimization of trusses [Ben-Tal and Nemirovski 1997]. The authors proposed methods for robustness analysis and robust optimization of structures [Takewaki and Ben-Haim 2005; Kanno and Takewaki 2006a; Kanno and Takewaki 2006b] based on the info-gap uncertainty model [Ben-Haim 2001].

A serious difficulty in worst-case detection arises when the worst case is defined as an optimal solution of a nonconvex optimization problem in terms of the uncertain parameters. Conventional methods for linear worst-case analysis, for example the convex model of [Ben-Haim and Elishakoff 1990], can be applied only to cases in which sufficiently small variation of the uncertain parameters is allowed, or in which the structural response considered is represented as a linear function of the uncertain parameters. In these cases, the worst case can be detected by solving a convex optimization problem.

Unfortunately, in many practical situations, the variation of uncertain parameters is not small and we are interested in nonlinear responses of structures. Then the worst case is defined through a nonconvex optimization problem. In general, the conventional nonlinear programming approach converges to a local optimal solution of that problem. However, a local minimum solution, which is not globally optimal, does not correspond to the worst case. Obviously, the worst case corresponds to a global optimal solution. Thus, we have to find a global optimal solution of the nonconvex problem and guarantee that the solution obtained is globally optimal, which prevent us from using the conventional nonlinear programming algorithms.

In this paper, we aim at developing a global optimization method for worst-case detection. We consider the limit load factor of a truss structure subjected to uncertain loads. The external forces applied to a truss are supposed to consist of a constant part and a part proportional to a load factor, where the former part cannot be known precisely, but is assumed to be bounded. The worst-case limit load factor is defined as the minimum value among all the possible limit load factors realized by some uncertain parameters belonging to the given closed set.

We define the worst-case limit load factor by using a nonconvex optimization problem, which can be rewritten as a mixed 0-1 programming problem. Based on a linear programming (LP) relaxation, we propose a simple branch-and-bound algorithm to obtain a global optimal solution of the mixed 0-1 programming problem. To strengthen the LP relaxation, we generate some cutting planes at the root node of the branch-and-bound tree. This approach is called the cut-and-branch method [Cordier et al. 1999]. We formulate an LP problem to generate the deepest disjunctive cut. By adding generated cuts to LP relaxation problems, we drastically reduce the number of LP problems to be solved in the branch-and-bound method.

The solution obtained by using the cut-and-branch method is a global optimal solution of the worst-case determination problem, that is, it is assured that there exists no uncertain parameter with which the limit load factor becomes smaller than the obtained optimal value. Through the numerical experiments in

Section 6 we show that the limit load factor can be reduced greatly from its nominal value, by nontrivial combination of uncertain external forces. We also show that the critical load yielding the worst-case limit load factor cannot be detected easily by generating a large sample of loading scenarios. The upper bound of the worst-case limit load factor obtained from such a sample is shown to be too optimistic.

Recently, there has been renewed interest in cutting planes, or cuts, that are valid linear inequalities of a mixed integer programming problem; see for example the review paper [Marchand et al. 2002]. In particular, the branch-and-cut method [Balas et al. 1996; Cordier et al. 1999], that is, a branch-and-bound method with cuts added, is considered as the most successful approach to solving the mixed integer program. Among various types of cuts, a disjunctive cut (or lift-and-project cut) is defined as a linear inequality selected among inequalities valid for a disjunctive programming relaxation of the mixed 0-1 program [Balas et al. 1996; Ceria and Soares 1997; Balas and Perregaard 2002]. We utilize disjunctive cuts to strengthen the LP relaxation problems that are solved at nodes of the branch-and-bound tree.

This paper is organized as follows. In Section 2, we prepare the LP problems for conventional limit analysis and define the notation used in this paper. In Section 3 we introduce the notion of uncertain limit analysis by defining the info-gap model for uncertainty of external load and the worst-case limit load factor. In Section 4, we present the mixed 0-1 programming formulation for the uncertain limit analysis, and propose a branch-and-bound method as the solution. In Section 5 we propose an LP problem that generates the disjunctive cutting plane, to strengthen the LP relaxation problems solved in the branch-and-bound tree. In Section 6 we present numerical experiments for various trusses, made using our cut-and-branch method. Finally, in Section 7 we draw conclusions.

## 2. Notation and preliminary results

**2.1. Notation.** In this paper, we assume all vectors to be column vectors. For an  $n$ -tuple  $p_{m+1}, \dots, p_{m+n}$ , we let  $(p_i | i = m + 1, \dots, m + n)$  and  $\{p_i | i = m + 1, \dots, m + n\}$ , respectively, denote the  $n$ -dimensional vector  $(p_{m+1}, \dots, p_{m+n})^\top$  and the set consisting of  $p_{m+1}, \dots, p_{m+n}$ . The vector  $(p_i | i = 1, \dots, n) \in \mathbb{R}^n$  is often simplified as  $(p_i) \in \mathbb{R}^n$ . The  $\ell^1$ ,  $\ell^2$  (or standard Euclidean), and  $\ell^\infty$  norms of the vector  $\mathbf{p} = (p_i) \in \mathbb{R}^n$ , denoted by  $\|\mathbf{p}\|_1$ ,  $\|\mathbf{p}\|_2$ , and  $\|\mathbf{p}\|_\infty$ , respectively, are defined as

$$\begin{aligned} \|\mathbf{p}\|_1 &= \sum_{i=1}^n |p_i|, \\ \|\mathbf{p}\|_2 &= (\mathbf{p}^\top \mathbf{p})^{1/2}, \\ \|\mathbf{p}\|_\infty &= \max_{i \in \{1, \dots, n\}} |p_i|. \end{aligned}$$

For vectors  $\mathbf{p} = (p_i) \in \mathbb{R}^n$  and  $\mathbf{q} = (q_i) \in \mathbb{R}^n$ , we write  $\mathbf{p} \geq \mathbf{0}$  and  $\mathbf{p} \geq \mathbf{q}$ , respectively, if  $p_i \geq 0$ ,  $i = 1, \dots, n$  and  $\mathbf{p} - \mathbf{q} \geq \mathbf{0}$ . The  $(m + n)$ -dimensional column vector  $(\mathbf{p}^\top, \mathbf{q}^\top)^\top$  is often written simply as  $(\mathbf{p}, \mathbf{q})$ . Moreover,  $(\mathbf{p}, \mathbf{q})_i$  denotes the  $i$ -th component of the vector  $(\mathbf{p}^\top, \mathbf{q}^\top)^\top$ .



We define the vectors  $\mathbf{1} \in \mathbb{R}^n$  and  $\mathbf{e}^j \in \mathbb{R}^n$ ,  $j = 1, \dots, n$ , as

$$\mathbf{1} = (1, \dots, 1)^\top,$$

$$\mathbf{e}^j = (e_i^j \mid i = 1, \dots, n), \quad e_i^j = \begin{cases} e_i^j = 0, & \text{for } i \neq j, \\ e_j^j = 1, \end{cases}$$

that is,  $\mathbf{e}^j$  is the  $j$ -th column vector of the identity matrix. We define  $\mathbb{R}_+^n \subset \mathbb{R}^n$  as

$$\mathbb{R}_+^n = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0}\}.$$

For two sets  $\mathcal{A} \subseteq \mathbb{R}^m$  and  $\mathcal{B} \subseteq \mathbb{R}^n$ , the Cartesian product is defined as  $\mathcal{A} \times \mathcal{B} = \{(\mathbf{a}^\top, \mathbf{b}^\top)^\top \in \mathbb{R}^{m+n} \mid \mathbf{a} \in \mathcal{A}, \mathbf{b} \in \mathcal{B}\}$ . In particular, we write  $\mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$ .

The convex hull of  $\mathcal{A}$ , that is the smallest convex set that contains  $\mathcal{A}$  is denoted 'conv  $\mathcal{A}$ .' The closure of  $\mathcal{A}$ , that is smallest closed set that contains  $\mathcal{A}$ , is denoted 'cl  $\mathcal{A}$ .' The cardinality of the set  $\mathcal{A}$  is denoted  $|\mathcal{A}|$ . The empty set is denoted as  $\emptyset$ .

**2.2. Basic problem for plastic limit analysis.** Consider an elastic and perfectly-plastic truss in two- or three-dimensional space. Small rotations and small strains are assumed. Let  $\mathbf{f} \in \mathbb{R}^{n^d}$  denote the vector of the external forces, where  $n^d$  denotes the number of degrees of freedom of displacements. Where  $n^m$  is the number of members, the vector of member axial forces is  $\mathbf{q} = (q_i) \in \mathbb{R}^{n^m}$ . The system of equilibrium equations in terms of  $\mathbf{f}$  and  $\mathbf{q}$  can be written

$$\mathbf{B}\mathbf{q} = \mathbf{f}, \quad (1)$$

where  $\mathbf{B} \in \mathbb{R}^{n^d \times n^m}$  is a constant matrix.

Let  $\mathbf{u} \in \mathbb{R}^{n^d}$  and  $c_i$  denote the vector of nodal displacements and the corresponding elongation of the  $i$ -th member, respectively. We often write  $\mathbf{c} = (c_i) \in \mathbb{R}^{n^m}$ . The  $i$ -th column vector of  $\mathbf{B}$  is denoted by  $\mathbf{b}_i \in \mathbb{R}^{n^d}$ ,  $i = 1, \dots, n^m$ . The compatibility relation between  $\mathbf{u}$  and  $c_i$  can be written as

$$c_i = \mathbf{b}_i^\top \mathbf{u}, \quad i = 1, \dots, n^m. \quad (2)$$

The external load  $\mathbf{f}$  consists of a constant part  $\mathbf{f}_D$  and a proportionally increasing part  $\lambda \mathbf{f}_R$ , that is,

$$\mathbf{f} = \mathbf{f}_D + \lambda \mathbf{f}_R. \quad (3)$$

Here,  $\lambda \mathbf{f}_R$  is defined by the monotonically increasing load parameter  $\lambda \in \mathbb{R}$  and the constant reference load  $\mathbb{R}^{n^d} \ni \mathbf{f}_R \neq \mathbf{0}$ . In civil engineering,  $\mathbf{f}_D$  consists of the dead load, live load, etc., while  $\lambda \mathbf{f}_R$  consists of live or disturbance load caused by earthquakes, winds, and the like. For the sake of simplicity, in this paper  $\mathbf{f}_D$  is simply called *dead load* and  $\mathbf{f}_R$  is called *reference disturbance load*.

Let  $\sigma_i^y > 0$  and  $-\sigma_i^y$  denote the yield stresses of the  $i$ -th member in tension and in compression, respectively. We assume for simplicity that the yield stresses in tension and compression share the common absolute value. The member cross-sectional area is denoted by  $a_i > 0$ . The absolute value of the admissible axial force can be expressed as

$$q_i^y = a_i \sigma_i^y, \quad i = 1, \dots, n^m.$$

Then, the yield functions can be written as

$$|q_i| - q_i^y \leq 0, \quad i = 1, \dots, n^m. \quad (4)$$

From the static or lower-bound principle [Hodge 1959], and by using (1), (3), and (4), the limit load factor is obtained by solving the following LP problem

$$\max \{ \lambda : \mathbf{B}\mathbf{q} = \mathbf{f}_D + \lambda \mathbf{f}_R, \quad |q_i| - q_i^y \leq 0, \quad i = 1, \dots, n^m \}, \quad (5)$$

where the variables are  $\lambda$  and  $\mathbf{q}$ .

### 3. Uncertain limit analysis

In this section, we introduce the uncertainty model of the external load, and rigorously define the worst-case limit load factor.

**3.1. Uncertainty model.** In this paper, we suppose that only  $\mathbf{f}_D$  in Equation (3) possesses uncertainty, that is, it cannot be known precisely. The model of the uncertainty of  $\mathbf{f}_D$  is motivated by a nonprobabilistic information-gap model [Ben-Haim 2001].

Let  $\tilde{\mathbf{f}}_D \in \mathbb{R}^{n^d}$  denote the nominal value (or the best estimate) of  $\mathbf{f}_D$ . We describe the uncertainty of  $\mathbf{f}_D$  in terms of the  $m$ -dimensional vector  $\boldsymbol{\zeta} \in \mathbb{R}^m$ , which is considered to be unknown but bounded. Suppose that  $\mathbf{f}_D$  depends on  $\boldsymbol{\zeta}$  affinely as

$$\mathbf{f}_D = \tilde{\mathbf{f}}_D + \mathbf{T}\boldsymbol{\zeta}, \quad (6)$$

where  $\mathbf{T} \in \mathbb{R}^{n^d \times m}$  is a constant matrix satisfying the following assumption:

**Assumption 3.1.** The matrix  $\mathbf{T}$  in Equation (6) satisfies the following conditions:

- (i)  $\{\mathbf{T}^\top \mathbf{u} \mid \mathbf{u} \in \mathbb{R}^{n^d}\} = \mathbb{R}^m$ ;
- (ii)  $\mathbf{f}_R^\top \mathbf{T}\boldsymbol{\zeta} = 0$  for any  $\boldsymbol{\zeta} \in \mathbb{R}^m$ .

Assumption 3.1 (ii) implies that the reference disturbance load  $\mathbf{f}_R$  does not have uncertainty.

For a given parameter  $\alpha \in \mathbb{R}_+$ , the uncertain set  $\mathcal{L}(\alpha) \subset \mathbb{R}^m$  is defined as

$$\mathcal{L}(\alpha) = \{ \boldsymbol{\zeta} \in \mathbb{R}^m \mid \alpha \geq \|\boldsymbol{\zeta}\|_\infty \}, \quad (7)$$

where the uncertain parameters vector  $\boldsymbol{\zeta}$  is assumed to be running through  $\mathcal{L}(\alpha)$ , that is,

$$\boldsymbol{\zeta} \in \mathcal{L}(\alpha). \quad (8)$$

From (6), (7), and (8) it follows that the uncertain  $\mathbf{f}_D$  satisfies

$$\mathbf{f}_D \in \mathcal{F}_D(\alpha) := \left\{ \mathbf{f} \in \mathbb{R}^{n^d} \mid \mathbf{f} = \tilde{\mathbf{f}}_D + \mathbf{T}\boldsymbol{\zeta}, \quad \alpha \geq \|\boldsymbol{\zeta}\|_\infty \right\}. \quad (9)$$

Roughly speaking,  $\mathbf{f}_D$  moves around the center-point  $\tilde{\mathbf{f}}_D$ . The greater the value of  $\alpha$ , the greater the range of possible variation of  $\mathbf{f}_D$ . In the context of the info-gap uncertainty model [Ben-Haim 2001],  $\alpha$  is called the uncertainty parameter. Throughout, we suppose that the bound  $\alpha$  on the uncertain variation given in Equation (9) is a constant. Note that the uncertain set  $\mathcal{F}_D(\alpha)$  is bounded for any  $\alpha \in \mathbb{R}_+$ .

Moreover,  $\mathcal{F}_D(\alpha)$  satisfies the two basic axioms for the info-gap model: nesting, in which  $0 \leq \alpha_1 < \alpha_2$  implies  $\mathcal{F}_D(\alpha_1) \subset \mathcal{F}_D(\alpha_2)$ , and contraction, in which  $\mathcal{F}_D(0)$  is the singleton set  $\{\tilde{\mathbf{f}}_D\}$ .

**3.2. Worst-case limit load factor.** For a given (but uncertain) dead load  $\mathbf{f}_D \in \mathbb{R}^{n^d}$ , we define a set  $\mathcal{Q}(\mathbf{f}_D) \subseteq \mathbb{R}^{n^m+1}$  as

$$\mathcal{Q}(\mathbf{f}_D) := \left\{ (\lambda, \mathbf{q}) \in \mathbb{R} \times \mathbb{R}^{n^m} \mid \mathbf{B}\mathbf{q} = \mathbf{f}_D + \lambda\mathbf{f}_R, |q_i| - q_i^y \leq 0, i = 1, \dots, n^m \right\}, \quad (10)$$

which is the set of all statically admissible vectors  $(\lambda, \mathbf{q}) \in \mathbb{R} \times \mathbb{R}^{n^m}$  associated with the fixed  $\mathbf{f}_D$ . We also define  $\lambda^* : \mathbb{R}^{n^d} \rightarrow \mathbb{R}$  as

$$\lambda^*(\mathbf{f}_D) = \max_{\lambda, \mathbf{q}} \left\{ \lambda : (\lambda, \mathbf{q}) \in \mathcal{Q}(\mathbf{f}_D) \right\}. \quad (11)$$

According to the static principle (5),  $\lambda^*(\mathbf{f}_D)$  corresponds to the limit load factor under the dead load  $\mathbf{f}_D$ , the external load in the decomposition (3) of the total external load  $\mathbf{f}$ , while  $\mathbf{f}_D$  is called the dead load, for the sake of simplicity. In civil engineering, it is usually held that  $\mathbf{f}_D$  consists of the conventional dead load caused by the weight of the truss itself and the live load caused by the nonstructural masses. In contrast,  $\mathbf{f}_R$  is the reference disturbance load, and  $\lambda\mathbf{f}_R$  is regarded as the load caused by earthquakes, winds, and the like. Hence, the uncertainty of live load, as well as that of dead load, can be represented by the uncertainty of  $\mathbf{f}_D$ . For consistency in the concept of the limit load factor, we make the following assumption throughout our study:

**Assumption 3.2.** The uncertainty set  $\mathcal{F}_D(\alpha)$  is chosen so that

$$\lambda^*(\mathbf{f}_D) > 0, \quad \text{for all } \mathbf{f}_D \in \mathcal{F}_D(\alpha).$$

This assumption guarantees that the structure does not collapse without applying the load  $\lambda\mathbf{f}_R$  ( $\lambda > 0$ ). In other words, if it is not satisfied, there exists  $\mathbf{f}'_D \in \mathcal{F}_D(\alpha)$  such that the structure collapses only with the dead load  $\mathbf{f}'_D$ .

We can now introduce a concept of the worst-case limit load factor, by considering that the limit load factor can be regarded as a function of  $\mathbf{f}_D$  as seen in Equation (11), while  $\mathbf{f}_D$  is uncertain and running through  $\mathcal{F}_D(\alpha)$ . Certainly, to evaluate robustness of trusses quantitatively, we are interested in the most severe situation, if any, in which the limit load factor happens to decrease unexpectedly from the nominal limit load factor corresponding to  $\tilde{\mathbf{f}}_D$  because of the uncertainty of  $\mathbf{f}_D$ . To this end, we attempt to compute the minimum value of the limit load factor that can be attained at some load satisfying  $\mathbf{f}_D \in \mathcal{F}_D(\alpha)$ . This is naturally realized by introducing  $\lambda_{\min} : \mathbb{R}_+ \rightarrow \mathbb{R}$  as

$$\lambda_{\min}(\alpha) = \min_{\mathbf{f}_D} \left\{ \lambda^*(\mathbf{f}_D) : \mathbf{f}_D \in \mathcal{F}_D(\alpha) \right\}. \quad (12)$$

Substitution of (9) into (12) yields

$$\lambda_{\min}(\alpha) = \min_{\boldsymbol{\zeta}} \left\{ \lambda^*(\mathbf{f}_D(\boldsymbol{\zeta})) : \alpha \geq \|\boldsymbol{\zeta}\|_{\infty} \right\}. \quad (13)$$

Let  $\boldsymbol{\zeta}^{\text{cr}}$ , which we call the *critical uncertain parameters vector*, denote an optimal solution of (13). Given the uncertainty parameter  $\alpha$ , we refer to  $\lambda_{\min}(\alpha)$  defined by (13) as the worst-case limit load factor, that is the minimum value among limit load factors  $\lambda^*(\mathbf{f}_D)$  corresponding to  $\mathbf{f}_D \in \mathcal{F}_D(\alpha)$ . The corresponding

dead load  $f_D(\zeta^{cr})$  is called the critical load. In the case without uncertainty, we easily see that the following relationship holds:

$$\lambda_{\min}(0) = \lambda^*(\tilde{f}_D) = \lambda^*(f_D(\mathbf{0})),$$

where  $\lambda^*(\tilde{f}_D)$  is the nominal limit load factor, that is, the limit load factor corresponding to the nominal dead load  $\tilde{f}_D$ .

The objective of this paper, then, is to propose a solution technique for computing both  $\lambda_{\min}(\alpha)$  and  $\zeta^{cr}$ .

**Remark 3.3.** In this paper we suppose that only the external load has uncertainty, and that the strength (or the admissible axial force)  $q_i^y$  of each member is certain. This is because the worst case associated with the uncertainty of member strength can be found easily for the limit analysis. Indeed, the limit load factor monotonically decreases if member strength decreases. Hence, the set of critical member strength corresponds to the trivial case in which the strength of each member coincides with its lower bound. On the contrary, the loading pattern that gives the worst case is not trivial, which motivates us to confine attention to the uncertainty of  $f_D$ . We also assume throughout that  $f_R$  is certain. Suppose that  $f_R$  varies proportionally with the fixed direction. Then the worst-case limit load factor is obtained simply by scaling the limit load factor corresponding to the nominal value of  $f_R$ . In future work, it may be interesting to consider the case where the distribution and/or the direction of  $f_R$  include uncertainty.

**Remark 3.4.** The limit load factor can be computed easily if the loading pattern of the additional dead load is fixed. Suppose that  $\zeta$  in Equation (9) is defined as  $\zeta = \beta \zeta^0$  with a given constant  $\zeta^0$  and a parameter  $\beta$ . After finding the nominal limit load factor  $\lambda^*(f_D(\mathbf{0}))$  by employing conventional limit analysis, the variation of  $\lambda^*(f_D(\beta \zeta^0))$  with respect to  $\beta$  can be computed by simply using a parametric linear programming [Chvátal 1983] approach. In our problem, the direction of  $\zeta$  is unknown and should be determined so as to minimize  $\lambda^*(f_D(\zeta))$ . Again, note that here we are trying to find a global optimal solution of (13) which is essentially nonconvex.

**3.3. Some relevant problems.** In the remainder of this section, we prepare a reformulation of (13) into the mixed 0-1 programming problem we will present in Section 4. Defining

$$\mathcal{U} = \left\{ (\mathbf{u}, \mathbf{z}) \in \mathbb{R}^{n^m} \times \mathbb{R}^{n^d} \mid \mathbf{f}_R^\top \mathbf{u} = 1, z_i \geq |\mathbf{b}_i^\top \mathbf{u}|, i = 1, \dots, n^m \right\}, \tag{14}$$

consider the following problem in the variables  $(\mathbf{u}, \mathbf{z}) \in \mathbb{R}^{n^d} \times \mathbb{R}^{n^m}$ :

$$v^*(f_D) := \min_{\mathbf{u}, \mathbf{z}} \left\{ -\mathbf{f}_D^\top \mathbf{u} + \mathbf{q}^y \mathbf{z} : (\mathbf{u}, \mathbf{z}) \in \mathcal{U} \right\}. \tag{15}$$

**Proposition 3.5** (Relation between problems (5) and (15)). Let  $(\bar{\mathbf{u}}, \bar{\mathbf{z}})$  denote an optimal solution of Problem (15). Then,

- (i)  $v^*(f_D) = \lambda^*(f_D)$ ;
- (ii)  $\bar{\mathbf{u}}$  corresponds to a collapse mode associated with  $f_D$ ;
- (iii)  $\bar{z}_i$  corresponds to the member elongation compatible to  $\bar{\mathbf{u}}$ .

*Proof.* We prove this proposition by showing that Equation (15) is dual to the static principle problem (5). Regarding the constraint

$$q_i^y \geq |q_i| \tag{16}$$

in (5), observe that  $q_i$  satisfies (16) if and only if

$$(q_i^y, q_i) \cdot (z_i, w_i) \geq 0$$

holds for any  $(z_i, w_i)$  where

$$z_i \geq |w_i|. \tag{17}$$

As a result, the function

$$L(\lambda, \mathbf{q}, \mathbf{u}, \mathbf{z}, \mathbf{w}) = \begin{cases} \lambda + \mathbf{u}^\top (\mathbf{B}\mathbf{q} - \mathbf{f}_D - \lambda \mathbf{f}_R) + (\mathbf{z}^\top \mathbf{q}^y + \mathbf{w}^\top \mathbf{q}), & \text{if } z_i \geq |w_i|, \ i = 1, \dots, n^m, \\ -\infty, & \text{otherwise,} \end{cases}$$

corresponds to the Lagrangian of (5), where  $\mathbf{u} \in \mathbb{R}^{n^d}$ ,  $\mathbf{z} \in \mathbb{R}^{n^m}$ , and  $\mathbf{w} \in \mathbb{R}^{n^m}$  are the Lagrangian multipliers. Then the Lagrangian dual of (5) is

$$\min_{\mathbf{u}, \mathbf{z}, \mathbf{w}} \sup \left\{ L(\lambda, \mathbf{q}, \mathbf{u}, \mathbf{z}, \mathbf{w}) : (\lambda, \mathbf{q}) \in \mathbb{R} \times \mathbb{R}^{n^m} \right\},$$

the explicit form of which is easily obtained as

$$\min \left\{ -\mathbf{f}_D^\top \mathbf{u} + \mathbf{q}^{y\top} \mathbf{z} : w_i = -\mathbf{b}_i^\top \mathbf{u}, \ \mathbf{f}_R^\top \mathbf{u} = 1, \ z_i \geq |w_i|, \ i = 1, \dots, n^m \right\}. \tag{18}$$

Eliminating  $\mathbf{w}$  from (18) yields (15). Hence, the LP problem (15) is dual to the LP problem (5). From this, and the strong duality of LP [Chvátal 1983], we obtain the assertions (i) and (ii). Optimal solutions of (5) and (18) satisfy the complementarity condition

$$z_i q_i^y + w_i q_i = 0 \tag{19}$$

over the constraints (16) and (17). Since (16), (17), and (19) imply  $z_i = -w_i$ , we see that  $\bar{z}_i = \mathbf{b}_i^\top \bar{\mathbf{u}}$  is satisfied at an optimal solution of (15), and we obtain assertion (iii).  $\square$

**Remark 3.6.** Note that the upper bound principle (15) is different from the well-known formulation for trusses (see, for example, [Muralidhar and Jagannatha Rao 1997]). Observe that the yield condition (4) in (5) can be rewritten as

$$q_i^y - q_i \geq 0, \quad q_i^y + q_i \geq 0, \quad i = 1, \dots, n^m. \tag{20}$$

The elongation  $c_i$  defined in (2) is divided into the two parts as

$$c_i = c_i^+ - c_i^-, \quad c_i^+ \geq 0, \quad c_i^- \geq 0, \quad i = 1, \dots, n^m. \tag{21}$$

Using (20) and (21), the set of relations governing the elastic and plastic behavior can be written as

$$\begin{aligned}
 \mathbf{B}\mathbf{q} &= \mathbf{f}_D + \lambda \mathbf{f}_R, & (\text{equilibrium}) \\
 \mathbf{q}^y - \mathbf{q} &\geq \mathbf{0}, \quad \mathbf{q}^y + \mathbf{q} \geq \mathbf{0}, & (\text{yield conditions}) \\
 \mathbf{f}_R^\top \mathbf{u} &= 1, & (\text{normalization}) \\
 \mathbf{c}^+ - \mathbf{c}^- &= \mathbf{B}^\top \mathbf{u}, & (\text{compatibility}) \\
 \mathbf{c}^+ &\geq \mathbf{0}, \quad \mathbf{c}^- \geq \mathbf{0}, & (\text{plastic elongation}) \\
 (\mathbf{q}^y - \mathbf{q})^\top \mathbf{c}^+ &= 0, \quad (\mathbf{q}^y + \mathbf{q})^\top \mathbf{c}^- = 0. & (\text{complementarity})
 \end{aligned} \tag{22}$$

From (21) and (22) it also follows that the dual to (5) can be formulated in the variables  $\mathbf{u} \in \mathbb{R}^{n^d}$ ,  $\mathbf{c}^+ \in \mathbb{R}^{n^m}$ , and  $\mathbf{c}^- \in \mathbb{R}^{n^m}$  as

$$\min \left\{ -\mathbf{f}_D^\top \mathbf{u} + \mathbf{q}^{y\top} (\mathbf{c}^+ + \mathbf{c}^-) : \mathbf{f}_R^\top \mathbf{u} = 1, \quad \mathbf{c}^+ - \mathbf{c}^- = \mathbf{B}^\top \mathbf{u}, \quad \mathbf{c}^+ \geq \mathbf{0}, \quad \mathbf{c}^- \geq \mathbf{0} \right\}, \tag{23}$$

which coincides with the conventional formulation of upper-bound principle [Muralidhar and Jagannatha Rao 1997]. However, the number of variables in (23) is larger than that of (15), which may imply an advantage of (15) over (23).

For  $\alpha \in \mathbb{R}_+$ , consider the following nonconvex problem in the variables  $(\mathbf{u}, \mathbf{z}, \boldsymbol{\zeta}) \in \mathbb{R}^{n^d} \times \mathbb{R}^{n^m} \times \mathbb{R}^m$ :

$$v_{\min}(\alpha) := \min_{\mathbf{u}, \mathbf{z}, \boldsymbol{\zeta}} \left\{ -(\tilde{\mathbf{f}}_D + \mathbf{T}\boldsymbol{\zeta})^\top \mathbf{u} + \mathbf{q}^{y\top} \mathbf{z} : (\mathbf{u}, \mathbf{z}) \in \mathcal{U}, \quad \alpha \geq \|\boldsymbol{\zeta}\|_\infty \right\}. \tag{24}$$

The following proposition shows that (24) corresponds to the kinematic version of the worst-case limit analysis (13):

**Proposition 3.7** (Relation between problems (13) and (24)). Let  $(\bar{\mathbf{u}}, \bar{\mathbf{z}}, \bar{\boldsymbol{\zeta}})$  denote an optimal solution of Problem (24). Then,

- (i)  $v_{\min}(\alpha) = \lambda_{\min}(\alpha)$ ;
- (ii)  $\bar{\boldsymbol{\zeta}}$  is an optimal solution of (13);
- (iii)  $\bar{\mathbf{u}}$  corresponds to a collapse mode associated with the external dead load  $\mathbf{f}_D(\bar{\boldsymbol{\zeta}})$ ;
- (iv)  $\bar{z}_i$  corresponds to the member elongation compatible with  $\bar{\mathbf{u}}$ .

*Proof.* By using the definition (11) of  $\lambda^*$ , (13) can be rewritten equivalently as

$$\min_{\boldsymbol{\zeta}} \left\{ \max_{\lambda, \mathbf{q}} \left\{ \lambda : (\lambda, \mathbf{q}) \in \mathcal{Q}(\mathbf{f}_D(\boldsymbol{\zeta})) \right\} : \alpha \geq \|\boldsymbol{\zeta}\|_\infty \right\} \tag{25}$$

without changing the optimal value. Let  $(\hat{\boldsymbol{\zeta}}, \hat{\lambda}, \hat{\mathbf{q}})$  denote an optimal solution of (25). It is obvious that  $\hat{\boldsymbol{\zeta}}$  is an optimal solution of (13), and that  $\hat{\lambda} = \lambda_{\min}(\alpha)$ . Since the inner problem of (25) coincides with the static principle (13),  $\hat{\mathbf{q}}$  corresponds to the vector of axial forces at the collapse mode. By using Proposition 3.5, we can rewrite the inner problem of (25) as

$$\min_{\boldsymbol{\zeta}} \left\{ \min_{\mathbf{u}, \mathbf{z}} \left\{ -\mathbf{f}_D(\boldsymbol{\zeta})^\top \mathbf{u} + \mathbf{q}^{y\top} \mathbf{z} : (\mathbf{u}, \mathbf{z}) \in \mathcal{U} \right\} : \alpha \geq \|\boldsymbol{\zeta}\|_\infty \right\} \tag{26}$$

without changing the optimal value. Obviously,  $\widehat{\boldsymbol{\zeta}}$  is an optimal solution of (25) if and only if it is an optimal solution of (26). Moreover, Proposition 3.5 guarantees that, at an optimal solution of (26),  $\mathbf{u}$  and  $z_i$ , respectively, coincide with the collapse mode and the member elongation corresponding to  $\widehat{\mathbf{q}}$ . From (6), we see that Problems (24) and (26) share the same optimal value and same optimal solutions, which concludes the proof.  $\square$

Proposition 3.7 justifies solving (24) instead of the bilevel optimization problem (13), that is, the worst-case limit load factor is obtained as the optimal value of (24). The critical load and the corresponding collapse mode can be obtained simultaneously as the optimal variables of (24). Note that (24) is a nonconvex (but single-level) problem, since the objective function includes the nonconvex quadratic term  $\boldsymbol{\zeta}^\top \mathbf{T}^\top \mathbf{u}$ . Hence, the conventional nonlinear programming approach converges to a local optimal or stationary solution in general. It should be emphasized that, for the purpose of the robustness analysis, the proof of global optimum of (24) is strongly desired, since it guarantees that the limit load factor cannot be smaller than the obtained optimal objective value. This is the major difficulty of the worst-case limit analysis. To overcome this difficulty, in the following section we propose an algorithm that converges to a global optimal solution of (24).

#### 4. Global optimization for uncertain limit analysis

In this section, we propose an algorithm to find a global optimal solution of (13) based on enumeration.

**4.1. Mixed 0-1 programming formulation.** We start by reformulating (13) as a mixed 0-1 programming problem. Letting

$$\mathcal{C}^0 := \mathbb{R}^{n^d} \times \mathbb{R}^{n^m} \times \mathbb{R}^m \times \mathbb{R}^m,$$

we define a set  $\mathcal{H} \subseteq \mathcal{C}^0$  as

$$\mathcal{H} = \left\{ (\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in \mathcal{C}^0 \left| \begin{array}{l} \mathbf{f}_R^\top \mathbf{u} = 1, \\ \mathbf{z} - \mathbf{B}^\top \mathbf{u} \geq \mathbf{0}, \quad \mathbf{z} + \mathbf{B}^\top \mathbf{u} \geq \mathbf{0}, \\ \boldsymbol{\gamma} - \mathbf{T}^\top \mathbf{u} \leq M(\mathbf{1} - \boldsymbol{\tau}), \quad \boldsymbol{\gamma} + \mathbf{T}^\top \mathbf{u} \leq M\boldsymbol{\tau}, \\ \mathbf{0} \leq \boldsymbol{\tau} \leq \mathbf{1} \end{array} \right. \right\}, \quad (27)$$

where  $M \in \mathbb{R}_+$  is a sufficiently large constant. Let

$$\mathcal{H}^Z = \{ (\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in \mathcal{H} \mid \boldsymbol{\tau} \in \{0, 1\}^m \}. \quad (28)$$

Consider the following optimization problem in the variables  $(\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in \mathcal{C}^0$ :

$$\min \left\{ -\alpha \mathbf{1}^\top \boldsymbol{\gamma} - \widetilde{\mathbf{f}}_D^\top \mathbf{u} + \mathbf{q}^{y^\top} \mathbf{z} : (\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in \mathcal{H}^Z \right\}. \quad (29)$$

We refer to (29) as the mixed 0-1 programming problem. It has binary constraints on  $\boldsymbol{\tau}$ , linear inequality constraints, and a linear objective function.

**Proposition 4.1** (Relation between problems (24) and (29)). A feasible solution  $(\bar{\mathbf{u}}, \bar{\mathbf{z}}, \bar{\boldsymbol{\zeta}})$  of (24) satisfying

$$\bar{\boldsymbol{\zeta}}^\top \mathbf{T}^\top \bar{\mathbf{u}} = \alpha \|\mathbf{T}^\top \bar{\mathbf{u}}\|_1 \quad (30)$$

is optimal if and only if a feasible solution  $(\bar{\mathbf{u}}, \bar{\mathbf{z}}, \bar{\boldsymbol{\gamma}}, \bar{\boldsymbol{\tau}})$  of Problem (29) satisfying

$$\bar{\gamma}_j = |\mathbf{t}_j^\top \bar{\mathbf{u}}|, \quad j = 1, \dots, m, \quad (31)$$

$$\begin{cases} \bar{\tau}_j = 1, & \text{if } \mathbf{t}_j^\top \bar{\mathbf{u}} > 0, \\ \bar{\tau}_j = 0, & \text{if } \mathbf{t}_j^\top \bar{\mathbf{u}} < 0, \\ \bar{\tau}_j \in \{0, 1\}, & \text{if } \mathbf{t}_j^\top \bar{\mathbf{u}} = 0, \end{cases} \quad (32)$$

is optimal. Moreover, Problems (24) and (29) share the same optimal value, that is, equal to  $\lambda_{\min}(\alpha)$ .

*Proof.* Observe that, in (24), only  $\alpha \geq \|\boldsymbol{\zeta}\|_\infty$  is the constraint on  $\boldsymbol{\zeta}$ , which is independent of the remaining variables  $\mathbf{z}$  and  $\mathbf{u}$ . Hence, (24) is equivalently rewritten as

$$\min_{\mathbf{u}, \mathbf{z}} \left\{ \min_{\boldsymbol{\zeta}} \left\{ -(\tilde{\mathbf{f}}_D + \mathbf{T}\boldsymbol{\zeta})^\top \mathbf{u} : \alpha \geq \|\boldsymbol{\zeta}\|_\infty \right\} + \mathbf{q}^y \top \mathbf{z} : (\mathbf{u}, \mathbf{z}) \in \mathcal{U} \right\} \quad (33)$$

without changing the optimal value and optimal solution. From the Hölder inequality [Marti and Stoeckel 2004, Chap. 9], we see that

$$(\mathbf{T}\boldsymbol{\zeta})^\top \mathbf{u} \geq \|\boldsymbol{\zeta}\|_\infty \|\mathbf{T}^\top \mathbf{u}\|_1 \quad (34)$$

holds for any fixed  $\mathbf{u}$ . Moreover, Assumption 3.1 (i) guarantees that there exists a  $\boldsymbol{\zeta}$  satisfying

$$(\mathbf{T}\boldsymbol{\zeta})^\top \mathbf{u} = \|\boldsymbol{\zeta}\|_\infty \|\mathbf{T}^\top \mathbf{u}\|_1. \quad (35)$$

From (34) and (35), we obtain

$$\min_{\boldsymbol{\zeta}} \left\{ -(\mathbf{T}\boldsymbol{\zeta})^\top \mathbf{u} : \alpha \geq \|\boldsymbol{\zeta}\|_\infty \right\} = \min_{\boldsymbol{\zeta}} \left\{ -\|\boldsymbol{\zeta}\|_\infty \|\mathbf{T}^\top \mathbf{u}\|_1 : \alpha \geq \|\boldsymbol{\zeta}\|_\infty \right\} = -\alpha \|\mathbf{T}^\top \mathbf{u}\|_1,$$

where an optimal  $\boldsymbol{\zeta}$  satisfies (35). Consequently, the variable  $\boldsymbol{\zeta}$  can be eliminated from (33) as

$$\min_{\mathbf{u}, \mathbf{z}} \left\{ -\alpha \|\mathbf{T}^\top \mathbf{u}\|_1 - \tilde{\mathbf{f}}_D^\top \mathbf{u} + \mathbf{q}^y \top \mathbf{z} : (\mathbf{u}, \mathbf{z}) \in \mathcal{U} \right\}. \quad (36)$$

Note that an optimal solution  $(\bar{\mathbf{u}}, \bar{\mathbf{z}})$  of (36) can be converted to an optimal solution  $(\bar{\mathbf{u}}, \bar{\mathbf{z}}, \bar{\boldsymbol{\xi}})$  of (33) by defining  $\bar{\boldsymbol{\xi}}$  as in (30), and these two problems share the same objective value. By introducing new variables  $\boldsymbol{\gamma} \in \mathbb{R}^m$ , (36) is equivalently rewritten as

$$\min \left\{ -\alpha \mathbf{1}^\top \boldsymbol{\gamma} - \tilde{\mathbf{f}}_D^\top \mathbf{u} + \mathbf{q}^y \top \mathbf{z} : (\mathbf{u}, \mathbf{z}) \in \mathcal{U}, (\gamma_j = \mathbf{t}_j^\top \mathbf{u}) \vee (\gamma_j = -\mathbf{t}_j^\top \mathbf{u}), j = 1, \dots, m \right\}. \quad (37)$$

where  $\vee$  denotes logical ‘or’. Note that (31) holds at an optimal solution of (37). By using a sufficiently large constant  $M$ , the disjunction

$$\left( \gamma_j \leq \mathbf{t}_j^\top \mathbf{u} \right) \vee \left( \gamma_j \leq -\mathbf{t}_j^\top \mathbf{u} \right)$$

is equivalently rewritten as

$$\gamma_j \leq \mathbf{t}_j^\top \mathbf{u} + M(1 - \tau_j), \quad \gamma_j \leq -\mathbf{t}_j^\top \mathbf{u} + M\tau_j, \quad \tau_j \in \{0, 1\},$$

with the relation (32), which completes the proof. □



**4.2. Branch-and-bound method for problem (29).** The LP relaxation of the mixed 0-1 programming problem (29) is obtained by ignoring the binary constraints on  $\boldsymbol{\tau}$  as

$$\min \left\{ -\alpha \mathbf{1}^\top \boldsymbol{\gamma} - \tilde{\mathbf{f}}_D^\top \mathbf{u} + \mathbf{q}^{y^\top} \mathbf{z} : (\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in \mathcal{H} \right\}. \quad (38)$$

Define a set  $\mathcal{C}$  as

$$\mathcal{C} = \left\{ (\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in \mathcal{C}^0 \mid \mathbf{A}_u^\top \mathbf{u} + \mathbf{A}_z^\top \mathbf{z} + \mathbf{A}_\gamma^\top \boldsymbol{\gamma} + \mathbf{A}_\tau^\top \boldsymbol{\tau} \geq \mathbf{b} \right\}, \quad (39)$$

where  $\mathbf{A}_u \in \mathbb{R}^{n^d \times n^c}$ ,  $\mathbf{A}_z \in \mathbb{R}^{n^m \times n^c}$ ,  $\mathbf{A}_\gamma \in \mathbb{R}^{m \times n^c}$ ,  $\mathbf{A}_\tau \in \mathbb{R}^{m \times n^c}$ , and  $\mathbf{b} \in \mathbb{R}^{n^c}$  are constant matrices and a constant vector. Assume that  $\mathcal{C}$  satisfies

$$\text{cl conv } \mathcal{H} \subseteq \mathcal{C} \subseteq \mathcal{C}^0. \quad (40)$$

Note that in this section we set  $\mathcal{C} := \mathcal{C}^0$ , while in Section 5 we discuss how to generate a proper subset  $\mathcal{C}$  of  $\mathcal{C}^0$ .

Let  $\mathcal{F}_0^k$  and  $\mathcal{F}_1^k$  denote the subsets of indices satisfying

$$\mathcal{F}_0^k \subseteq \{1, \dots, m\}, \quad \mathcal{F}_1^k \subseteq \{1, \dots, m\}, \quad \mathcal{F}_0^k \cap \mathcal{F}_1^k = \emptyset.$$

Let

$$\mathcal{H}(\mathcal{C}, \mathcal{F}_0^k, \mathcal{F}_1^k) = \left\{ (\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in \mathcal{H} \cap \mathcal{C} \mid \tau_j = 0 \text{ for } j \in \mathcal{F}_0^k, \tau_j = 1 \text{ for } j \in \mathcal{F}_1^k \right\},$$

where  $\mathcal{H}$  and  $\mathcal{C}$  have been defined in (27) and (39). Consider the following LP problem in the variables  $(\mathbf{z}, \mathbf{u}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in \mathcal{C}^0$ :

$$\text{LP}(\mathcal{C}, \mathcal{F}_0^k, \mathcal{F}_1^k) : v^k := \min \left\{ -\alpha \mathbf{1}^\top \boldsymbol{\gamma} - \tilde{\mathbf{f}}_D^\top \mathbf{u} + \mathbf{q}^{y^\top} \mathbf{z} : (\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in \mathcal{H}(\mathcal{C}, \mathcal{F}_0^k, \mathcal{F}_1^k) \right\}. \quad (41)$$

Explicitly, (41) is written as

$$\min \quad \left\{ -\alpha \mathbf{1}^\top \boldsymbol{\gamma} - \tilde{\mathbf{f}}_D^\top \mathbf{u} + \mathbf{q}^{y^\top} \mathbf{z} : \right. \quad (42a)$$

$$\left. \mathbf{f}_R^\top \mathbf{u} = 1, \right. \quad (42b)$$

$$\left. \mathbf{z} - \mathbf{B}^\top \mathbf{u} \geq \mathbf{0}, \right. \quad (42c)$$

$$\left. \mathbf{z} + \mathbf{B}^\top \mathbf{u} \geq \mathbf{0}, \right. \quad (42d)$$

$$\left. \boldsymbol{\gamma} - \mathbf{T}^\top \mathbf{u} \leq M(\mathbf{1} - \boldsymbol{\tau}), \right. \quad (42e)$$

$$\left. \boldsymbol{\gamma} + \mathbf{T}^\top \mathbf{u} \leq M\boldsymbol{\tau}, \right. \quad (42f)$$

$$\left. \mathbf{0} \leq \boldsymbol{\tau} \leq \mathbf{1}, \right. \quad (42g)$$

$$\left. \mathbf{A}_u^\top \mathbf{u} + \mathbf{A}_z^\top \mathbf{z} + \mathbf{A}_\gamma^\top \boldsymbol{\gamma} + \mathbf{A}_\tau^\top \boldsymbol{\tau} \geq \mathbf{b}, \right. \quad (42h)$$

$$\left. \tau_j = 0 \text{ for } j \in \mathcal{F}_0^k, \tau_j = 1 \text{ for } j \in \mathcal{F}_1^k \right\}. \quad (42i)$$

We solve  $\text{LP}(\mathcal{C}, \mathcal{F}_0^k, \mathcal{F}_1^k)$  at the nodes of enumeration tree. Note that  $\text{LP}(\mathcal{C}^0, \emptyset, \emptyset)$  coincides with the LP relaxation (38).

The following is a branch-and-bound method for solving the mixed 0-1 programming problem (29) based on the LP relaxation.

**Algorithm 4.2.** Branch-and-bound algorithm for (29).

**Step 0:** Initialization. Set  $k = 0$ ,  $\mathcal{F}_1^0 = \emptyset$ ,  $\mathcal{F}_0^0 = \emptyset$ , and  $v^U = \infty$ . Choose the small tolerance  $\epsilon > 0$  and  $\mathcal{C}$  satisfying (40) (set  $\mathcal{C} := \mathcal{C}^0$  in this section).

**Step 1:** Solving the subproblem. Solve the linear program  $\text{LP}(\mathcal{C}, \mathcal{F}_0^k, \mathcal{F}_1^k)$  defined in (41). If the problem is infeasible, go to Step 5; otherwise, let  $(\mathbf{u}^k, \mathbf{z}^k, \boldsymbol{\gamma}^k, \boldsymbol{\tau}^k)$  and  $v^k$  denote its optimal solution and optimal objective value, respectively.

**Step 2:** Fathoming. If  $v^k \geq v^U$ , go to Step 5.

**Step 3:** Branching. If  $(\boldsymbol{\tau}^k)^\top (1 - \boldsymbol{\tau}^k) \leq \epsilon$ , go to Step 4; otherwise, select an index  $j_1$  such that

$$j_1 = \arg \max_{j \in \{1, \dots, m\}} \left\{ \tau_j^k (1 - \tau_j^k) \right\}.$$

Set  $\mathcal{F}_1^{k+1} := \mathcal{F}_1^k \cup \{j_1\}$  and  $\mathbf{p}^{k+1} := (\mathbf{p}^k, j_1)^\top$ . Update  $k \leftarrow k + 1$ , and go to Step 1.

**Step 4:** Updating. Put  $v^U := v^k$  and  $(\bar{\mathbf{u}}, \bar{\mathbf{z}}, \bar{\boldsymbol{\gamma}}, \bar{\boldsymbol{\tau}}) := (\mathbf{u}^k, \mathbf{z}^k, \boldsymbol{\gamma}^k, \boldsymbol{\tau}^k)$ . Go to Step 5.

**Step 5:** Backtracking. If  $\mathbf{p}^k < \mathbf{0}$ , go to Step 6; otherwise branch to a new live node as follows:

Define  $l_1 = \max \{l \in \{1, \dots, L\} \mid p_l^k > 0\}$ , where  $L$  denotes the size of the vector  $\mathbf{p}^k$ . Divide  $\mathbf{p}^k$  into the three parts

$$\mathbf{p}_1 = (p_l^k \mid l = 1, \dots, l_1 - 1), \quad \mathbf{p}_2 = p_{l_1}^k, \quad \mathcal{P}_3 = \{-p_l^k \mid l = l_1 + 1, \dots, L\}.$$

Set

$$\begin{aligned} \mathbf{p}^{k+1} &:= (\mathbf{p}_1^\top, -p_2)^\top, \\ \mathcal{F}_0^{k+1} &:= \{\mathcal{F}_0^k \cup \{p_2\}\} \setminus \mathcal{P}_3, \\ \mathcal{F}_1^{k+1} &:= \mathcal{F}_1^k \setminus \{p_2\}. \end{aligned}$$

Update  $k \leftarrow k + 1$ , and go to Step 1.

**Step 6:** Termination. Declare  $(\bar{\mathbf{z}}, \bar{\mathbf{u}}, \bar{\boldsymbol{\gamma}}, \bar{\boldsymbol{\tau}})$  as the optimal solution, and stop.

**Remark 4.3.** Essentially, Algorithm 4.2 is designed by using depth-first search (see, for example [Clausen and Perregaard 1999]) as a strategy for selecting the next live subproblem at Step 5. The condition  $\mathbf{p}^k < \mathbf{0}$  implies that there exists no live node. Among the live subproblems, we always select the subproblem with the largest level in the branch-and-bound tree. The vector  $\mathbf{p}^k$  plays the role of bookkeeping the path from the root node to the current node in the branch-and-bound tree. The size  $L$  of  $\mathbf{p}^k = (p_j^k)$  coincides with the current depth of the tree, and we see that the following relations hold:

$$\begin{aligned} \mathcal{F}_0^k &= \left\{ p_j^k \mid p_j^k \leq 0, \quad j = 1, \dots, L \right\}, \\ \mathcal{F}_1^k &= \left\{ -p_j^k \mid p_j^k \geq 0, \quad j = 1, \dots, L \right\}, \end{aligned}$$

that is, the components of  $\mathbf{p}^k$  correspond to the indices of  $\tau_j$ , possibly with opposite signs, which are fixed in the current subproblem  $\text{LP}(\mathcal{C}, \mathcal{F}_0^k, \mathcal{F}_1^k)$ . The remaining  $\tau_j$  are not fixed in  $\text{LP}(\mathcal{C}, \mathcal{F}_0^k, \mathcal{F}_1^k)$ . The order of an element  $p_j^k$  of  $\mathbf{p}^k$  is determined by its level in the tree.

**Remark 4.4.** Observe that the binary constraints  $\boldsymbol{\tau} \in \{0, 1\}^m$  are equivalent to the following complementarity conditions:

$$\boldsymbol{\tau} \geq \mathbf{0}, \quad \mathbf{1} - \boldsymbol{\tau} \geq \mathbf{0}, \tag{43}$$

$$\tau_j(1 - \tau_j) = 0, \quad j = 1, \dots, m. \tag{44}$$

Notice here that any feasible solution of  $\text{LP}(\mathcal{C}, \mathcal{F}_0^k, \mathcal{F}_1^k)$  satisfies (43). At Step 3, we make a check if the current solution  $(\mathbf{u}^k, \mathbf{z}^k, \boldsymbol{\gamma}^k, \boldsymbol{\tau}^k)$  satisfies the complementarity conditions (44) or not. Satisfaction (possibly with small tolerance in practice) implies that  $(\mathbf{u}^k, \mathbf{z}^k, \boldsymbol{\gamma}^k, \boldsymbol{\tau}^k)$  is a feasible solution of (29). Alternatively, if (44) is not satisfied, then the variable  $\tau_j$  with the largest residual of the complementarity (44) is used as the branching variable in Step 3.

**Remark 4.5.** Note that it is not difficult to randomly generate  $\mathbf{f}'_D$  satisfying  $\mathbf{f}'_D \in \mathcal{F}_D(\alpha)$ . Then the corresponding limit load factor  $\lambda^*(\mathbf{f}'_D)$  provides an upper bound of the mixed 0-1 (29). At Step 0, we can obtain an upper bound  $v^U$  by solving (11) several times for randomly sampled  $\mathbf{f}'_D$ . We simply set  $v^U = \infty$  if this process is skipped.

**Remark 4.6.** In this paper, we focus on finding the global optimal solution of (13). Alternatively, for very large structures, it is also important to develop an efficient algorithm for computing a lower bound of  $\lambda_{\min}(\alpha)$ . For this purpose, it may be interesting to investigate a relaxation of (13) based on the concept of uncertain LP, as presented in [Ben-Tal et al. 2004].

**4.3. Duality and simplification.** The remainder of this section is devoted to some practical issues regarding implementation of Algorithm 4.2. In fact, to obtain  $(\mathbf{u}^k, \mathbf{z}^k, \boldsymbol{\gamma}^k, \boldsymbol{\tau}^k)$  at Step 1, we do not solve (41) directly but use the simplex method to solve its Lagrangian dual problem, denoted by  $\text{LP}^*(\mathcal{C}, \mathcal{F}_0^k, \mathcal{F}_1^k)$ . Then we obtain the solution of (41) as the optimal Lagrange multipliers. From preliminary numerical experiments, we observed that the CPU time required to solve the dual problem is much smaller than that required to solve the original (41). Indeed, after the branching process of Step 3, at the new node it is easy to obtain a feasible solution of the dual problem from an optimal solution of the dual problem solved at the previous node. Let

$$\mathcal{C}^{0*} = \mathbb{R} \times \mathbb{R}^{n^m} \times \mathbb{R}^{n^m} \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m.$$

From the LP duality [Chvátal 1983] it follows that the dual problem of the LP relaxation  $\text{LP}(\mathcal{C}, \emptyset, \emptyset)$  is formulated in the variables  $(\rho^\lambda, \mathbf{q}^+, \mathbf{q}^-, \boldsymbol{\zeta}^+, \boldsymbol{\zeta}^-, \boldsymbol{\rho}^+, \boldsymbol{\rho}^-, \boldsymbol{\mu}) \in \mathcal{C}^{0*} \times \mathbb{R}^{n^c}$  as

$$\text{LP}^*(\mathcal{C}, \emptyset, \emptyset) : \quad \max \quad \{\rho^\lambda - \mathbf{1}^\top (M\boldsymbol{\rho}^+ + \boldsymbol{\rho}^-) + \mathbf{b}^\top \boldsymbol{\mu} : \tag{45a}$$

$$\mathbf{B}(\mathbf{q}^+ - \mathbf{q}^-) = \tilde{\mathbf{f}}_D + \mathbf{T}(\boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-) + \mathbf{f}_R \rho^\lambda + \mathbf{A}_u \boldsymbol{\mu}, \tag{45b}$$

$$\mathbf{q}^+ + \mathbf{q}^- + \mathbf{A}_z \boldsymbol{\mu} = \mathbf{q}^y, \tag{45c}$$

$$\boldsymbol{\zeta}^+ + \boldsymbol{\zeta}^- = \alpha \mathbf{1} + \mathbf{A}_\gamma \boldsymbol{\mu}, \tag{45d}$$

$$M(\boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-) = \boldsymbol{\rho}^+ - \boldsymbol{\rho}^- + \mathbf{A}_\tau \boldsymbol{\mu}, \tag{45e}$$

$$\mathbf{q}^+, \mathbf{q}^-, \boldsymbol{\zeta}^+, \boldsymbol{\zeta}^-, \boldsymbol{\rho}^+, \boldsymbol{\rho}^-, \boldsymbol{\mu} \geq \mathbf{0}\}. \tag{45f}$$

Let  $\text{LP}^*(\mathcal{C}, \mathcal{F}_0^k, \mathcal{F}_1^k)$  denote the dual of  $\text{LP}(\mathcal{C}, \mathcal{F}_0^k, \mathcal{F}_1^k)$  in which some variables are fixed. To obtain  $\text{LP}^*(\mathcal{C}, \mathcal{F}_0^k, \mathcal{F}_1^k)$ , we modify and simplify (45) as follows:

(a) For  $j \in \mathcal{F}_0^k$ , the variable  $\tau_j$  in the primal problem (42) is set as  $\tau_j = 0$ . This is realized in the dual problem as follows:

(i) The  $j$ -th row of (42f) should be rewritten as

$$-\mathbf{t}_j^\top \mathbf{u} - \gamma_j \geq 0.$$

Consequently, the variable  $\zeta_j^-$  should be eliminated from the  $j$ -th row of (45e).

- (ii) The  $j$ -th row of (42e) becomes redundant. Hence, the variable  $\zeta_j^+$  can be eliminated from (45).
  - (iii) The  $j$ -th row of (42g) becomes redundant. Hence, the variables  $\rho_j^+$  and  $\rho_j^-$  can be eliminated from (45).
  - (iv) If the  $j$ -th row vector of  $\mathbf{A}_\tau$  is a zero vector, then the constraint (45e) itself can be eliminated.
- (b) For  $j \in \mathcal{F}_1^k$ , the variable  $\tau_j$  in the primal problem (42) is set as  $\tau_j = 1$ . This is realized in the dual problem as follows:

(i) The  $j$ -th row of (42e) should be rewritten as

$$\mathbf{t}_j^\top \mathbf{u} - \gamma_j \geq 0.$$

Consequently, the variable  $\zeta_j^+$  should be eliminated from the  $j$ -th row of (45e).

- (ii) The  $j$ -th row of (42f) becomes redundant. Hence, the variable  $\zeta_j^-$  can be eliminated from (45).
- (iii) The  $j$ -th row of (42g) becomes redundant. Hence, the variables  $\rho_j^+$  and  $\rho_j^-$  can be eliminated from (45).
- (iv) If the  $j$ -th row vector of  $\mathbf{A}_\tau$  is a zero vector, then the constraint (45e) itself can be eliminated.

Note that the LP (45) originally has  $(2n^m + 4m + n^c + 1)$  variables and  $(n^d + n^m + 2m)$  linear equality constraints besides side constraints. With the simplification proposed in this section, we can reduce the number of variables and constraints when some binary variables  $\tau_j$  are fixed in (41).

### 5. Cutting plane algorithm

It is guaranteed that Algorithm 4.2 converges to a global optimal solution of the mixed 0-1 programming problem (29). However, it is possible that the algorithm is no better than the enumeration of all binary variables  $\boldsymbol{\tau}$ . The efficiency of the algorithm depends partially on the tightness of the LP relaxation problem solved at each node of the branch-and-bound tree.

**5.1. Disjunctive cut generation.** To strengthen the LP relaxation problems, we now propose an algorithm that generates the disjunctive cutting planes. Recall that  $\mathcal{H}^Z$  and  $\mathcal{H}$ , defined in (28) and (27), correspond to the feasible sets of the mixed 0-1 programming problem (29) and its LP relaxation (38), respectively. Let  $(\hat{\mathbf{u}}, \hat{\mathbf{z}}, \hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\tau}})$  denote the optimal solution of the LP relaxation (38). Suppose that  $\hat{\boldsymbol{\tau}}$  does not satisfy the binary constraints in  $\mathcal{H}^Z$ , that is,

$$(\hat{\mathbf{u}}, \hat{\mathbf{z}}, \hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\tau}}) \notin \mathcal{H}^Z.$$

The cutting plane, then, is an additional linear inequality that the point  $(\hat{\mathbf{u}}, \hat{\mathbf{z}}, \hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\tau}})$  does not satisfy, but is valid for  $\mathcal{H}^Z$ . If a cutting plane is generated successfully, we can add it to the LP relaxation as the constraint without cutting off any feasible solution in  $\mathcal{H}^Z$ . If the new optimal solution of the obtained LP

problem is feasible for  $\mathcal{K}^{\mathbb{Z}}$ , it is a global optimal solution of the original mixed 0-1 program problem (29); otherwise, we may continue to generate cutting planes.

In the following, the cutting plane generation is performed over the so-called *disjunctive programming* relaxation of  $\mathcal{K}^{\mathbb{Z}}$  instead of  $\mathcal{K}^{\mathbb{Z}}$  itself. A valid inequality obtained in this way is called the disjunctive cut [Ceria and Soares 1997]. We define the sets

$$P_j(\mathcal{K}) = \text{cl conv} \{(\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in \mathcal{K} \mid \tau_j \in \{0, 1\}\}, \quad j = 1, \dots, m,$$

each of which is a disjunctive programming relaxation of the closure of  $\text{conv } \mathcal{K}^{\mathbb{Z}}$ . We attempt to find a linear inequality that cuts off  $(\widehat{\mathbf{u}}, \widehat{\mathbf{z}}, \widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\tau}})$  but is valid for  $P_j(\mathcal{K})$ . Although the characterization of  $P_j(\mathcal{K})$  is essentially nonlinear, a polyhedral representation can be obtained easily [Balas et al. 1996; Balas and Perregaard 2002], that is, the condition

$$(\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in P_j(\mathcal{K})$$

is satisfied if and only if there exist

$$(\mathbf{w}^u, \mathbf{w}^z, \mathbf{w}^\gamma, \mathbf{w}^\tau) \in \mathcal{C}^0, \quad w_0 \in \mathbb{R},$$

$$(\mathbf{y}^u, \mathbf{y}^z, \mathbf{y}^\gamma, \mathbf{y}^\tau) \in \mathcal{C}^0, \quad y_0 \in \mathbb{R}$$

satisfying

$$(\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) = (\mathbf{w}^u, \mathbf{w}^z, \mathbf{w}^\gamma, \mathbf{w}^\tau) + (\mathbf{y}^u, \mathbf{y}^z, \mathbf{y}^\gamma, \mathbf{y}^\tau), \quad (46a)$$

$$\mathbf{f}_R^\top \mathbf{w}^u = w_0, \quad (46b)$$

$$\mathbf{w}^z - \mathbf{B}^\top \mathbf{w}^u \geq \mathbf{0}, \quad \mathbf{w}^z + \mathbf{B}^\top \mathbf{w}^u \geq \mathbf{0}, \quad (46c)$$

$$\mathbf{T}^\top \mathbf{w}^u - \mathbf{w}^\gamma - M \mathbf{w}^\tau \geq -M w_0 \mathbf{1}, \quad -\mathbf{T}^\top \mathbf{w}^u - \mathbf{w}^\gamma + M \mathbf{w}^\tau \geq \mathbf{0}, \quad (46d)$$

$$\mathbf{0} \leq \mathbf{w}^\tau \leq w_0 \mathbf{1}, \quad (46e)$$

$$w_j^\tau \leq 0, \quad (46f)$$

$$\mathbf{f}_R^\top \mathbf{y}^u = y_0, \quad (46g)$$

$$\mathbf{y}^z - \mathbf{B}^\top \mathbf{y}^u \geq \mathbf{0}, \quad \mathbf{y}^z + \mathbf{B}^\top \mathbf{y}^u \geq \mathbf{0}, \quad (46h)$$

$$\mathbf{T}^\top \mathbf{y}^u - \mathbf{y}^\gamma - M \mathbf{y}^\tau \geq -M y_0 \mathbf{1}, \quad -\mathbf{T}^\top \mathbf{y}^u - \mathbf{y}^\gamma + M \mathbf{y}^\tau \geq \mathbf{0}, \quad (46i)$$

$$\mathbf{0} \leq \mathbf{y}^\tau \leq y_0 \mathbf{1}, \quad (46j)$$

$$y_j^\tau \leq y_0, \quad (46k)$$

$$w_0 + y_0 = 1. \quad (46l)$$

We define a set  $P_j^*(\mathcal{K}) \subseteq \mathcal{C}^0 \times \mathbb{R}$  so that

$$(\boldsymbol{\alpha}_u, \boldsymbol{\alpha}_z, \boldsymbol{\alpha}_\gamma, \boldsymbol{\alpha}_\tau, \beta) \in P_j^*(\mathcal{K}) \quad (47)$$

holds if and only if there exist

$$\begin{aligned} (\xi^\lambda, \xi^{q+}, \xi^{q-}, \xi^{\zeta+}, \xi^{\zeta-}, \xi^{\rho+}, \xi^{\rho-}) &\in \mathcal{C}^{0*}, \quad \xi^0 \in \mathbb{R}, \\ (\eta^\lambda, \eta^{q+}, \eta^{q-}, \eta^{\zeta+}, \eta^{\zeta-}, \eta^{\rho+}, \eta^{\rho-}) &\in \mathcal{C}^{0*}, \quad \eta^0 \in \mathbb{R} \end{aligned}$$

satisfying

$$\alpha_u = f_R \xi^\lambda - \mathbf{B}(\xi^{q+} - \xi^{q-}) + \mathbf{T}(\xi^{\zeta+} - \xi^{\zeta-}), \quad (48a)$$

$$\alpha_z = \xi^{q+} + \xi^{q-}, \quad (48b)$$

$$\alpha_\gamma = -\xi^{\zeta+} - \xi^{\zeta-}, \quad (48c)$$

$$\alpha_\tau = -M(\xi^{\zeta+} - \xi^{\zeta-}) + \xi^{\rho+} - \xi^{\rho-} - \xi^0 \mathbf{e}^j, \quad (48d)$$

$$\beta = \xi^\lambda - \mathbf{1}^\top (M\xi^{\zeta+} + \xi^{\rho-}), \quad (48e)$$

$$\xi^{q+}, \xi^{q-}, \xi^{\zeta+}, \xi^{\zeta-}, \xi^{\rho+}, \xi^{\rho-}, \xi^0 \geq \mathbf{0}, \quad (48f)$$

$$\alpha_u = f_R \eta^\lambda - \mathbf{B}(\eta^{q+} - \eta^{q-}) + \mathbf{T}(\eta^{\zeta+} - \eta^{\zeta-}), \quad (48g)$$

$$\alpha_z = \eta^{q+} + \eta^{q-}, \quad (48h)$$

$$\alpha_\gamma = -\eta^{\zeta+} - \eta^{\zeta-}, \quad (48i)$$

$$\alpha_\tau = -M(\eta^{\zeta+} - \eta^{\zeta-}) + \eta^{\rho+} - \eta^{\rho-} + \eta^0 \mathbf{e}^j, \quad (48j)$$

$$\beta = \eta^\lambda - \mathbf{1}^\top (M\eta^{\zeta+} + \eta^{\rho-}) + \eta^0, \quad (48k)$$

$$\eta^{q+}, \eta^{q-}, \eta^{\zeta+}, \eta^{\zeta-}, \eta^{\rho+}, \eta^{\rho-}, \eta^0 \geq \mathbf{0}. \quad (48l)$$

Then, the following inequality is valid for  $P_j(\mathcal{K})$  if it satisfies (47):

$$(\alpha_u, \alpha_z, \alpha_\gamma, \alpha_\tau) \cdot (\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \geq \beta. \quad (49)$$

Thus, for a point  $(\widehat{\mathbf{u}}, \widehat{\mathbf{z}}, \widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\tau}}) \notin P_j(\mathcal{K})$ , we are interested in the following problem in the variables  $(\alpha_u, \alpha_z, \alpha_\gamma, \alpha_\tau, \beta) \in \mathcal{C} \times \mathbb{R}$ :

$$\max \{ \beta - (\alpha_u, \alpha_z, \alpha_\gamma, \alpha_\tau) \cdot (\widehat{\mathbf{u}}, \widehat{\mathbf{z}}, \widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\tau}}) : (\alpha_u, \alpha_z, \alpha_\gamma, \alpha_\tau, \beta) \in P_j^*(\mathcal{K}) \}, \quad (50)$$

because a feasible solution of (50) defines a valid inequality (in the form of (49)) for  $P_j(\mathcal{K})$ , that is violated at  $(\widehat{\mathbf{u}}, \widehat{\mathbf{z}}, \widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\tau}})$ .

However, since (50) itself is unbounded, some normalization constraints should be appended to it. We add constraints restricting the magnitude of the vector  $(\alpha_u, \alpha_z, \alpha_\gamma, \alpha_\tau)$  [Ceria and Soares 1997]. The index sets  $\mathcal{F}$  and  $\overline{\mathcal{F}}$  are defined as subsets of  $\{1, \dots, n^d + n^m + 2m\}$  by

$$\begin{aligned} \overline{\mathcal{F}} &= \{i \in \{1, \dots, n^d + n^m + 2m\} \mid (\widehat{\mathbf{u}}, \widehat{\mathbf{z}}, \widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\tau}})_i = 0\}, \\ \mathcal{F} &= \{1, \dots, n^d + n^m + 2m\} \setminus \overline{\mathcal{F}}. \end{aligned}$$

We assume  $\mathcal{J} \neq \emptyset$ . We then consider the cut generation problem

$$\begin{aligned}
 (\text{CGLP})_j : \max \quad & \left\{ \beta - (\alpha_u, \alpha_z, \alpha_\gamma, \alpha_\tau) \cdot (\widehat{u}, \widehat{z}, \widehat{\gamma}, \widehat{\tau}) : \right. \\
 & (\alpha_u, \alpha_z, \alpha_\gamma, \alpha_\tau, \beta) \in P_j^*(\mathcal{H}), \\
 & \left. \|(\alpha_u, \alpha_z, \alpha_\gamma, \alpha_\tau)\|_\infty \leq 1 \right\},
 \end{aligned} \tag{51}$$

where  $P_j^*(\mathcal{H})$  has been defined in (48). In (51), which is an LP problem, we attempt to find the deepest cut in the sense that a distance from  $(\widehat{u}, \widehat{z}, \widehat{\gamma}, \widehat{\tau})$  to a separating hyperplane is maximized. The dual to (51) is formulated as in [Ceria and Soares 1997] as

$$\begin{aligned}
 \min \quad & \|(\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) - (\widehat{u}, \widehat{z}, \widehat{\gamma}, \widehat{\tau})\|_1 : \\
 & (\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in P_j(\mathcal{H}), \\
 & (\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau})_{\overline{\mathcal{J}}} = (\widehat{u}, \widehat{z}, \widehat{\gamma}, \widehat{\tau})_{\overline{\mathcal{J}}},
 \end{aligned} \tag{52}$$

where  $P_j(\mathcal{H})$  has been defined in (46). At the root node of the enumeration tree of Algorithm 4.2, we employ the following procedure for generating some disjunctive cuts.

**Algorithm 5.1.** Cut generation for problem (29).

**Step 0:** Set  $\mathcal{C}^0 = \mathbb{R}^{nm} \times \mathbb{R}^{nd} \times \mathbb{R}^m \times \mathbb{R}^m$ ,  $\mathcal{J}_{\text{res}}^0 = \{1, \dots, m\}$ ,  $\mathcal{J} = \{1, \dots, m\}$ , and  $k = 1$ .  
 Let  $(\mathbf{u}^0, \mathbf{z}^0, \boldsymbol{\gamma}^0, \boldsymbol{\tau}^0)$  denote an optimal solution of  $\text{LP}(\mathcal{C}^0, \emptyset, \emptyset)$ .

**Step 1:** Select  $j_2 \in \mathcal{J}$  by

$$j_2 = \arg \max_{j \in \mathcal{J}} \left\{ \tau_j^{k-1} (1 - \tau_j^{k-1}) \right\}.$$

**Step 2:** Solve  $(\text{CGLP})_{j_2}$ , with the definition (48) of  $P_j^*(\mathcal{H})$ , at

$$(\widehat{u}, \widehat{z}, \widehat{\gamma}, \widehat{\tau}) = (\mathbf{u}^{k-1}, \mathbf{z}^{k-1}, \boldsymbol{\gamma}^{k-1}, \boldsymbol{\tau}^{k-1})$$

to find an optimal solution  $(\alpha_u^k, \alpha_z^k, \alpha_\gamma^k, \alpha_\tau^k, \beta^k)$ .

**Step 3:** Letting

$$\mathcal{C}_{\text{cur}} := \left\{ (\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \in \mathcal{C}^{k-1} \mid (\alpha_u^k, \alpha_z^k, \alpha_\gamma^k, \alpha_\tau^k) \cdot (\mathbf{u}, \mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\tau}) \geq \beta^k \right\},$$

solve  $\text{LP}(\mathcal{C}_{\text{cur}}, \emptyset, \emptyset)$  to find an optimal solution  $(\mathbf{u}^k, \mathbf{z}^k, \boldsymbol{\gamma}^k, \boldsymbol{\tau}^k)$ .

**Step 4:** Let

$$\mathcal{J}_{\text{res}}^k = \left\{ j \in \{1, \dots, m\} \mid \tau_j^k (1 - \tau_j^k) > \epsilon \right\}.$$

If  $|\mathcal{J}_{\text{res}}^k| \leq |\mathcal{J}_{\text{res}}^{k-1}|$ , let  $\mathcal{C}^k := \mathcal{C}_{\text{cur}}$  and  $\mathcal{J} := \{1, \dots, m\}$ ; otherwise,

$$\mathcal{C}^k := \mathcal{C}^{k-1} \quad \text{and} \quad \mathcal{J} := \{1, \dots, m\} \setminus j_2.$$

**Step 5:** If the termination condition is satisfied, stop; otherwise, update  $k \leftarrow k + 1$ , and go to Step 1.

**Remark 5.2.** If  $\mathcal{J}_{\text{res}}^k = \emptyset$  at Step 4, then stop, because the current solution  $(\mathbf{u}^k, \mathbf{z}^k, \boldsymbol{\gamma}^k, \boldsymbol{\tau}^k)$  is a global optimal solution of the original problem (29). However, it often requires large computational time to solve (29) only by Algorithm 5.1. In practice, we restrict the maximum number of iterations as  $k \leq 1.8m$

and then employ Algorithm 4.2. The set of disjunctive cuts  $\mathcal{C}^k$  generated by Algorithm 5.1 plays a role in strengthening the LP relaxation problems solved in Algorithm 4.2.

**Remark 5.3.** At Step 1, as is done in Step 3 of Algorithm 4.2, we select the variable  $\tau_j$  with the largest residual of the complementarity condition (44). Then the variable  $\tau_{j_2}$  is used at Step 2 to define the disjunctive constraint.

**5.2. Simplification of cut generating LP.** In this section, we further analyze simplifications to the cut generating LP problem (51), motivated by the fact that in the dual problem (52), we can eliminate some of variables corresponding to the index set  $\bar{\mathcal{F}}$ . Recall that  $P_j(\mathcal{H})$  in the constraints of (52) has been defined in (46). Then we can consider the following points:

(a) For  $i$  such that  $\widehat{z}_i = 0$ :

Observe that  $(\alpha_z)_i$  does not contribute to the objective function of (51). By replacing the  $i$ -th rows of (48b) and (48h) with

$$\xi_i^{q+} + \xi_i^{q-} = \eta_i^{q+} + \eta_i^{q-}, \quad (53)$$

we can remove the variable  $(\alpha_z)_i$  from (48) without changing the optimal solution. At an optimal solution, we can complete  $(\bar{\alpha}_z)_i$  by letting

$$(\bar{\alpha}_z)_i := \bar{\xi}_i^{q+} + \bar{\xi}_i^{q-}, \quad (54)$$

where  $\bar{\xi}_i^{q+}$  and  $\bar{\xi}_i^{q-}$  are components of the optimal solution obtained. However, instead of (53), we append more restrictive constraints

$$\xi_i^{q+} = \eta_i^{q-}, \quad \xi_i^{q-} = \eta_i^{q+}$$

to (48), which enables us to remove the variables  $\eta_i^{q+}$  and  $\eta_i^{q-}$ . Then an optimal solution of the simplified problem can be completed to an optimal solution of (51) by using (54).

(b) For  $i$  such that  $\widehat{\tau}_i = 0$ :

In the system of (46), observe that (46a), (46e), (46j), and  $\widehat{\tau}_i = 0$  imply

$$w_i^\tau = y_i^\tau = 0. \quad (55)$$

(i) Assume that there exists an  $l \in \{1, \dots, m\}$  such that  $\widehat{\tau}_l \neq 0$ . Then, in the system (46), (55) and the  $l$ -th row of (46e) make the constraint

$$w_i^\tau \leq w_0 \quad (56)$$

redundant. Similarly, it follows from the  $l$ -th row of (46j) and (55) that the constraint

$$y_i^\tau \leq y_0 \quad (57)$$

is redundant. Then, we see that eliminating (56) and (57) from (46) is equivalent to eliminating the variables  $\xi_i^{\rho-}$  and  $\eta_i^{\rho-}$  from (48).

(ii) In (46), it follows from (55) that the  $j$ -th rows of (46d) can be replaced with

$$t_i^\top w^u - w_i^\gamma \geq -Mw_0, \quad -t_i^\top w^u - w_i^\gamma \geq 0$$



without changing  $P_j(\mathcal{K})$ . Then, in (48), the  $i$ -th row of (48d) is replaced with

$$(\boldsymbol{\alpha}^\tau)_i = \xi_i^{\rho+} - \xi_0 e_i^j. \quad (58)$$

Similarly, the  $i$ -th row of (48j) is replaced with

$$(\boldsymbol{\alpha}^\tau)_i = \eta_i^{\rho+} + \eta_0 e_i^j. \quad (59)$$

Note that, in (51), the variables  $\xi_i^{\rho+}$  and  $\eta_i^{\rho+}$  appear only in the constraints (58) and (59), respectively. Moreover,  $(\boldsymbol{\alpha}^\tau)_i$  does not contribute to the objective function. Consequently, from (51), we can eliminate the variables  $(\boldsymbol{\alpha}^\tau)_i$ ,  $\xi_i^{\rho+}$ , and  $\eta_i^{\rho+}$  and the constraints (58) and (59). From the nonnegativity of  $\xi_i^{\rho+}$  and  $\eta_i^{\rho+}$  it follows that an optimal solution of the simplified problem can be completed to an optimal solution of (51) by defining the eliminated variables as

$$(\bar{\boldsymbol{\alpha}}^\tau)_i := \max \left\{ -\bar{\xi}_0 e_i^j, \bar{\eta}_0 e_i^j \right\}.$$

The LP problem (51) originally has  $(n^d + 5n^m + 10m + 3)$  variables and  $(2n^d + 2n^m + 4m + 2)$  linear equality constraints besides side constraints. However, we shall show in the numerical examples in Section 6.2 that the simplification proposed in this section greatly reduces the numbers both of variables and constraints. The size of the simplified problem depends on  $(\hat{\boldsymbol{u}}, \hat{\boldsymbol{z}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{\tau}})$ , and hence differs at each iteration of Algorithm 5.1.

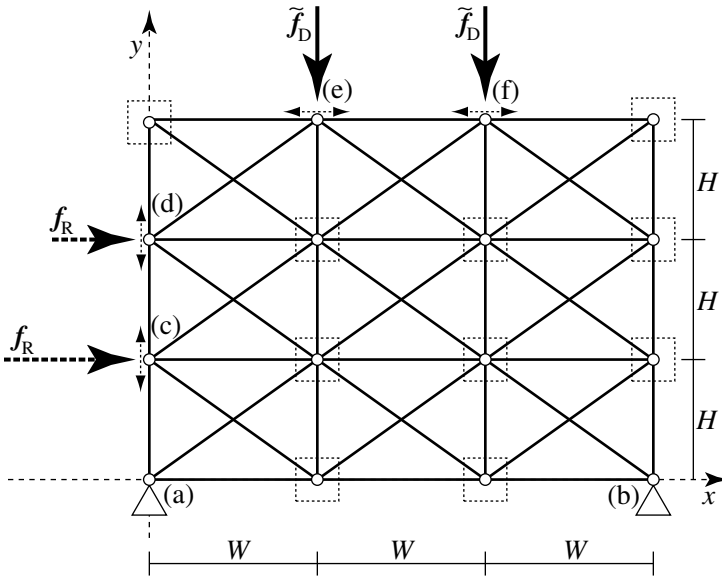
## 6. Numerical experiments

The worst-case limit load factors are computed for trusses by using Algorithms 4.2 and 5.1. Computation was carried out on a Pentium M (1.5 GHz with 1 GB memory) with MATLAB Version 6.5.1. The LP problems are solved by using the simplex method at Step 1 of Algorithm 4.2 and at Steps 2 and 3 of Algorithm 5.1. As an implementation of the simplex method, we use MATLAB built-in function `linprog` of Optimization Toolbox, Version 2.1 [MATLAB 2000], with the options ‘LargeScale’ set to ‘off’, and ‘Simplex’ set to ‘on’.

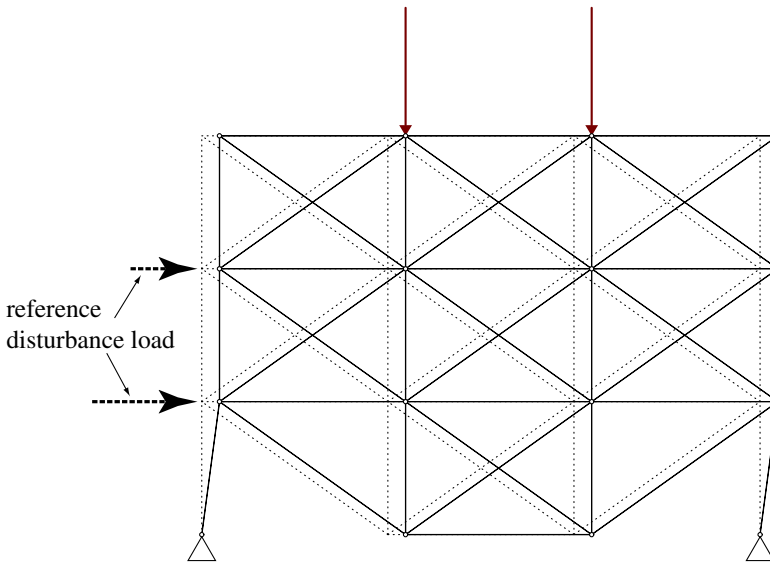
In the following examples, the yield stress is  $\sigma_i^y = 400$  MPa and cross-sectional area is  $a_i = 20.0$  cm<sup>2</sup> for each member. We set  $M = 5.0$  in Algorithm 4.2 and Algorithm 5.1.

**6.1. 3 × 3 truss.** Consider a plane truss illustrated in Figure 1, where  $W = 70.0$  cm,  $H = 50.0$  cm,  $n^d = 28$ , and  $n^m = 42$ . The nodes (a) and (b) are pin-supported.

As the nominal dead load  $\tilde{\boldsymbol{f}}_D$ , we apply the external forces  $(0, -120.0)$  kN at the nodes (e) and (f) as shown in Figure 1. Note again that  $\boldsymbol{f}_D$  represents the sum of conventional live load and dead load. It is possible that the conventional live load has moderately large magnitude of uncertain variation. In this case, the level of uncertainty of  $\boldsymbol{f}_D$ , that is,  $\alpha$  in (9), is supposed to be a moderately large value. The reference disturbance load  $\boldsymbol{f}_R$  is defined such that  $(40.0, 0)$  kN and  $(20.0, 0)$  kN, respectively, are applied at the nodes (c) and (d). The limit load factor under the nominal dead loads is computed as  $\lambda^*(\tilde{\boldsymbol{f}}_D) = 48.4$  by employing the usual limit analysis, that is, by solving the LP problem (5). The collapse mode corresponds to the sway-type with horizontal displacements of the joints shown in Figure 2, where the vanishing members experience plastic deformations.



**Figure 1.** 3 × 3 plane grid truss.



**Figure 2.** Collapse mode and the dead load of the 3 × 3 truss without the uncertainty in dead load ( $\lambda^*(\tilde{f}_D) = 48.4$ ).

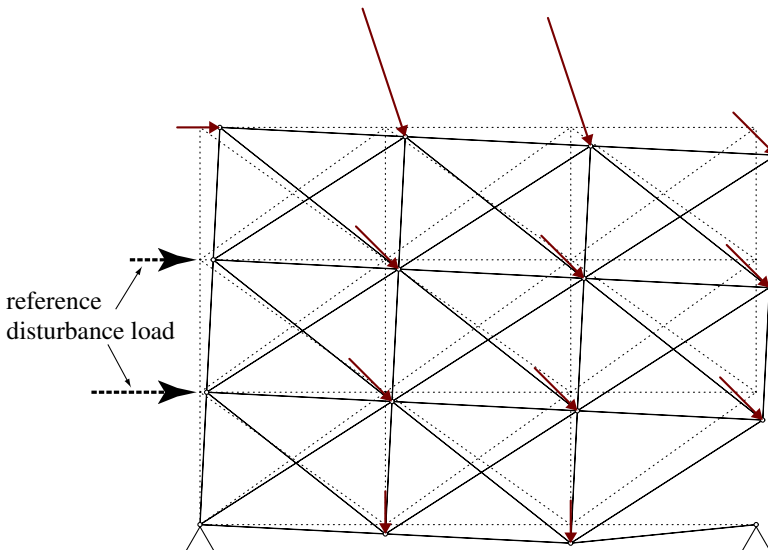
We assume that the uncertain load  $T\zeta$  can possibly exist at free nodes, so that Assumption 3.1 is satisfied and the condition

$$(T\zeta)^\top \tilde{f}_D = 0, \quad \zeta \in \mathbb{R}^m \tag{60}$$

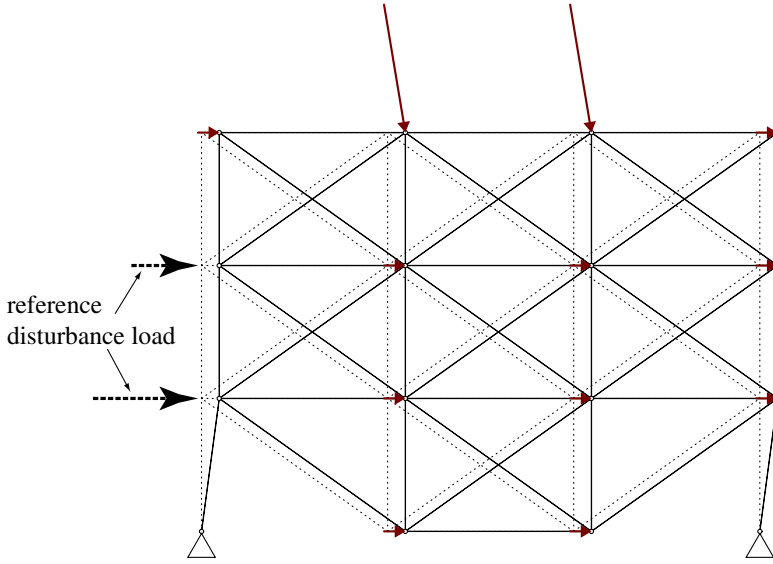
holds with  $m = 24$ . Accordingly, the uncertain load  $T\zeta$  is running through the squares and arrows depicted with the dotted lines in Figure 1. For  $\alpha_1 = 40.0$  kN, the worst-case limit load factor is computed as  $\lambda_{\min}(\alpha_1) = 37.0$  by using Algorithms 4.2 and 5.1. Let  $\zeta_1^{cr}$  denote the optimal solution of (13). The corresponding critical load  $f_D(\zeta_1^{cr})$  and collapse mode are shown in Figure 3.

Figure 3 shows that the collapse mode in the worst case is different from the sway-type mode observed in the nominal case of Figure 2. On the contrary, for  $\alpha_2 = 20.0$  kN, the collapse mode in the worst case coincides with the sway-type as illustrated in Figure 4. The corresponding worst-case limit load factor is  $\lambda_{\min}(\alpha_2) = 44.4$ . The distribution of critical load  $f_D(\zeta_2^{cr})$  is shown in Figure 4, which is different from the critical load in the case of Figure 3.

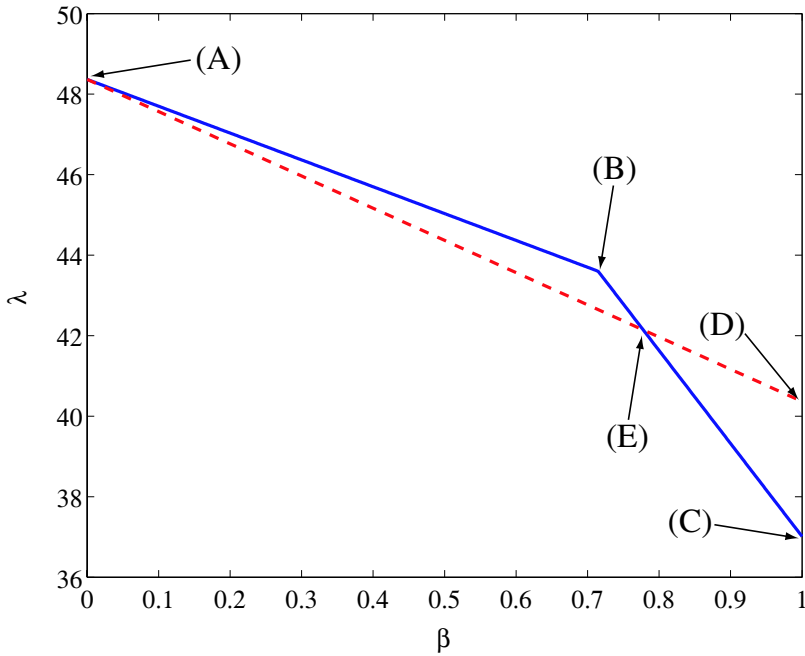
We next investigate the variation of the limit load factor by proportionally increasing the uncertain dead load, that is, we employ the usual limit analyses repeatedly by putting  $\zeta = \beta\zeta_1^{cr}$  and gradually increasing  $\beta$ . In Figure 5, the solid curve (A)→(B)→(C) shows the variation of  $\lambda^*(f_D(\beta\zeta_1^{cr}))$  with respect to  $\beta$ . The collapse mode coincides with the sway-type shown in Figure 2 between the points (A) and (B), while the mode of Figure 3 is observed between (B) and (C). The variation of  $\lambda^*(f_D(2\beta\zeta_2^{cr}))$  with respect to  $\beta$  is indicated by the dashed line (A)→(D) in Figure 5. Note that  $\alpha_1 = 2\alpha_2$  implies  $\|\zeta_1^{cr}\| = 2\|\zeta_2^{cr}\|$ . The collapse mode coincides with the sway type shown in Figure 4 between (A) and (D). The curve (A)→(E)→(C) corresponds to the variation of the worst-case limit load factor  $\lambda_{\min}(\beta\alpha_1)$  with respect to  $\beta$ . This illustrates that the critical loads as well as the corresponding collapse modes depend on the level of uncertainty  $\alpha$ .



**Figure 3.** Collapse mode and the critical load of the  $3 \times 3$  truss in the worst case for  $\alpha_1 = 40.0$  kN ( $\lambda_{\min}(\alpha_1) = 37.0$ ).



**Figure 4.** Collapse mode and the critical load of the  $3 \times 3$  truss in the worst case for  $\alpha_2 = 20.0 \text{ kN}$  ( $\lambda_{\min}(\alpha_2) = 44.4$ ).



**Figure 5.** Limit load factor of the  $3 \times 3$  truss;  $\lambda^*(f_D(\beta \xi_1))$ : solid line;  $\lambda^*(f_D(2\beta \xi_2))$ : dashed line.

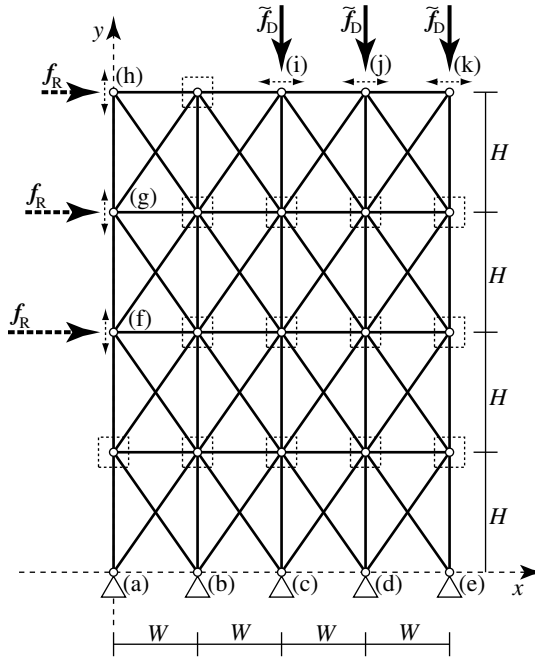
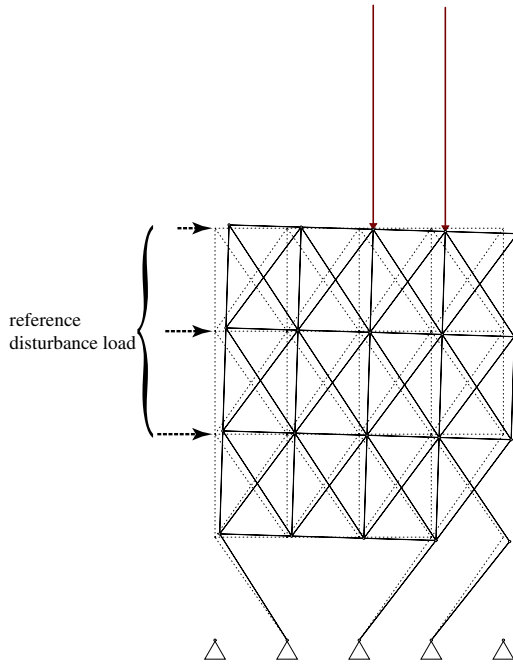


Figure 6.  $4 \times 4$  plane truss.

**6.2.  $4 \times 4$  truss.** Consider the 68-bar plane truss illustrated in Figure 6, where  $n^m = 68$ ,  $n^d = 40$ ,  $W = 35.0$  cm, and  $H = 50.0$  cm. The nodes (a)–(e) are pin-supported. As the nominal dead load  $\tilde{f}_D$ , we apply the external forces  $(0, -800.0)$  kN at the nodes (i)–(k). The reference disturbance load  $f_R$  is defined such that  $(52.0, 0)$  kN,  $(40.0, 0)$  kN, and  $(28.0, 0)$  kN are applied at the nodes (f), (g), and (h), respectively. The nominal limit load factor is computed as  $\lambda^*(\tilde{f}_D) = 14.3$  by employing the usual limit analysis. The corresponding collapse mode is shown in Figure 7, where the vanishing members experience plastic deformations.

In a manner similar to Section 6.1, we assume that the uncertain load  $T\zeta$  can possibly exist at all free nodes so that Assumption 3.1 and (60) are satisfied with  $m = 34$ . Accordingly, the uncertain load  $T\zeta$  is running through the squares and arrows depicted with the dotted lines in Figure 6. We set  $\alpha = 40.0$  kN. By using Algorithms 4.2 and 5.1, the worst-case limit load factor is computed as  $\lambda_{\min}(\alpha) = 7.73$ . The CPU time required by Algorithm 5.1 is 229.5 sec, and 42 cutting planes are generated within 61 iterations. Afterward, Algorithm 4.2 terminates by solving only 9 LP problems, where the CPU time required is 28.5 sec. This result demonstrates that the generated cutting planes at the root node of the branch-and-bound tree can reduce the number of nodes drastically that have to be visited in Algorithm 4.2.

In Step 2 of Algorithm 5.1, the LP problem (51) originally has 723 variables and 354 linear equality constraints besides side constraints in this example. By using the simplification proposed in Section 5.2, the problems to be solved have 534 variables and 240 linear equality constraints in average. Thus, the simplification scheme can reduce the numbers both of variables and constraints drastically. In Step 1 of Algorithm 4.2, the LP problem (45) originally has 315 variables and 176 linear equality constraints



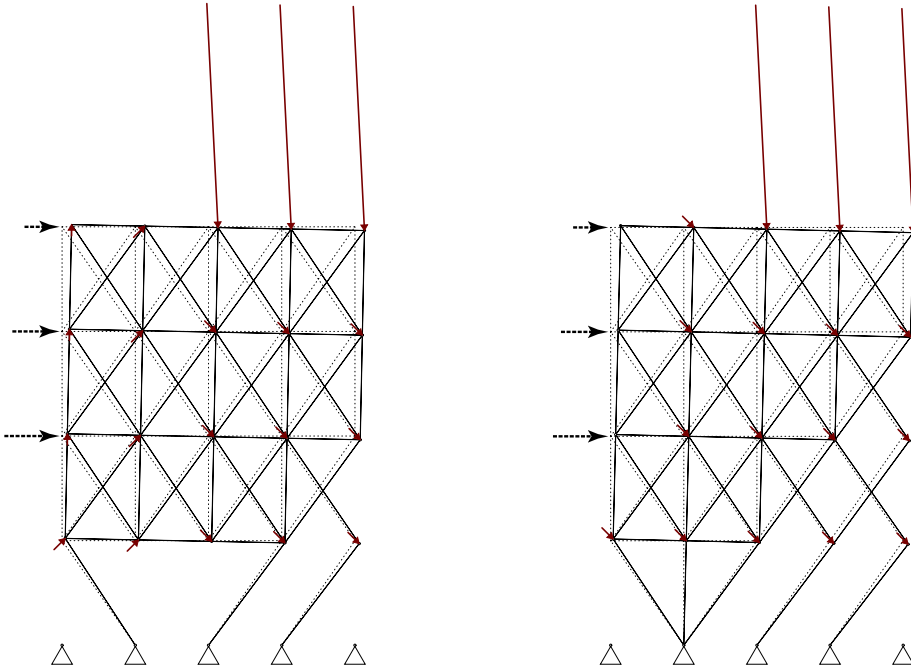
**Figure 7.** Collapse mode and the dead load of the  $4 \times 4$  truss without the uncertainty in dead load ( $\lambda^*(\tilde{\mathbf{f}}_D) = 14.3$ ).

besides side constraints, because  $n^c = 42$ . After applying the simplification proposed in Section 4.3, the problems to be solved have 309 variables and 174 linear equality constraints in average.

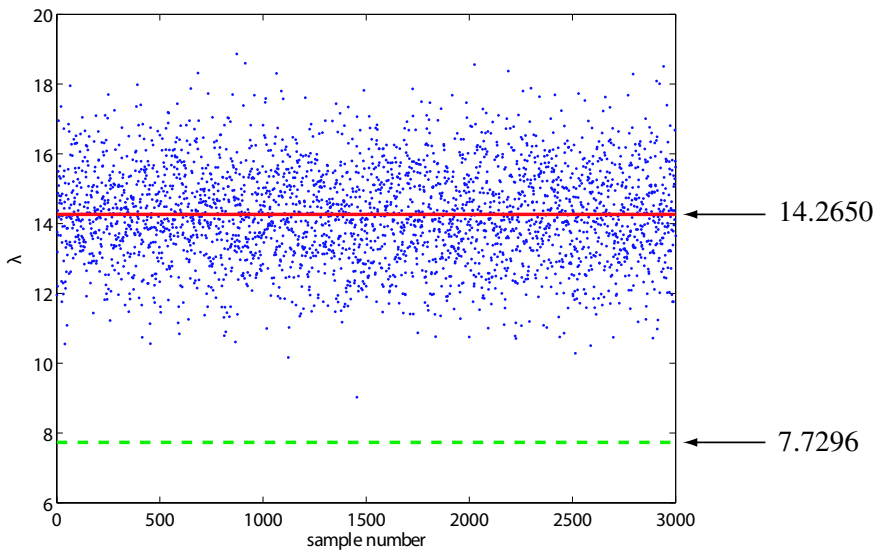
Note that the worst-case limit load factor is almost half of the nominal one, in spite of the fact that the level of uncertainty  $\alpha$  is relatively small compared with the norm of the nominal dead loads vector  $\tilde{\mathbf{f}}_D$ . The critical load  $\mathbf{f}_D(\boldsymbol{\zeta}^{cr})$  and the corresponding collapse mode are shown in Figure 8 (left). It is observed from Figure 8 (left) that the collapse mode in the worst case is the same as that in the nominal case illustrated in Figure 7.

For comparison, we select a sample of the uncertain parameters vector  $\boldsymbol{\zeta}'$  satisfying  $\|\boldsymbol{\zeta}'\|_\infty = \alpha$  as the nodal forces shown in Figure 8 (right). The corresponding limit load factor is  $\lambda^*(\mathbf{f}_D(\boldsymbol{\zeta}')) = 8.43$ , which is larger than the worst case. The corresponding collapse mode is shown in Figure 8 (right), which is different from the mode shown in Figure 8 (left). Thus, it is not easy to find the critical loads vector in a heuristic way.

We randomly generate a number of  $\boldsymbol{\zeta}$  satisfying  $\|\boldsymbol{\zeta}\|_\infty = \alpha$ , and perform conventional limit analyses. The limit load factors  $\lambda^*(\mathbf{f}_D(\boldsymbol{\zeta}))$  obtained are plotted in Figure 9. We observe that all generated  $\lambda^*(\boldsymbol{\zeta})$  are larger than the worst-case limit load factor  $\lambda^*(\boldsymbol{\zeta}^{cr}) = \lambda_{\min}(\alpha)$ . This supports the assertion that using our proposed algorithms yields a global optimal solution of the nonconvex problem (13).



**Figure 8.** Left: collapse mode and the critical load of the  $4 \times 4$  truss in the worst case for  $\alpha = 40.0$  kN ( $\lambda_{\min}(\alpha) = 7.73$ ). Right: definition of the dead load with  $\zeta'$  and the corresponding collapse mode of the  $4 \times 4$  truss ( $\lambda^*(f_D(\zeta')) = 8.43$ ).



**Figure 9.** Limit load factor for randomly generated  $\zeta$ : (solid line)  $\lambda^*(\tilde{f}_D)$ ; (dashed line)  $\lambda_{\min}(\alpha)$ .

## 7. Conclusions

In this paper, we have investigated the worst-case detection in the plastic limit analysis of trusses affected by unknown-but-bounded dead and live loads. While the imprecisely known dead and live loads are constrained into a given bounded set, the live or disturbance loads are amplified with the load factor. A global optimization technique has been presented to compute the worst-case limit load factor as well as the critical load.

We supposed that the dead and live loads applied to a truss contain bounded errors around the nominal values. The level, or ‘width’, of uncertain variation is assumed to be known. We defined the worst case limit load factor as the minimum value among limit load factors attained with some loading patterns belonging to a given closed set. Then the worst-case detection problem has been formulated as a mixed 0-1 programming problem. To obtain a global optimal solution of the present problem, we have proposed a cut-and-branch method based on the LP relaxation and the disjunctive cut, where a cutting plane is generated by solving another LP problem. Since the proposed method converges to a global optimal solution, it is theoretically guaranteed that there exists no uncertain parameter such that the limit load factor becomes smaller than the obtained optimal value.

We showed in the numerical examples that the proposed cut-and-branch method can find the worst-case limit load factors. The comparison with the limit load factors for randomly generated dead and live loads demonstrates that the limit load factors we obtained correspond to the global optimal solutions of the mixed 0-1 programming problem presented. We have also illustrated through numerical examples that the process of cutting plane generation at the root node of the enumeration tree can reduce the number of LP relaxation problems that should be solved in the successive branch-and-bound procedure, although no theoretical result is to date available that suggests how many cutting planes should be generated.

## References

- [Alefeld and Mayer 2000] G. Alefeld and G. Mayer, “Interval analysis: theory and applications”, *J. Comput. Appl. Math.* **121**:1–2 (2000), 421–464. MR 2001d:65068 Zbl 0995.65056
- [Andersen et al. 1998] K. D. Andersen, E. Christiansen, and M. L. Overton, “Computing limit loads by minimizing a sum of norms”, *SIAM J. Sci. Comput.* **19**:3 (1998), 1046–1062. MR 99a:73059 Zbl 0924.73074
- [Balas and Perregaard 2002] E. Balas and M. Perregaard, “Lift-and-project for mixed 0-1 programming: recent progress”, *Discrete Appl. Math.* **123**:1–3 (2002), 129–154. MR 2004c:90023 Zbl 1076.90031
- [Balas et al. 1996] E. Balas, S. Ceria, and G. Cornuéjols, “Mixed 0-1 programming by lift-and-project in a branch-and-cut framework”, *Manage. Sci.* **42**:9 (1996), 1229–1246.
- [Ben-Haim 2001] Y. Ben-Haim, *Information-gap decision theory: decisions under severe uncertainty*, Series on Decision and Risk, Academic Press, San Diego, CA, 2001. MR 2003f:91001 Zbl 0985.91013
- [Ben-Haim and Elishakoff 1990] Y. Ben-Haim and I. Elishakoff, *Convex models of uncertainty in applied mechanics*, Studies in Applied Mechanics **25**, Elsevier, Amsterdam, 1990. MR 94g:73027 Zbl 0703.73100
- [Ben-Tal and Nemirovski 1997] A. Ben-Tal and A. Nemirovski, “Robust truss topology design via semidefinite programming”, *SIAM J. Optim.* **7**:4 (1997), 991–1016. MR 98m:73068 Zbl 0899.90133
- [Ben-Tal and Nemirovski 2002] A. Ben-Tal and A. Nemirovski, “Robust optimization: methodology and applications”, *Math. Program.* **92**:3, Ser. B (2002), 453–480. MR 2003d:90074 Zbl 1007.90047
- [Ben-Tal et al. 2004] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski, “Adjustable robust solutions of uncertain linear programs”, *Math. Program.* **99**:2, Ser. A (2004), 351–376. MR 2004j:90057 Zbl 1089.90037



- [Ceria and Soares 1997] S. Ceria and J. Soares, “Disjunctive cut generation for mixed 0-1 programs: duality and lifting”, Computational Operations Research Center, Columbia University, New York, 1997.
- [Chen et al. 2002] S. Chen, H. Lian, and X. Yang, “Interval static displacement analysis for structures with interval parameters”, *Int. J. Numer. Methods Eng.* **53**:2 (2002), 393–407. Zbl pre01742227
- [Cheng et al. 2004] Y. S. Cheng, F. T. K. Au, L. G. Tham, and G. W. Zeng, “Optimal and robust design of docking blocks with uncertainty”, *Eng. Struct.* **26**:4 (2004), 499–510.
- [Chvátal 1983] V. Chvátal, *Linear programming*, W. H. Freeman, New York, 1983. MR 86g:90062 Zbl 0537.90067
- [Clausen and Perregaard 1999] J. Clausen and M. Perregaard, “On the best search strategy in parallel branch-and-bound: best-first search versus lazy depth-first search”, *Ann. Oper. Res.* **90** (1999), 1–17. Zbl 0937.90089
- [Cocchetti and Maier 2003] G. Cocchetti and G. Maier, “Elastic-plastic and limit-state analyses of frames with softening plastic-hinge models by mathematical programming”, *Int. J. Solids Struct.* **40**:25 (2003), 7219–7244. Zbl 1065.74039
- [Cordier et al. 1999] C. Cordier, H. Marchand, R. Laundry, and L. A. Wolsey, “Bc-opt: a branch-and-cut code for mixed integer programs”, *Math. Program.* **86**:2, Ser. A (1999), 335–353. MR 2000h:90052 Zbl 0939.90025
- [Craig et al. 2003] K. J. Craig, N. Stander, and S. Balasubramanyam, “Worst-case design in head impact crashworthiness optimization”, *Int. J. Numer. Methods Eng.* **57**:6 (2003), 795–817.
- [Elishakoff et al. 1994] I. Elishakoff, R. T. Haftka, and J. Fang, “Structural design under bounded uncertainty: optimization with anti-optimization”, *Comput. Struct.* **53**:6 (1994), 1401–1405.
- [Ganzerli and Pantelides 1999] S. Ganzerli and C. P. Pantelides, “Load and resistance convex models for optimum design”, *Struct. Multidiscip. O.* **17**:4 (1999), 259–268.
- [Gu et al. 2000] X. Gu, J. E. Renaud, S. M. Batill, R. M. Brach, and A. S. Budhiraja, “Worst-case propagated uncertainty of multidisciplinary systems in robust design optimization”, *Struct. Multidiscip. O.* **20**:3 (2000), 190–213.
- [Hodge 1959] P. G. Hodge, Jr., *Plastic analysis of structures*, McGraw-Hill, New York, 1959. MR 22 #4237 Zbl 0113.17903
- [Kanno and Takewaki 2006a] Y. Kanno and I. Takewaki, “Robustness analysis of trusses with separable load and structural uncertainties”, *Int. J. Solids Struct.* **43**:9 (2006), 2646–2669. MR 2006j:74050
- [Kanno and Takewaki 2006b] Y. Kanno and I. Takewaki, “Sequential semidefinite program for maximum robustness design of structures under load uncertainty”, *J. Optim. Theory Appl.* **130**:2 (2006), 265–287.
- [Kharmanda et al. 2004] G. Kharmanda, N. Olhoff, A. Mohamed, and M. Lemaire, “Reliability-based topology optimization”, *Struct. Multidiscip. O.* **26**:5 (2004), 295–307.
- [Krabbenhoft and Damkilde 2003] K. Krabbenhoft and L. Damkilde, “A general non-linear optimization algorithm for lower bound limit analysis”, *Int. J. Numer. Methods Eng.* **56**:2 (2003), 165–184.
- [Lloyd Smith et al. 1990] D. Lloyd Smith, P.-H. Chuang, and J. Munro, “Fuzzy linear programming in plastic limit design”, pp. 425–435 in *Mathematical programming methods in structural plasticity*, edited by D. Lloyd Smith, CISM Courses and Lectures **299**, Springer, Vienna, 1990.
- [Lyamin et al. 2005] A. V. Lyamin, S. W. Sloan, K. Krabbenhoft, and M. Hjiaj, “Lower bound limit analysis with adaptive remeshing”, *Int. J. Numer. Methods Eng.* **63**:14 (2005), 1961–1974.
- [Marchand et al. 2002] H. Marchand, A. Martin, R. Weismantel, and L. Wolsey, “Cutting planes in integer and mixed integer programming”, *Discrete Appl. Math.* **123**:1–3 (2002), 397–446. MR 2003g:90045 Zbl pre01801739
- [Marti and Stoeckel 2004] K. Marti and G. Stoeckel, “Stochastic linear programming methods in limit load analysis and optimal plastic design under stochastic uncertainty”, *Z. Angew. Math. Mech.* **84**:10 (2004), 666–677.
- [MATLAB 2000] *Optimization toolbox user’s guide*, The MathWorks, Natick, MA, 2000.
- [MATLAB 2002] *Using MATLAB*, The MathWorks, Natick, MA, 2002.
- [Muralidhar and Jagannatha Rao 1997] R. Muralidhar and J. R. Jagannatha Rao, “New models for optimal truss topology in limit design based on unified elastic/plastic analysis”, *Comput. Methods Appl. Mech. Eng.* **140**:1–2 (1997), 109–138. MR 97i:73078 Zbl 0898.73046
- [Rocho and Sonnenberg 2003] P. Rocho and S. Sonnenberg, “Limit analysis of frames: application to structural reliability”, pp. 269–282 in *Numerical methods for limit and shakedown analysis*, edited by M. Staat and M. Heitzer, NIC **15**, John von Neumann Institute for Computing, Jülich, 2003.

- [Staat and Heitzer 2003] M. Staat and M. Heitzer, “Probabilistic limit and shakedown problems”, pp. 217–268 in *Numerical methods for limit and shakedown analysis*, edited by M. Staat and M. Heitzer, NIC 15, John von Neumann Institute for Computing, Jülich, 2003.
- [Takewaki and Ben-Haim 2005] I. Takewaki and Y. Ben-Haim, “Info-gap robust design with load and model uncertainties”, *J. Sound Vib.* **288**:3 (2005), 551–570.
- [Tin-Loi and Ngo 2003] F. Tin-Loi and N. S. Ngo, “Performance of the  $p$ -version finite element method for limit analysis”, *Int. J. Mech. Sci.* **45**:6–7 (2003), 1149–1166.
- [Zang et al. 2005] C. Zang, M. I. Friswell, and J. E. Mottershead, “A review of robust optimal design and its application in dynamics”, *Comput. Struct.* **83**:4–5 (2005), 315–326.

Received 30 Jun 2006. Accepted 23 Sep 2006.

YOSHIHIRO KANNO: [kanno@mist.i.u-tokyo.ac.jp](mailto:kanno@mist.i.u-tokyo.ac.jp)

*Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan*

<http://www.simplex.t.u-tokyo.ac.jp>

IZURU TAKEWAKI: [takewaki@archi.kyoto-u.ac.jp](mailto:takewaki@archi.kyoto-u.ac.jp)

*Department of Urban and Environmental Engineering, Graduate School of Engineering, Kyoto University, Kyotodaigaku-Katsura, Nishikyo, Kyoto 615-8540, Japan*

<http://www.archi.kyoto-u.ac.jp/%7ebge/tkww/takewaki-new.html>

## SHAPE OPTIMIZATION OF A RIGID INCLUSION IN A SHEAR-LOADED ELASTIC PLANE

SHMUEL VIGDERGAUZ

A rigid inclusion perfectly embedded in a thin plate is considered as a two-dimensional elastostatic composite structure to solve the inverse problem of finding the inclusion shape around which the local maximum of the von Mises equivalent stresses attain the global minimum under shear loading at infinity. Absent optimality preconditions such as the equistress principle for bulk-type loading, a fast and accurate assessment of a given shape is developed by combining complex-valued series expansions with a new infinite summation scheme. This approach to solving the direct problem is then included into a genetic algorithm optimization over the set of shapes obtained from a circle by a finite-term conformal mapping with square symmetry. Compared to a circular inclusion, the stresses may thus be lowered by 15–25%, depending on the Poisson's ratio of the plate. The numerical results presented allow us to conjecture that the von Mises stresses around the optimal shape are *uniform*. The inclusion that minimizes the induced energy increment is also identified by the same approach. Both shapes appear to be very similar, though not identical.

### 1. Introduction

Grained and fibre-reinforced composites find broad use as construction elements. For this reason, their mechanical behavior is the subject of extensive study in elasticity. Such two-phase structures are idealized here as an infinite zero-thickness plate with perfectly bonded foreign inclusions, which may represent either a two-dimensional grain or the cross-section of a long cylindrical fibre. These mathematically equivalent cases are known as *plane stress* and *plane strain*, respectively [Muskhelishvili 1975].

We further simplify the model by assuming that the inclusions are remote from each other, so that one may neglect their interaction and consider a single inclusion in an unbounded matrix under a homogeneous elastic field  $\sigma_{xx}^\infty, \sigma_{yy}^\infty, \sigma_{xy} = 0$  at infinity.

At finite distances, this field is distorted by the inclusion. With the layout above, the resulting stresses depend only on the inclusion geometry. This brings us to the optimization problem of finding the inclusion shape that minimizes a certain criterion. Two practically important criteria are commonly used in shape optimization: the inclusion-induced strain energy increment  $\delta W$  divided by the inclusion area, and the maximum  $\mathcal{M}$  of von Mises equivalent stresses  $\sigma_M$  [Cherepanov 1997] along the material interface. On a free-traction hole,  $\sigma_M$  coincide with the hoop stresses, but on a rigid or elastic inclusion they involve bilinearly all the boundary stresses which, in a general way, should differ from zero to ensure the assumed perfect contact between the phases.

The exact formulation of the problem reads:

---

*Keywords:* plane elasticity problem, Kolosov–Muskhelishvili potentials, shape optimization, effective energy, extremal elastic structures, genetic algorithm.

*Given the topology of the single inclusion in the matrix, the uniform static far loading and the elastic properties of the matrix, find, among all continuous curves, the rigid inclusion shape that minimizes either of the two above-mentioned criteria.*

Beginning with the pioneering work of Cherepanov [1974], it was realized that the sign of the ratio  $\beta = \sigma_{yy}^\infty / \sigma_{xx}^\infty$  is the main factor driving the complexity of this shape optimization problem. As applied to rigid inclusions, two basic cases are of particular interest:

- *The far loading is hydrostatic tension or compression ( $\beta = 1$ ).* Here, attention is focused on the equistress contours [Cherepanov 1974; Vigdergauz 1976; Grabovsky and Kohn 1994], at which matrix-side boundary tractions reduce to uniform normal stresses. Remarkably, these contours not only exist for a variety of geometries but also provide the global minima for both criteria simultaneously. The single equistress inclusion is simply a circle. For interacting inclusions the equistress condition is used as a prerequisite to find their shapes by analytical or numerical solving of the coupled Dirichlet problem; see, for instance, [Grabovsky and Kohn 1994]. In other words, the optima are found without touching the direct problem of solving the stress field around any given inclusion.

- *The far load is that of pure shear ( $\beta = -1$ ).* In this context, neither  $\delta W$ - nor  $\mathcal{M}$ -optimal shapes are known so far, because the equistress principle is no longer valid and no other optimal precondition has yet been developed instead. For this reason, the more effective strategy is the standard numerical one, where a search space of admissible shapes is first defined through a finite number of parameters, which are then optimized by repetitively solving the criterion function over many curves.

As parameters, we choose the first  $N$  Laurent coefficients  $d_k, k = \overline{1, N}, d_k = 0, k > N$  of the conformal map from the exterior of the unit circle to the exterior of a given inclusion. The adoption of these coefficients not only allows the use the whole machinery of complex variable theory, but they are also naturally ordered, in the sense that the higher the coefficient, the less its global impact on the inclusion shape. Owing to this, the finite-term mapping approach of [Muskhelishvili 1975], or its revised version (see [Kalandia 1975]), performs well in the numerical treatment of the direct problem, as demonstrated in the classical book by Savin [1961] and more recently in [Jasiuk 1995].

However, such solving of the local stresses involves sums of Fourier series that are actually infinite at any  $N$ , and hence may imply an unacceptably large  $N$  to attain a reasonable accuracy in optimization practice. The finite differences scheme proposed by us in the similar context of hole shape optimization [Vigdergauz 2006] overcomes this drawback by computing the infinite sums involved *exactly*, as rational functions of  $\{d_k\}, k \leq N$ . This can be done with the proviso that the first  $N$  items of the sums are solved from the  $N \times N$  system of linear algebraic equations expressing the matrix-inclusion contact conditions. In fact, just this novelty makes the considered optimization problem numerically tractable. A further computational saving can be achieved due to an assumed rotational symmetry of the inclusion by which the set  $\{d_k\}$  partially vanishes. For the most part, we treat here the case of square symmetry, which appears to provide the best optimal response like for a shear-loaded plate with a hole [Vigdergauz 2006].

After all these considerations, the problem to be solved is a moderate-size problem of minimizing a rational function of the arguments  $d_k, k = \overline{1, N}$ , subject to the two-side inequalities of mapping uniqueness [Schinzinger and Laura 1991]. This nonlinear optimization problem with linear constraints is nevertheless a computationally expensive task, which is solved by a genetic algorithm (GA), an optimization

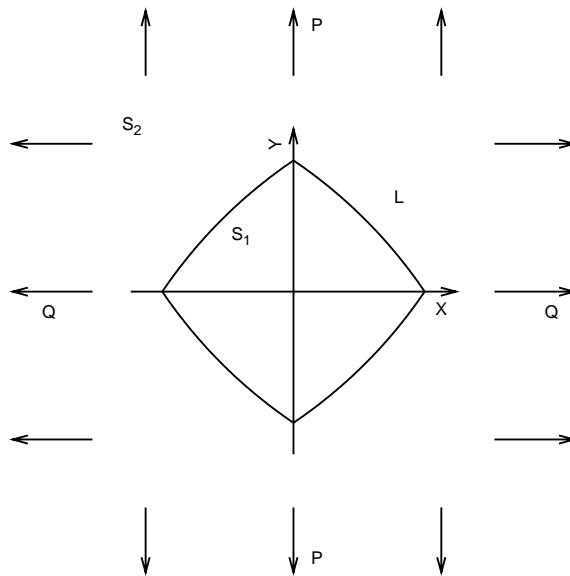
method imitating adaptive schemes found in nature (see [Gen and Cheng 1997; Osyczka 2001] and references therein). GAs have a twofold benefit over gradient-based schemes: the latter find a local optimum, while a GA is a global optimization algorithm; and no differentiation with respect to the optimized parameters is needed at any stage.

The main results so obtained show that the  $\delta W$ - and  $\mathcal{M}$ -optimal square symmetric rigid inclusions are almost the same except for a hardly compressible or incompressible matrix material, and that their shape is close to a slightly rounded square rotated by 45 degrees with respect to the main stresses  $\sigma_{xx}^\infty, \sigma_{yy}^\infty$ .

The paper is structured as follows. Section 2 covers some basics of conformal mapping in two-dimensional elasticity, together with new algebraic identities derived for later use. Section 3 describes the summation scheme for the boundary stresses in the direct problem which is the basis for further considerations. After these lengthy preliminaries, the reformulated optimization problem is rather straightforward and simple. First, the one-term mapping approximation is derived analytically (Section 4) to obtain qualitative properties of the solution. The multiterm approximations are next found by the standard GA briefly described in Section 5. Numerical results are displayed and discussed in Section 6. Finally, conclusions and proposed extensions of the work can be found in Section 7.

## 2. Complex variable technique in plane elastostatics

**2.1. The local stress-strain relations.** Consider the setup in Figure 1. Let an infinite elastic plane  $E$  be strengthened by a perfectly bonded rigid inclusion with a piecewise smooth boundary  $\Gamma$  enclosing the origin of the Cartesian system  $xOy$ . The curve  $\Gamma$  divides the plane into the inclusion region  $S_1$ , of finite area  $f_1$ , and the unbounded region  $S_2 = E \setminus S_1$ , filled with a linearly elastic phase. Let the plane be



**Figure 1.** The problem schematic: an infinite plate with a perfectly bonded inclusion under uniform stresses, the cases  $P = Q$  and  $P = -Q$  correspond to remote bulk and shear, respectively. The piecewise smooth inclusion boundary possesses a certain rotational symmetry and may have a finite number of angular points.

remotely loaded by uniform nontangential stresses

$$\sigma_{xx}^\infty = P, \quad \sigma_{yy}^\infty = Q, \quad \sigma_{xy}^\infty = 0. \tag{1}$$

The load-induced stresses in  $S_2$  are governed by the biharmonic Airy function which, though useful as a theoretical tool, becomes ineffective in computing local displacements. Far more advantageous is the complex variable approach of replacing the Airy function with a pair of holomorphic functions (the KM potentials; see [England 1971; Muskhelishvili 1975, §§47–51]) linked along the material interface through the contact conditions. The resulting boundary value problem is then solved using complex variable techniques. For holes and rigid inclusions the boundary conditions are further simplified by use of a conformal mapping  $z = \omega(\zeta)$  from the connected matrix domain  $S_2$  to the exterior  $\Sigma_2$  of a unit circle  $\gamma$ . The reason is that the following identities hold on  $\gamma$ , in contrast to any other shape:

$$\bar{\xi} = \xi^{-1}, \quad \int_{\gamma} \bar{\xi}^m \xi^n d\xi = \int_{\gamma} \xi^{n-m} d\xi = 2\pi i \delta_{m-n,1}, \quad \xi \in \gamma, \tag{2}$$

where  $\delta_{i,j}$  is the Kronecker delta. The mutual orthogonality of (2) makes the stress-strain computations much easier.

The KM potentials  $\Phi(\zeta)$ ,  $\Psi(\zeta)$  present the transformed stresses in  $\Sigma_2 + \gamma$  as [Muskhelishvili 1975]

$$\sigma_{\rho\rho}(\zeta) + \sigma_{\theta\theta}(\zeta) = 4 \operatorname{Re} \Phi(\zeta), \quad \zeta = \rho \exp(i\theta) : \zeta \in \Sigma_2 + \gamma, \rho \geq 1, \tag{3a}$$

$$\sigma_{\rho\rho}(\zeta) - i\sigma_{r\theta}(\zeta) = 2 \operatorname{Re} \Phi(\zeta) - \frac{\zeta^2}{\rho^2 \omega'(\zeta)} (\overline{\omega(\zeta)} \Phi'(\zeta) + \omega'(\zeta) \Psi(\zeta)). \tag{3b}$$

The zero-displacement condition at the rigid interface  $\gamma$  is [Muskhelishvili 1975]

$$\overline{\lambda \varphi(\xi)} - \overline{\omega(\xi)} \frac{\varphi'(\xi)}{\omega'(\xi)} - \psi(\xi) = 0; \quad \xi \in \gamma, \tag{4}$$

$$\lambda = \frac{3 - \nu_{\text{stress}}}{1 + \nu_{\text{stress}}}, \quad \nu_{\text{stress}} \in [0; 1] \quad \text{for plane stress,} \tag{5a}$$

$$\lambda = 3 - 4\nu_{\text{strain}}, \quad \nu_{\text{strain}} = \frac{\nu_{\text{stress}}}{1 + \nu_{\text{stress}}} \in [0.5; 1] \quad \text{for plane strain,} \tag{5b}$$

where

$$\varphi'(\zeta) \equiv \Phi(\zeta)\omega'(\zeta), \quad \psi'(\zeta) \equiv \Psi(\zeta)\omega'(\zeta). \tag{6}$$

To save room, nonobservable negative values of  $\nu$  will be excluded from the numerical simulations (Section 6). Then, both cases form the same interval  $1 \leq \lambda \leq 3$  for  $\lambda$ , which thus serves as a unified elastic parameter of the problem. The minimum value  $\lambda = 1$  corresponds to a rigid inclusion in a soft incompressible matrix.

Finally, integration of (3b) with the multiplier  $\overline{\xi^2 \omega'(\xi)} d\xi$  along  $\gamma$  results in the boundary stress condition [Muskhelishvili 1975] structurally similar to (4):

$$\overline{\varphi(\xi)} + \overline{\omega(\xi)} \frac{\varphi'(\xi)}{\omega'(\xi)} + \psi(\xi) = f(\xi), \quad \xi \in \gamma, \tag{7}$$

where the tractions  $f(\xi)$  take the form

$$f(\xi) = \int_{\gamma} (\sigma_{\rho\rho}(\xi) - i\sigma_{r\theta}(\xi)) \overline{\xi^2 \omega'(\xi)} d\xi. \quad (8)$$

**2.2. One-potential formulation of the forward rigid inclusion problem.** The homogeneous boundary condition (4) allows us to derive new relations which, though simple, have not, to our knowledge, been noticed in the literature so far. They give the stress components along  $\gamma$  through only one rather than both KM potentials, and are crucial for further analysis. To this end, we combine (4) and (7) to obtain

$$(\lambda + 1)\overline{\phi(\xi)} = f(\xi), \quad \xi \in \gamma. \quad (9)$$

Differentiating (9) in view of Equations (3a), (6), (8) yields for the real and imaginary parts of  $\Phi(\xi)$

$$\sigma_{\rho\rho}(\xi) = (\lambda + 1) \operatorname{Re} \Phi(\xi), \quad \sigma_{\theta\theta}(\xi) = (3 - \lambda) \operatorname{Re} \Phi(\xi), \quad \sigma_{\rho\theta}(\xi) = (\lambda + 1) \operatorname{Im} \Phi(\xi), \quad \xi \in \gamma. \quad (10)$$

At  $\lambda = -1$  (the Dundurs correspondence, [Jasiuk 1995]), the identities in (10) formally turn into the traction-free conditions along a hole:

$$\sigma_{\rho\rho}(\xi), \sigma_{\rho\theta}(\xi) = 0, \quad \sigma_{\theta\theta}(\xi) = 4 \operatorname{Re} \Phi(\xi), \quad (11)$$

or, equivalently, in KM terms, the homogeneous condition (7)

$$\overline{\varphi(\xi)} + \overline{\omega(\xi)} \frac{\varphi'(\xi)}{\omega'(\xi)} + \psi(\xi) = 0, \quad \xi \in \gamma, \quad (12)$$

explicitly repeated here for future reference. Namely the resemblance of the zero-traction and the zero-displacements boundary conditions (12), (4) leads to the resolving relations (10).

Except in these two homogeneous cases, the computation of boundary stresses does involve the second potential  $\Psi(\xi)$ , which presents additional difficulties.

In this context, a disadvantage of the mapping technique is that it returns the local stresses (10) from  $\Sigma_2$  to the physical domain  $S_2$  only as two invariant combinations  $I_1, I_2$  rather than separately [Muskhelishvili 1975]:

$$I_1(\zeta) = \sigma_{\rho\rho}(\zeta) + \sigma_{\theta\theta}(\zeta) = I_1(z) = \sigma_{rr}(z) + \sigma_{tt}(z), \quad (13a)$$

$$I_2(\zeta) = \sigma_{\rho\theta}^2 - \sigma_{\rho\rho}(z)\sigma_{\theta\theta}(z) = I_2(z) = \sigma_{r\vartheta}^2 - \sigma_{rr}(z)\sigma_{\vartheta\vartheta}(z), \quad (13b)$$

$$z = \omega(\zeta) = r \exp(i\vartheta).$$

However, this is sufficient for our purpose of minimizing the von Mises equivalent stresses  $\sigma_M(\zeta)$  along  $\Gamma$ :

$$\sigma_M^2(z) = \sigma_M^2(\zeta) = I_1^2(\zeta) + 3I_2(\zeta), \quad \zeta \in \Sigma_2 + \gamma, \quad z = \omega(\zeta) \in S_2 + \Gamma. \quad (14)$$

Through the use of (13), we can express the stresses in terms of  $\Phi(\xi)$  ( $\xi = \exp i\theta \in \gamma, t = \omega(\xi) \in \Gamma$ ):

$$\sigma_M^2(t) = D_1(\lambda) \operatorname{Re}^2 \Phi(\xi) + D_2(\lambda) \operatorname{Im}^2 \Phi(\xi), \quad D_1(\lambda) = 7 - 6\lambda + 3\lambda^2, \quad D_2(\lambda) = 3(1 + \lambda)^2, \quad (15a)$$

$$\mathcal{M} \equiv \max_{t \in \Gamma} |\sigma_M(t)| = \max_{0 \leq \theta \leq 2\pi} |\sigma_M(\theta)|. \quad (15b)$$

Further, the boundary value function  $\Psi(\xi)$  is routinely eliminated from the homogeneous condition (9) by integrating over  $\gamma$  with the kernel  $(\xi - \eta)^{-1}$ ,  $\xi, \eta \in \gamma$  [Muskhelishvili 1975]. This leads to a Cauchy-type singular integral equation in  $\Phi(\xi)$ :

$$\frac{1}{2}F(\eta) + \frac{1}{2\pi i} \int_{\gamma} \frac{F(\xi)}{\eta - \xi} d\xi = (1 - \lambda)\Phi(\infty) + \Psi(\infty), \tag{16a}$$

$$F(\xi) \equiv 2\overline{\xi^2 \omega'(\xi)} \operatorname{Re} \Phi(\xi) - \overline{\omega(\xi)} \Phi'(\xi), \quad \xi, \eta \in \gamma, \tag{16b}$$

where the function's asymptotic behavior is dictated by the applied loads (1) (see [Muskhelishvili 1975]):

$$4\Phi(\zeta) = P + Q + O(|\zeta|^{-2}), \quad 2\Psi(\zeta) = Q - P + O(|\zeta|^{-2}), \quad |\zeta| \rightarrow \infty. \tag{17}$$

For a circular integration path, substitution of the convergent Laurent expansions

$$\omega(\zeta) = \zeta + \sum_{k=1}^{\infty} d_k \zeta^{-k}, \quad \Phi(\zeta) = \frac{P + Q}{4} + \sum_{k=2}^{\infty} a_k \zeta^{-k}, \quad \zeta \in \Sigma_2 + \gamma \tag{18}$$

transforms (16a), with the use of (2), into a more easily handled infinite system of linear algebraic equations in  $\{a_k\}$ ,  $k \geq 2$ :

$$-\lambda a_{m+2} + \lambda \sum_{k=2}^m (m - k + 1) \bar{d}_{m-k+1} a_k - (m + 1) \sum_{k=2}^{\infty} \bar{d}_{m+k+1} \bar{a}_k = A_m, \tag{19}$$

$$m = 0, 1, \dots, \quad A_0 = P, \quad A_1 = 0, \quad 2A_m = -(P + Q)(m + 1) \bar{d}_{m+1}, \quad m \geq 2.$$

In conformity with the asymptotics (17) we have  $a_1 = 0$ , so the first sum is omitted in (19) when  $m = 0, 1$ .

At  $\lambda = -1$ , Equations (19) again go over to the resolving system for a traction-free circular hole derived by [Kalandia 1975] and extensively used in [Vigdergauz 2006].

Finally, with pure shear

$$Q = -P = 1, \quad \Phi(\zeta) = O(|\zeta|^{-2}), \tag{20}$$

the normalized energy increment possesses the form (see for instance [Jasiuk 1995])

$$\delta W(\Gamma) = 4\pi f_1^{-1} a_2, \quad Q = -P = 1, \tag{21}$$

while all but the first equation in (19) become homogeneous:

$$A_0 = 1, \quad A_m = 0 \text{ for } m > 0. \tag{22}$$

This completes the one-potential formulation (19) of the forward problem of finding the criteria (15b), (21) for inclusion shape  $\Gamma$  as given by the mapping function (18). Concluding this section, we note that the energy criterion (21) is computationally much simpler than  $\mathcal{M}$ , since it involves only the first Laurent term  $a_2$ . In the current context, the  $\delta W$  value is a byproduct of the direct problem solution for  $\sigma_M$ . For this reason, the  $\delta W$ -optimization is further addressed as a problem accompanying the  $\mathcal{M}$ -shape optimization.



### 3. Analytical truncation of the forward problem

In numerical applications, the infinite system (19) has to be truncated in some way. This is usually done by keeping only the first  $N$  equations, that is, by dropping high-order expansion terms both in  $\omega(\zeta)$ , so that

$$\omega(\zeta) = \zeta + \sum_{k=1}^N d_k \zeta^{-k} \quad (d_k = 0 \text{ for } k > N), \quad (23)$$

and in  $\Phi(\zeta)$ , so that

$$\Phi(\zeta) = \sum_{k=2}^N a_k \zeta^{-k}, \quad (a_k = 0 \text{ for } k > N); \quad (24)$$

see [Jasiuk 1995; Cherkaev et al. 1998] and [Tsukrov and Novak 2004], where the resolving system for  $\Phi(\zeta)$  is derived by collocating the contact conditions at a finite number of points along the interface.

The  $N$ -term solution so obtained amounts to a finite approximation in inverse powers for the potential  $\Phi(\zeta)$ . This algorithm solves well the energy criterion (21) but it is too time-consuming for finding the boundary stresses (13), (14). Indeed, due to their local nature, double truncation (DT) requires a rather large value of  $N$ ; otherwise spurious oscillations are produced along the interface, as exemplified in [Vigdergauz 2006]. As a remedy, we use here the specific structure of Equations (19) to restrict the truncations only to (23), while keeping all the coefficients  $a_k$ ,  $k = 2, 3, \dots$ , as in (18). This *single truncation* (ST) scheme allows us to perform this infinite sum *analytically* for any finite-term expansion of the mapping function  $\omega(\zeta)$ , and so yields a *finite* semianalytic approach for solving the *infinite* system (19). The ST scheme was proposed and detailed for a free-traction hole in [Vigdergauz 2006]. For reference, we go over it briefly, focusing on the differences between the earlier case and the one discussed here.

To begin with, the coefficients  $\{d_m\}$  obey the necessary conditions of mapping univalence, namely

$$-\frac{1}{\sqrt{m}} \leq d_m \leq \frac{1}{\sqrt{m}}, \quad m = 1, 2, \dots, \quad (25)$$

following from the nonnegativeness of the area  $f_1$  inside  $\Gamma$  [Schinzinger and Laura 1991]:

$$f_1 = \pi \left( 1 - \sum_{k=1}^{\infty} k |d_k|^2 \right) \geq 0.$$

Though the asymptotics for the interval width is only  $m^{-1/2}$  as  $m \rightarrow \infty$ , the coefficients  $\{d_m\}$  actually tend to zero much faster, at least as fast as  $m^{-2}$ , provided  $\Gamma$  is a continuous curve [Schinzinger and Laura 1991]. It is precisely this asymptotic behavior that makes the assumed mapping truncation (23) numerically reasonable.

With the truncation (23), the second sum in (19) is also finite. It contributes to the first  $N-2-m$  unknown entries  $a_2, \dots, a_{N-1-m}$  in the  $m$ -th equation, where  $m < N-2$ , and hence disappears in all equations from  $m = N-2$  onward. As a result, each matrix row contains either  $N$  (at  $m < N$ ) or  $N+1$  (at  $m = N, N+1, \dots$ ) nonzero entries arranged as shown schematically on the next page for for  $N = 3$  and the shear-type right-hand side (22).

$$\begin{pmatrix} \times & \times & \times & & & & & \\ \times & \times & \times & & & & & \\ \times & \times & \times & & & & & \\ \times & \times & \times & \times & & & & \\ & \times & \times & \times & \times & & & \\ & & \times & \times & \times & \times & & \\ & & & \times & \times & \times & \times & \\ & & & & \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ \dots \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \dots \end{pmatrix} \tag{26}$$

Put differently, the finite expansion (23) decomposes the initial system (19) into two subsystems. The first is just the DT-reduced  $N \times N$  nonhomogeneous system to solve the first  $N$  coefficients  $a_k$ , for  $k = 2, 3, \dots, N + 1$ . The second ST-specific homogeneous subsystem in the rest of unknowns  $a_k$ ,  $k > N + 1$ , is neglected in the DT-scheme. An important point is that its  $(N+1)$ -width band matrix comprises an infinite linear system of finite differences [Levy and Lessman 1959]

$$a_{m+2} - \sum_{k=1}^N kd_k a_{m-k+1} = 0, \quad m = N, N+1, \dots, \tag{27}$$

with constant coefficients  $1, 0, -d_1, -2d_2, \dots, -Nd_N$ , which define the  $N$ -term Laurent expansion of  $\omega'(\zeta)$ , as seen from (23). In contrast to the first subsystem, these equations are free of the matrix parameter  $\lambda$ .

Solving (27) by a standard finite-differences approach [Levy and Lessman 1959] we obtain  $\Phi(\zeta)$  after some algebra as a simple rational function of  $\zeta^{-1}$ :

$$\Phi(\zeta) = \frac{R_N(\zeta^{-1})}{\zeta \omega'(\zeta)} \tag{28}$$

Here  $R_N(\zeta)$  is a degree- $N$  polynomial in  $\zeta$  with coefficients linearly composed of the first  $\{d_k\}$ ,  $k \leq N$ :

$$R_N(\zeta) = \sum_{k=1}^N r_k \zeta^k, \quad \text{with } r_1 = a_2, \quad r_m = a_{m+1} + \sum_{j=1}^{m-1} (-1)^{m-j} (m-j+1) d_{m-j+1} a_j \text{ for } m \geq 2, \tag{29}$$

which in turn are to be solved for from the first  $N \times N$  subsystem. In passing we note that  $r_0 = 0$  to match the shear-type asymptotics (20).

The proposed algorithm works faster than its basic counterpart [Muskhelishvili 1975], which requires a time-consuming precomputation of the  $N$ -term expansion of the composed function  $\omega(\zeta)/\overline{\omega'(1/\zeta)}$ . This is of particular value when the same forward problem is repeatedly solved many times in evolutionary optimization methods.

The truncation (23) and the resulting finite expressions (28)–(29) allow us at last to reformulate the optimization problem in a numerically tractable form.

*For a given finite number  $N$  of mapping coefficients and pure shear field  $B = 0$ , find the  $\mathcal{M}$ -optimal rigid inclusion shape on which*

$$\mathcal{M} \equiv \max_{t \in \Gamma_N} \sigma_M(t) \equiv \max_{0 \leq \theta \leq 2\pi} \sigma_M(\xi \equiv \exp i\theta) \xrightarrow{\{\Gamma_N\}} \min, \tag{30}$$

where  $\{\Gamma_N\}$  denotes the set of all curves mapped onto the unit circle by  $\omega(\zeta)$  with any admissible finite set  $\{d_m, m = \overline{1, N}\}$ , and  $d_m = 0, m > N$ .

The accompanying  $\delta W$ -optimization problem is posed similarly.

#### 4. One-term analytical relations at a rigid interface

It is well known that even a one-term conformal mapping covers a wide variety shape geometries: from a rectilinear cut (with the Zhukovsky function  $\omega(\zeta) = \zeta + \zeta^{-1}$  — see [Schinzinger and Laura 1991]) to a  $(p+1)$ -cusped hypocycloid (with  $\omega(\zeta) = \zeta + p^{-1}\zeta^{-p}$ ,  $p = 2, 3, \dots$ ). With this in view, the one-term closed expressions for  $\delta W$  and  $\sigma_M$  are derived and analyzed below to gain qualitative insight into the optimal solution, which appears to depend on the rotational symmetry of the contour sought. This symmetry is defined through an integer parameter  $p = 1, 2, \dots$ , defined so that  $\alpha = 2\pi/(p+1)$  is the minimal rotation angle transforming the contour to itself:

$$\omega(e^{i\alpha}\zeta) = e^{i\alpha}\omega(\zeta),$$

or, equivalently,

$$\omega(\zeta) = \zeta + \sum_{k=1}^{\infty} d_{pk}\zeta^{-pk}.$$

In particular, the one-term mapping takes the form

$$\omega(\zeta) = \zeta + d_p\zeta^{-p}, \quad \omega'(\zeta) = 1 - pd_p\zeta^{-p-1}, \quad f_1 = \pi(1 - pd_p^2), \quad (31)$$

and hence the summation coefficients in (29) simplify:  $r_m = a_{m+1}$ ,  $m \geq 1$ . As in the hole case [Vigdergauz 2006], this results in

$$\Phi(\zeta) = \frac{a_2\zeta^{p-1} + a_{p-1}\zeta^2}{\zeta^{p+1} - pd_p}, \quad (32a)$$

$$\sigma_M(\theta) = D_1(\lambda)F_1(\theta) + D_2(\lambda)F_2(\theta), \quad (32b)$$

$$F_1(\theta) = \left( \frac{a_2(\cos 2\theta - pd_p \cos(p-1)\theta) + a_{p-1}(\cos(p-1)\theta - pd_p \cos 2\theta)}{1 - 2pd_p \cos(p+1)\theta + p^2d_p^2} \right)^2, \quad (32c)$$

$$F_2(\theta) = \left( \frac{a_2(\sin 2\theta + pd_p \sin(p-1)\theta) + a_{p-1}(\sin(p-1)\theta + pd_p \sin 2\theta)}{1 - 2pd_p \cos(p+1)\theta + p^2d_p^2} \right)^2, \quad (32d)$$

where the dependence on  $\lambda$  is hidden in the Laurent coefficients  $a_2, a_{p-1}$  of the potential  $\Phi(\zeta)$ . They are found by solving the first nonvanishing equations (19). For  $p \leq 3$ , the second term  $a_{p-1}$  is zero and should be dropped from (31). The one-term results so obtained are detailed below for different symmetries. For convenience, the criteria are henceforth normalized to the interval  $[0, 1]$  by their counterparts  $\delta W_0, \mathcal{M}_0$  for a circular inclusion, when  $\gamma \equiv \Gamma$ ,  $\omega(\zeta) = \zeta$ ,  $f_1 = \pi$ ,  $\Phi(\zeta) = -1/\lambda\zeta^2$  [Muskhelishvili 1975] and hence, with (21), (15),

$$\delta W_0 = -4/\lambda, \quad \sigma_M^2(\theta) = D_1(\lambda) \cos^2 2\theta + D_2(\lambda)^2 \sin^2 2\theta, \quad (33a)$$

$$\mathcal{M}_0 = \max_{\theta} \sigma_M(\theta) = \sigma_M(\pi/4) = D_2^{1/2}(\lambda)/\lambda = (\lambda + 1)\sqrt{3}/\lambda, \quad 1 \leq \lambda \leq 3. \quad (33b)$$

To avoid confusion, we note that the Dundurs correspondence ( $\lambda = -1$ , a traction-free circular hole under shear) does not hold for (33b). Here,

$$\begin{aligned} \max \sigma_M(\theta) &= \sigma_M(0) = D_1^{1/2}(-1) = 2, \\ \min \sigma_M(\theta) &= \sigma_M(\pi/4) = D_2^{1/2}(-1) = 0. \end{aligned}$$

**4.1. Twofold ( $p = 1$ ) and triangular symmetry ( $p = 2$ ).** Here

$$a_2 = -\frac{1}{\lambda}, \quad \delta W_2 = -\frac{1}{1 - 2d_2^2} : \quad d_p^{(\min)} = 0, \quad \min_{d_p} \delta W_p = 1. \tag{34}$$

A simple numerical check of (32b) in the univalence interval (25) shows further that  $\min_{d_p} \mathcal{M} = 1$  at  $d_p = 0$ , independently of  $\lambda$ .

**4.2. Square symmetry ( $p = 3$ ).** Here

$$\begin{aligned} a_2 &= -\frac{1}{\lambda + d_3}, \quad \delta W_3 = \frac{\lambda}{(\lambda + d_3)(1 - 3d_3^2)}, \quad |d_3| \leq \frac{1}{\sqrt{3}}, \\ d_3^{(\min)} &= \frac{\sqrt{\lambda^2 + 1} - \lambda}{3}; \quad \min_{d_3} \delta W_3 = \frac{9\lambda}{2(2\lambda + \sqrt{\lambda^2 + 1})(1 - \lambda^2 + \lambda\sqrt{\lambda^2 + 1})}. \end{aligned} \tag{35}$$

In contrast to the case (34), the square symmetric  $\delta W$ -optimal inclusion differs from a circle. It is shaped as a rounded square rotated by 45 degrees with respect to the axes  $X, Y$ . The roundness depends on  $\lambda$  and is measured by the deviation of  $d_3^{(\min)}$  from the first mapping coefficient  $d_3^{(\text{square})} = 1/6$  [Jasiuk 1995] of an exact square shape. It varies from  $\approx 0.0286$  at  $\lambda = 1$  to  $\approx 0.1126$  at  $\lambda = 3$ . As a result, the energy  $\delta W_3^{(\text{square})}$  exceeds the minimum (35) by  $\approx 0.00308$  at  $\lambda = 1$  ( $\min \delta W_3 \approx 0.93198$ ,  $\delta W_3^{(\text{square})} \approx 0.93506$ ) and by  $\approx 0.04251$  at  $\lambda = 3$  ( $\min \delta W_3 \approx 0.99099$ ,  $\delta W_3^{(\text{square})} \approx 1.03349$ ). The latter comparison shows that just small local improvements of an exact square shape near the angular points allow one to go beyond the circle-related value  $\delta W^{(\text{circle})} = 1$ . This is also true for the multiterm approximation (Section 6) with  $\delta W^{(\text{square})} = 1$  at  $\lambda \approx 1.88149$ .

Another observation shows that at  $\lambda = 1$  (an incompressible matrix) the optimal coefficient  $d_3^{(\min)} = (\sqrt{2} - 1)/3$  is equal in value but opposite in sign to its counterpart for a hole under shear [Cherkaev et al. 1998; Vigdergauz 2006]. In other words, both one-term shapes are the same, differing only in orientation about the axes.

Similarly, the  $\mathcal{M}$ -optimal shape is also a rotated rounded square though with somewhat different value of the mapping coefficient.

The already noted resemblance of the energy optimal hole and the rigid inclusion at  $\lambda = 1$  is of general nature and holds for any pair of mutually rotated square-symmetric shapes. Indeed, the mapping function

$$\omega(e^{i\frac{\pi}{4}}\zeta) = e^{\frac{\pi}{4}}\zeta + \sum_{k=1}^{\infty} e^{(1-4k)\frac{\pi}{4}} \frac{d_{4k-1}}{\zeta^{4k-1}} = e^{i\frac{\pi}{4}} \left( \zeta + \sum_{k=1}^{\infty} (-1)^k \frac{d_{4k-1}}{\zeta^{4k-1}} \right), \tag{36}$$

with alternating signs of the Laurent coefficients rotates a shape by 45 degrees. Substitution of (36) into the system (19) while putting  $\lambda = -1$  results in a system equivalent to the traction-free hole conditions

(12), whose solution evidently takes the form

$$a_{4k-2}^{(hole)} = (-1)^k a_{4k-2}^{(\lambda=1)}; \quad k = 1, 2, \dots$$

In particular  $a_2^{(hole)} = -a_2^{(\lambda=1)}$  and hence both (no normalized) energy increments are the same up to the sign.

### 4.3. Higher symmetry ( $p \geq 4$ ).

$$a_2 = -\frac{\lambda}{\lambda^2 - (p-2)d_p^2}, \quad a_{p-1} = \frac{(p-2)d_p}{\lambda^2 - (p-2)d_p^2}, \quad |d_p| \leq \frac{1}{\sqrt{p}},$$

$$\delta W_p = \frac{\lambda^2}{(\lambda^2 - (p-2)d_p^2)(1 - pd_p^2)}, \quad \min_{d_p} \delta W_p = 1, \quad d_p^{(min)} = 0.$$

Again, as in Section 4.1,  $\mathcal{M}$ -optimal contour appears to be a circle with  $\mathcal{M} = 1$ .

The one-term results suggest that the most promising for further analysis is the fourfold ( $p = 3$ ) symmetry case when the  $\delta W$ - and  $\mathcal{M}$ -optima are achieved at similarly rotated square-like shapes. As numerically shown in Section 6, this pattern is preserved for multiterm  $N > p = 3$  approximations.

## 5. The GA scheme

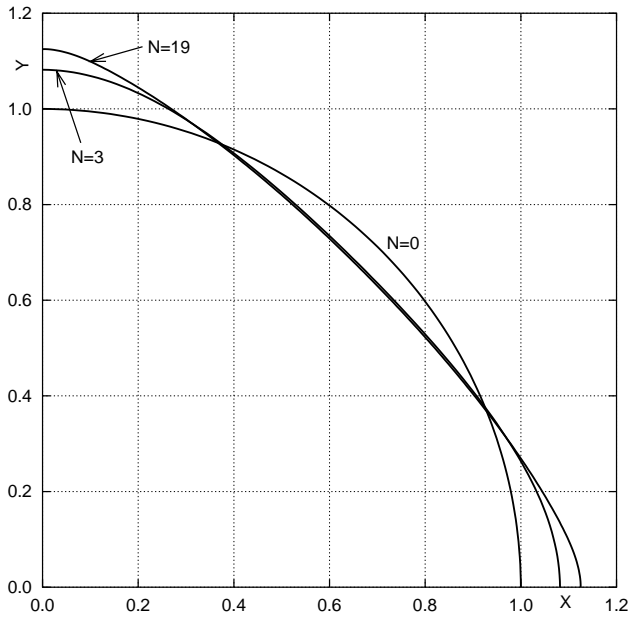
As described above, the shape optimization problem is solved by minimizing either  $\mathcal{M}$  or  $\delta W$  criterion over a  $N$ -parameter set of closed curves. In principle, the single truncation approach (27)–(29) explicitly provides all the information required by a standard gradient-based optimization. However, this results in too complicated analytical manipulations and time-consuming calculations. A good alternative is a relatively novel genetic algorithm, which mimics the Darwinian evolution paradigm, both in the concept and in the terms. Devised by Holland [1975], the GA finds wide application in engineering. In the current context, the relevant information is detailed in [Vigdergauz 2006], where a fairly standard GA is applied to find the stress-minimizing hole under remote shear.

The GA optimization process is governed by a number of parameters that control its performance. Here parameter tuning is of lesser importance due to the relatively small problem size (the first five nonzero mapping terms appear to provide a good convergence in the square symmetric case). For this reason, we use the parameter values found in the process of hole shape optimization [Vigdergauz 2006].

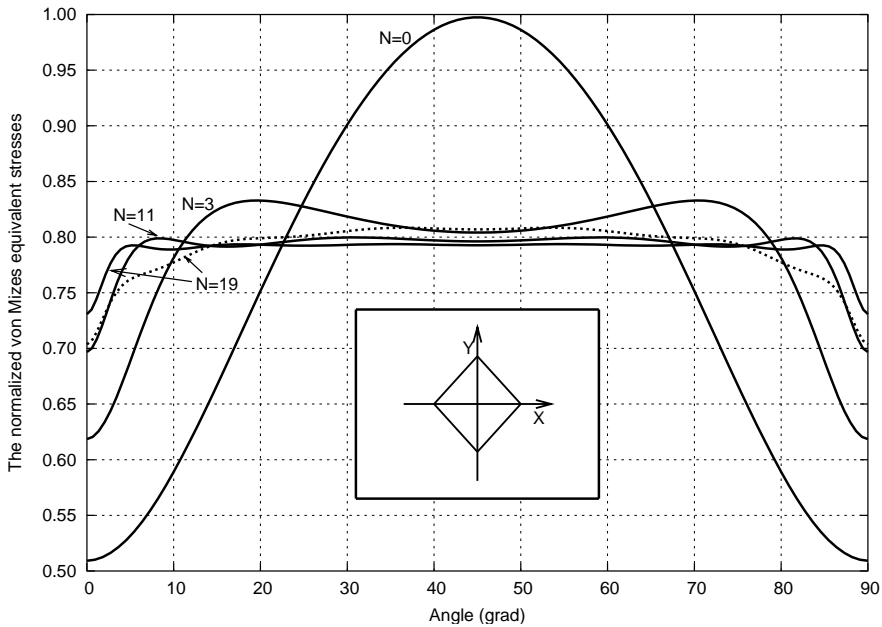
## 6. Multiterm numerical results

In quantifying the shape optimization effect on the matrix stress-strain state, we first find a compromise truncation number  $N$  to achieve a reasonable numerical accuracy at a not-too-large computational cost. Tables 1 and 2 display the optimization results obtained by sequentially truncating the resolving system (19) at the first nonvanishing equation in the square symmetric case.

The observed fast convergence is explained differently for either criterion. As noted in Section 2, the higher mapping coefficients go to zero rather rapidly and hence have little or no effect on the integral-type energy increment  $\delta W$ . On the other hand, they are responsible for forming large curvature isolated shape points (see Figure 2) which may only worsen the local criterion  $\mathcal{M}$ . In this respect, of special interest is the von Mises stress distribution along the  $\mathcal{M}$ -optimal interface as depicted in Figure 3 in dependence of



**Figure 2.** A quarter of the  $M$ -optimal rigid inclusion under simple shear for different number of the first mapping terms.



**Figure 3.** Distribution of the normalized von Mises equivalent stresses along the square-like  $M$ -optimal rigid inclusion in dependence of the problem mapping size beginning with a circle ( $N = 0$ ). The  $\delta W$ -optimal distribution (dotted) is also added for comparison. The inset shows the inclusion orientation.

$N$	$d_3$	$d_7$	$d_{11}$	$d_{15}$	$d_{19}$	$\delta W_{\min}$	$\delta W^{(\text{square})}$
3	0.078689					0.980356	1.006993
7	0.080125	0.006922				0.979970	1.005677
11	0.080356	0.007244	0.001794			0.979929	1.005269
15	0.080415	0.007319	0.001910	0.000671		0.979921	1.005102
19	0.080434	0.007342	0.001942	0.000722	0.000303	0.979919	1.005083

**Table 1.** A rigid square symmetric inclusion under remote shear at  $\lambda = 2$ : the global minimum  $\delta W_{\min}$  and its mapping coefficients for the first values of  $N$ . The energy values for an exact square shape are also added for comparison.

$N$ . The curves are seen to converge to the uniform distribution, with successively diminishing deviations occurring near the angles  $\theta_l = l\pi/2$ ,  $l = 0, 1, \dots$ . In some contrast, the  $\delta W$ -optimal von Mises stresses (the dotted line) show diminishing local peaks, though generally this is a very similar pattern. The same situation is numerically observed everywhere in the interval  $1 \leq \lambda \leq 3$ . This allows us to conservatively conclude that *the full-size optimal solution should lead to the exactly uniform distribution of the von Mises stresses along the rigid inclusion shape with their level depending on  $\lambda$ .*

Of course, it is only a numerical conjecture rather than a rigorous analytical proof which is outside of our current scope.

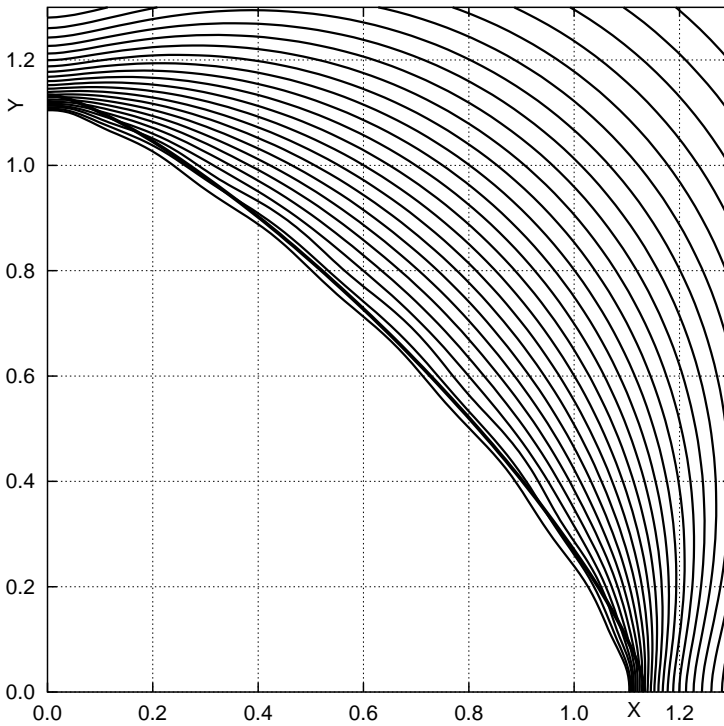
The rest of the numerical results is obtained by holding the first five terms  $d = 4_{4k-1}$ ,  $k = 1, \dots, 5$ , or, equivalently by truncating the system at  $N = 19$ .

Figure 4 exemplifies the von Mises stress levels around the  $\mathcal{M}$ -optimal inclusion at  $\nu_{\text{stress}} = 1/3$ , ( $\lambda = 2$ ). As one might expect, the maximum stress gradient occurs near the angles  $\theta_l$  where the level lines thicken.

Figure 5 displays the  $\mathcal{M}$ - and  $\delta W$ -optimal shapes for the limiting values of  $\lambda$ . The  $\delta W$ -optimal shapes, as compared to their  $\mathcal{M}$ -optimal counterparts, have angular points. Though distinct, they do not reach the critical openings  $\alpha(\lambda)$  where the stresses become singular, contrary to a physical possibility for real material.

$N$	$d_3$	$d_7$	$d_{11}$	$d_{15}$	$d_{19}$	$\mathcal{K}_{\min}$	$\delta W$
3	0.067933					0.834158	0.980727
7	0.007981	0.005158				0.808959	0.979996
11	0.084671	0.006868	0.001245			0.800220	0.979997
15	0.086647	0.007422	0.001631	0.000394		0.796231	0.980050
19	0.088398	0.008069	0.002011	0.000642	0.000194	0.793977	0.980122

**Table 2.** A rigid square symmetric inclusion under remote shear at  $\lambda = 2$ : the conformal mapping coefficients and the local criterion  $\mathcal{K}_{\min}$  resulted from the GA optimization process for different values of  $N$ . The corresponding global criterion  $\delta W$  is also shown, for comparison with its optimal values in Table 1.



**Figure 4.** A quarter of the  $\mathcal{M}$ -optimal rigid inclusion under simple shear: von Mises equivalent stress contours at  $\lambda = 2$ .

The formation of the angles is comprehensively analyzed by [Cherkaev et al. 1998]. Here we only note that  $\alpha(\lambda)$  are in the numerically found interval

$$\alpha(1) \leq \alpha(\lambda) \leq \alpha(3), \quad \alpha(1) \approx 102.5^\circ, \quad \alpha(3) \approx 152.5^\circ.$$

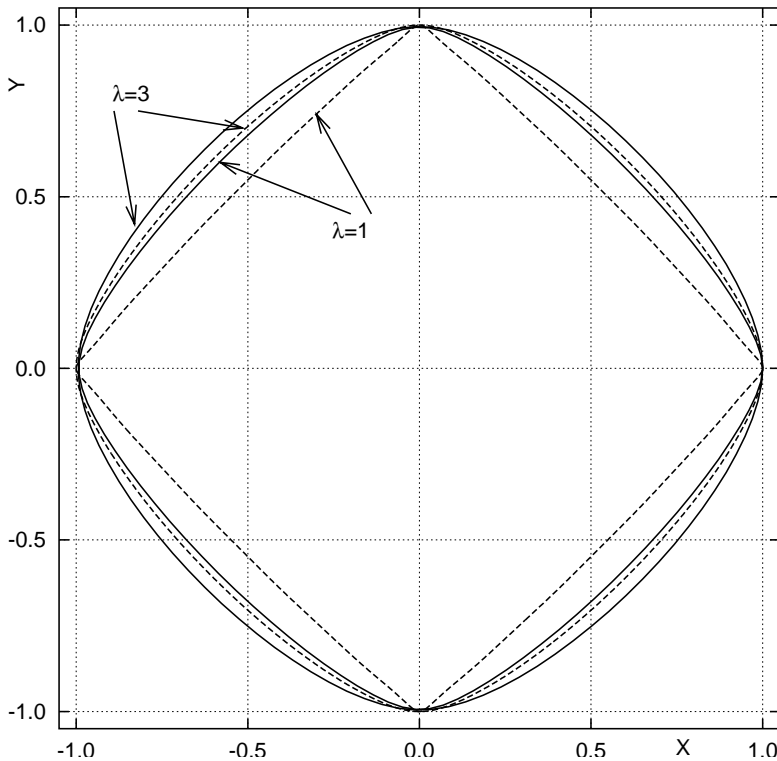
Another observation is that the  $\delta W$ -optimal shape is flatter than its  $\mathcal{M}$ -optimal counterpart at  $\lambda = 1$  and quite the reverse at  $\lambda = 3$ . This suggests that there is an intermediate value  $\lambda^*$  at which the shapes and their criteria come the closest together. The computation gives  $\lambda^* \approx 1.712 \dots$

Finally, Figure 6 gives the criteria at hand for either of the two optimal shapes in dependence of  $\lambda$ . It is seen that the  $\delta W$ -criterion of the  $\mathcal{M}$ -optimal shapes markedly exceeds  $\min \delta W$  only near the point  $\lambda = 1$ . Contrastingly, the  $\mathcal{M}$ -criterion of the  $\delta W$ -optimal shapes is substantially above its minimum everywhere except near the above-found point  $\lambda^*$ . Put it differently, the  $\delta W$ -optimization is performed partly at the penalty of increasing the  $\mathcal{M}$ -criterion

### 7. Conclusions and future applications

We have presented a novel semianalytical solution for isolated, nondeformable inclusion subjected to pure remote shear. An inclusion shape is presented as a finite-term conformal map of a circle rather than by more traditional nodal points. This allows to pose the optimization as a small-size numerical problem with linear constraints particularly amenable to the GA search. Another essential idea is to express the





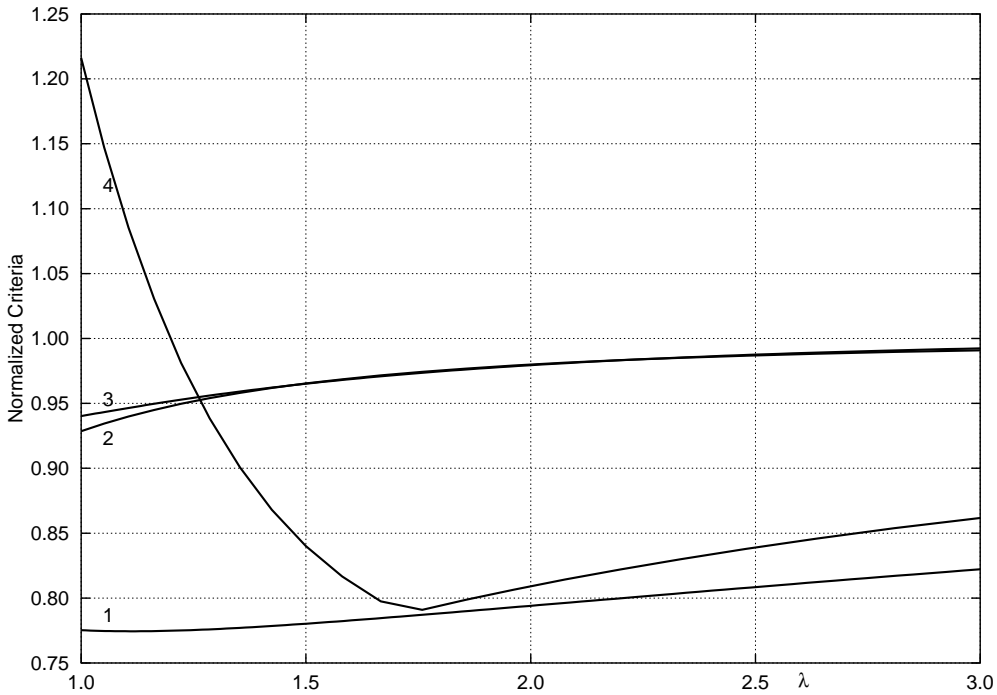
**Figure 5.** The  $\mathcal{M}$ -optimal and  $\delta W$ -optimal (dotted) inclusion shapes for the limiting values of the elastic parameters  $\lambda$ .

stresses through only one KM potential whose infinite Laurent tail is summed in a closed form. The numerical accuracy so achieved permits to solve the yet unconsidered problem of globally minimizing the von Mises local criterion over a broad range of continuous curves. The results obtained are in strong parallel to those for the optimal hole [Vigdergauz 2006]:

- The both shapes differ from a circle only in the square symmetric case when they are close to a slightly rounded square located in a specifically rotated position relative to each other;
- For the  $\delta W$ -criterion the roundness is combined with four distinct angular points whose opening is greater than the critical values which bring stress singularities;
- The  $\delta W$ - and  $\mathcal{M}$ -optimal shapes are very similar though not exactly the same;
- Under the  $\mathcal{M}$ -criteria, the von Mises stress distribution becomes uniform along the optimal interfaces.

Of course, these conclusions are arrived at only numerically and the analytical proof would be very desirable.

It is worthy of note that the conformal mapping is used not only for the shape description but also to transform the stress field around the shape. This may be done only for a hole or a rigid inclusion which exhausts the setup with a single deformable region. In order to fill gap between these limiting cases, an optimal shaping of elastic inclusion should be considered. This problem is far too complicated



**Figure 6.**  $\min \mathcal{M}$  (curve 1),  $\min \delta W$  (curve 2),  $\delta W$  of  $\mathcal{M}$ -optimal shapes (curve 3),  $\mathcal{M}$  of  $\delta W$ -optimal shapes (curve 4) as functions of  $\lambda$ .

because both elastic regions may not be mapped simultaneously. Nevertheless, the proposed mapping parameterization inside the GA seems to be potentially useful here too.

### References

- [Cherepanov 1974] G. P. Cherepanov, “Inverse problem of the plane theory of elasticity”, *Prikl. Mat. Mekh.* **38**:6 (1974), 963–979. In Russian; translated in *J. Appl. Math. Mech.* **38**:6 (1974), 915–931.
- [Cherepanov 1997] G. P. Cherepanov, *Methods of fracture mechanics: Solid matter physics*, Solid Mechanics and its Applications **51**, Kluwer, Boston, 1997.
- [Cherkaev et al. 1998] A. V. Cherkaev, Y. Grabovsky, A. B. Movchan, and S. K. Serkov, “The cavity of the optimal shape under the shear stresses”, *Int. J. Solids Struct.* **35**:33 (1998), 4391–4410.
- [England 1971] A. H. England, *Complex variable methods in elasticity*, Wiley, London, 1971. Reprinted Dover, New York, 203.
- [Gen and Cheng 1997] M. Gen and R. Cheng, *Generic algorithm and engineering design*, Wiley, New York, 1997.
- [Grabovsky and Kohn 1994] Y. Grabovsky and R. V. Kohn, “Microstructures minimizing the energy of a two phase elastic composite in two space dimensions, II: The Vigdergauz microstructure”, *J. Mech. Phys. Solids* **43** (1994), 949–972.
- [Holland 1975] J. H. Holland, *Adaptation in natural and artificial systems*, University of Michigan Press, Ann Arbor, 1975.
- [Jasiuk 1995] I. Jasiuk, “Cavities vis-a-vis rigid inclusions: Elastic moduli of materials with polygonal inclusions”, *Int. J. Solids Struct.* **32**:3-4 (1995), 407–422.
- [Kalandia 1975] A. I. Kalandia, *Mathematical methods of two-dimensional elasticity*, Mir, Moscow, 1975.

- [Levy and Lessman 1959] H. Levy and F. Lessman, *Finite difference equations*, Pitman, London, 1959. Reprinted Dover, New York, 1992.
- [Muskhelishvili 1975] N. I. Muskhelishvili, *Some basic problems of the mathematical theory of elasticity*, 2nd ed., Noordhoff, Leiden, 1975.
- [Osyczka 2001] A. Osyczka, *Evolutionary algorithms for single and multicriteria design optimization*, Springer Physica-Verlag, Heidelberg, 2001.
- [Savin 1961] G. N. Savin, *Stress concentration around holes*, Pergamon, New York, 1961.
- [Schinzinger and Laura 1991] R. Schinzinger and P. A. A. Laura, *Conformal mapping: methods and applications*, Elsevier, New York, 1991. Reprinted Dover, New York, 2003.
- [Tsukrov and Novak 2004] I. Tsukrov and J. Novak, "Effective elastic properties of solids with two-dimensional inclusions of irregular shapes", *Int. J. Solids Struct.* **41**:24-25 (2004), 6905–6924.
- [Vigdergauz 1976] S. B. Vigdergauz, "Integral equation of the inverse problem of the plane theory of elasticity", *Prik. Mat. Mekh.* **40**:3 (1976), 566–569. In Russian; translated in *J. Appl. Math. Mech.* **40**:3 (1976), 518–522.
- [Vigdergauz 2006] S. Vigdergauz, "The stress-minimizing hole in an elastic plate under remote shear", *J. Mech. Mater. Struct.* **1**:2 (2006), 387–406.

Received 24 Jul 2006. Accepted 9 Oct 2006.

SHMUEL VIGDERGAUZ: smuel@iec.co.il

Research and Development Division, The Israel Electric Corporation Ltd., P.O.Box 10, Haifa 31000, Israel

## SYMMETRY CONDITIONS AND FINITE MECHANISMS

SIMON D. GUEST AND PATRICK W. FOWLER

Using group representation theory, a simplified criterion for the detection of finite symmetric mechanisms is presented.

### 1. Introduction

The identification of finite mechanisms in statically and kinematically indeterminate structures is, in general, a difficult problem. However, Kangwai and Guest [1999] showed that in certain cases finiteness of mechanisms could be found using only symmetry arguments and a linear analysis. Here we revisit Kangwai and Guest's method to show that their symmetry arguments can be straightforwardly stated in terms of representations of mechanisms and states of self-stress in the point group of the structure, giving an immediate assessment of the finiteness of mechanisms for many cases.

For any kinematically indeterminate structure, it is possible to find a set of mechanisms, i.e., displacements which to first order cause no deformation of structural elements. (Here it is usual to exclude rigid body motions.) Mechanisms may be either *finite*, in which case there is a continuous displacement path that is compatible at every point with zero deformation of the structure, or *infinitesimal*, in which case there is deformation at second or higher order. Determination of the finite nature of a mechanism in general requires nonlinear analysis [Tarnai 1989; Calladine and Pellegrino 1992; Salerno 1992; Connelly and Servatius 1994; Tarnai and Szabó 2000; Garcea et al. 2005]. Kuznetsov [2000] has stressed the difficulties that may arise with 'singular' (e.g., highly symmetric) configurations, but nonetheless, the behaviour at points of high symmetry is often a useful guide to that of physical systems, where the symmetry may be only approximate. Kangwai and Guest [1999] introduced, for specific symmetric cases, a criterion that could determine the finiteness of a mechanism based on purely first-order analysis combined with a symmetry argument, and has proved to be applicable to a wide variety of structures [Kovács et al. 2004; Fowler and Guest 2005]. We show here that there is a simple and general way of determining finiteness according to this criterion, obviating the need for explicit calculation in every particular case.

The difficult cases for determining finiteness of mechanisms are those where structures are also statically indeterminate, and hence have states of self-stress, i.e., sets of internal stresses in self-equilibrium in the absence of externally applied loads. Here is the symmetry finiteness criterion, as stated in [Kangwai and Guest 1999]:

**Proposition 1.** *If a mechanism is fully-symmetric in some subgroup of the symmetry group of the structure, with no equisymmetric state of self-stress, then that mechanism must be finite.*

---

Guest acknowledges the support of the Leverhulme Trust, and Fowler the support of the Royal Society through the RS/Wolfson Research Merit Award Scheme.

However, the converse does not always hold: if such an equisymmetric state of self-stress exists, then the mechanism *may* be stiffened, and hence be only infinitesimal, or *may* still be finite. A celebrated example where the converse of the proposition would not apply is the cusp mechanism of Connelly and Servatius [1994].

The present paper reformulates the symmetry finiteness criterion in a way that avoids the need to consider subgroups of the symmetry group of the structure. Statement and proof of the new formulation in cases where there is a mechanism belonging to a nondegenerate representation follows in Section 2. This covers all cases that have been analysed so far with the symmetry finiteness criterion. For completeness, the present paper briefly considers, in Section 3, the consequences of degeneracy. Section 4 contains a number of examples of the criterion.

## 2. A symmetry finiteness criterion based on representations

For mechanisms that belong to a nondegenerate representation, it can be shown that the following proposition is equivalent to the symmetry criterion stated by Kangwai and Guest.

**Proposition 2.** *A mechanism that belongs to a nondegenerate representation will be finite if, in the point group of the undisplaced object, there is neither a state of self-stress that is equisymmetric with the mechanism, nor a totally symmetric state of self-stress.*

The proposition can be proved as follows.

Suppose that a structure has a configuration with point-group symmetry  $G$ , and in that configuration has mechanisms spanning the (reducible) representation  $\Gamma(m)$  of  $G$ , and states of self-stress spanning the representation  $\Gamma(s)$ . We will initially concentrate on one member of the set of mechanisms,  $m_1$ , a mechanism with nondegenerate, irreducible representation  $\Gamma_{m_1}$ . We can assume that the mechanism is not totally symmetric, as if it were, Proposition 1 would apply directly: a totally symmetric mechanism will be finite if there is no equisymmetric state of self-stress. Displacement of the structure along  $m_1$  gives a new configuration with point group symmetry  $H_1$ ;  $H_1$  is a subgroup of  $G$  defined entirely by  $\Gamma_{m_1}$ . (Using notation that will be defined in Section 3,  $H_1$  is the *kernel* of  $G$  under  $\Gamma_{m_1}$ .)

Let  $G$  consist of symmetry operations  $R_i, i = 1 \dots |G|$  and let the characters of  $\Gamma_{m_1}$  be  $\chi_{m_1}(R_i)$ . Then  $H_1$  is a subgroup of  $G$ , of order  $|H_1| = |G|/2$ , comprising those operations  $R_i$  of  $G$  for which  $\chi_{m_1}(R_i) = +1$ . It is easy to see that this condition on the characters defines a group. As the characters of a nondegenerate irreducible representation obey the group multiplication table, i.e.,  $\chi(R_i)\chi(R_j) = \chi(R_k)$  for  $R_i R_j = R_k$ , the set of operations with character  $+1$  is closed under multiplication, includes the identity, contains an inverse for every operation in the set, and inherits the associative property from  $G$ .

Suppose that  $\Gamma(s)$  is not empty, and consider a state of self-stress with irreducible representation  $\Gamma_s$ , say, as a candidate for ‘blocking’  $\Gamma_{m_1}$ , i.e., stiffening the mechanism  $m_1$ . There are three possibilities:

- (i)  $\Gamma_s$  is the totally symmetric representation,  $\Gamma_0$ , in  $G$ ;
- (ii)  $\Gamma_s$  is  $\Gamma_{m_1}$  in  $G$ ;
- (iii)  $\Gamma_s$  is neither  $\Gamma_0$  nor  $\Gamma_{m_1}$  in  $G$ .

As a nondegenerate and non-totally symmetric irreducible representation,  $\Gamma_{m_1}$  has character  $+1$  for exactly half of the operations  $R_i$  of  $G$ , and character  $-1$  for the other half (by orthogonality with  $\Gamma_0$ ). For convenience, we will choose an ordering of the operations such that  $\chi_{m_1}(R_i) = +1$  for  $i = 1, \dots, |G|/2$ ,

and  $\chi_{m_1}(R_i) = -1$  for  $i = |G|/2 + 1, \dots, |G|$ . With this ordering, let the characters of the representation of the state of self-stress,  $\Gamma_s$ , be  $\chi_s(R_i) = \alpha_i$  and  $\chi_s(R_{(|G|/2+i)}) = \beta_i$  for  $i = 1, \dots, |G|/2$  with

$$\alpha = \sum_{i=1}^{|G|/2} \alpha_i \quad ; \quad \beta = \sum_{i=1}^{|G|/2} \beta_i.$$

The various characters are summarized thus:

$G$	$R_1$	$\dots$	$R_{ G /2}$	$R_{ G /2+1}$	$\dots$	$R_{ G }$
$\Gamma_0$	+1	$\dots$	+1	+1	$\dots$	+1
$\Gamma_{m_1}$	+1	$\dots$	+1	-1	$\dots$	-1
$\Gamma_s$	$\alpha_1$	$\dots$	$\alpha_{ G /2}$	$\beta_1$	$\dots$	$\beta_{ G /2}$

In case (i), we have  $\alpha_i = \beta_i = +1$ , and  $\Gamma_s = \Gamma_0$  in both  $G$  and  $H_1$ . In case (ii),  $\alpha_i = -\beta_i = +1$ , and  $\Gamma_s = \Gamma_0$  in  $H_1$ , but not  $G$ . In case (iii), orthogonality of  $\Gamma_s$  to  $\Gamma_0$  gives

$$\alpha + \beta = 0,$$

and orthogonality to  $\Gamma_{m_1}$  gives

$$\alpha - \beta = 0,$$

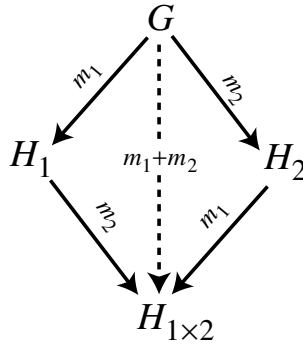
and hence  $\alpha = \beta = 0$ ;  $\alpha = 0$  implies that  $\Gamma_s$  remains orthogonal to  $\Gamma_0$  (and hence to  $\Gamma_{m_1}$ ) in  $H_1$ . Thus in case (i) state of self-stress  $s$  may block mechanism  $m_1$  in both  $G$  and  $H_1$ ; in case (ii)  $s$  may block  $m_1$  in  $H_1$ ; in case (iii)  $s$  does not block  $m_1$ . Notice that the above applies equally to degenerate and nondegenerate  $\Gamma_s$ . Case-by-case consideration has therefore shown the truth of Proposition 2.

Details of the identification of  $\Gamma_{m_1}$  and its associated group  $H_1$  can be filled in from standard character and descent in symmetry tables; see, for example, [Atkins et al. 1970; Salthouse and Ware 1972; Altmann and Herzig 1994].

So far we have considered a single nondegenerate mechanism. If the configuration that has  $G$  symmetry allows several such mechanisms, but displacement occurs along only one of them, the above reasoning applies directly. If, instead, displacement is along some linear combination of such mechanisms, the consequences are easily worked out. For example, suppose that we have mechanisms  $m_1$  and  $m_2$  of distinct symmetries in  $G$ ,  $\Gamma_{m_1}$  and  $\Gamma_{m_2}$ . Displacement along a linear combination of  $m_1$  and  $m_2$  can be analysed with the help of the character table below, where the operations of  $G$  have been separated into equal-sized blocks according to their characters for the irreducible representations  $\Gamma_{m_1}$  and  $\Gamma_{m_2}$ .

$G$	$R_1$	$\dots$	$R_{ G /4}$	$R'_1$	$\dots$	$R'_{ G /4}$	$R''_1$	$\dots$	$R''_{ G /4}$	$R'''_1$	$\dots$	$R'''_{ G /4}$
$\Gamma_0$	+1	$\dots$	+1	+1	$\dots$	+1	+1	$\dots$	+1	+1	$\dots$	+1
$\Gamma_{m_1}$	+1	$\dots$	+1	+1	$\dots$	+1	-1	$\dots$	-1	-1	$\dots$	-1
$\Gamma_{m_2}$	+1	$\dots$	+1	-1	$\dots$	-1	+1	$\dots$	+1	-1	$\dots$	-1
$\Gamma_{m_1} \times \Gamma_{m_2}$	+1	$\dots$	+1	-1	$\dots$	-1	-1	$\dots$	-1	+1	$\dots$	+1
$\Gamma_s$	$\alpha_1$	$\dots$	$\alpha_{ G /4}$	$\beta_1$	$\dots$	$\beta_{ G /4}$	$\gamma_1$	$\dots$	$\gamma_{ G /4}$	$\delta_1$	$\dots$	$\delta_{ G /4}$

The operations  $\{R_1 \dots R_{|G|/4}\} + \{R'_1 \dots R'_{|G|/4}\}$  constitute the group  $H_1$  which is reached from  $G$  by a pure  $m_1$  distortion. Similarly the group  $H_2$  reached from  $G$  by a pure  $m_2$  distortion consists of the  $R$  and  $R''$  operations. The  $R$  operations by themselves define the group  $H_{1 \times 2}$ , which is reached from  $G$  by a displacement along a generic combination of  $m_1$  and  $m_2$ . The relationships between the various



**Figure 1.** The descent in symmetry from  $G$  arising from displacement along mechanisms  $m_1$  and  $m_2$ , alone and in combination.

subgroups of  $G$  are shown schematically in Figure 1. By definition,  $\Gamma_{m_1}$  and  $\Gamma_0$  become totally symmetric in  $H_1$ , and  $\Gamma_{m_2}$  and  $\Gamma_0$  become totally symmetric in  $H_2$ . In the group  $H_{1 \times 2}$ ,  $\Gamma_{m_1}$ ,  $\Gamma_{m_2}$ ,  $\Gamma_{m_1} \times \Gamma_{m_2}$  and  $\Gamma_0$  become totally symmetric. Now consider a candidate state of self-stress,  $s$ . Its characters are defined in the table, and we define the partial sums

$$\alpha = \sum_{i=1}^{|G|/4} \alpha_i \quad ; \quad \beta = \sum_{i=1}^{|G|/4} \beta_i \quad ; \quad \gamma = \sum_{i=1}^{|G|/4} \gamma_i \quad ; \quad \delta = \sum_{i=1}^{|G|/4} \delta_i.$$

There are five possibilities for  $\Gamma_s$ :

- (i)  $\Gamma_s$  is the totally symmetric representation,  $\Gamma_0$ , in  $G$ :  $\alpha_i = \beta_i = \gamma_i = \delta_i = +1$ , and  $\Gamma_s = \Gamma_0$  in  $G$  and all subgroups. Thus state of self-stress  $s$  may block mechanism  $m_1$  and  $m_2$  in any combination.
- (ii)  $\Gamma_s$  is  $\Gamma_{m_1}$  in  $G$ :  $\alpha_i = \beta_i = -\gamma_i = -\delta_i = +1$ , and  $\Gamma_s = \Gamma_0$  in  $H_1$  and  $H_{1 \times 2}$ , but not  $H_2$ . Thus,  $s$  may block all but pure  $m_2$ .
- (iii)  $\Gamma_s$  is  $\Gamma_{m_2}$  in  $G$ :  $\alpha_i = -\beta_i = \gamma_i = -\delta_i = +1$ , and  $\Gamma_s = \Gamma_0$  in  $H_2$  and  $H_{1 \times 2}$ , but not  $H_1$ . Thus,  $s$  may block all but pure  $m_1$ ;
- (iv)  $\Gamma_s$  is  $\Gamma_{m_1} \times \Gamma_{m_2}$  in  $G$ :  $\alpha_i = -\beta_i = -\gamma_i = \delta_i = +1$ , and  $\Gamma_s = \Gamma_0$  in  $H_{1 \times 2}$ , but not  $H_1$  or  $H_2$ . Thus  $s$  may block all but pure  $m_1$  or pure  $m_2$ .
- (v)  $\Gamma_s$  is none of the above. Orthogonality gives:

$$\begin{aligned} \alpha + \beta + \gamma + \delta &= 0 \\ \alpha + \beta - \gamma - \delta &= 0 \\ \alpha - \beta + \gamma - \delta &= 0 \\ \alpha - \beta - \gamma + \delta &= 0 \end{aligned}$$

and hence  $\alpha = 0$ , implying that  $\Gamma_s$  remains orthogonal to  $\Gamma_0$  in  $H_1$ ,  $H_2$  and  $H_{1 \times 2}$ . Hence,  $s$  does not block  $m_1$ ,  $m_2$ , or any combination of  $m_1$  and  $m_2$ .

This reasoning can be extended to apply Proposition 2 to any combination of nondegenerate mechanisms.

### 3. Mechanisms described by degenerate representations

When a mechanism is  $d$ -fold degenerate, the symmetry possibilities for distortion and blocking by states of self-stress are more involved, as the system can visit different subgroups of  $G$  by following different combinations of the  $d$  components of the mechanism. An established notation for the relations between the various groups is used, for example, in vibrational spectroscopy McDowell [1965], and can be used to frame some general remarks on how degenerate and nondegenerate mechanisms are blocked.

Let the irreducible representation of the mechanism be the  $d$ -fold degenerate  $\Gamma_{md}$ . The lowest symmetry group, reached by a generic combination of the  $d$  components of the mechanism, is the *kernel* of  $\Gamma_{md}$ . The kernel is an invariant subgroup of  $G$  and consists simply of those elements of  $G$  whose characters for  $\Gamma_{md}$  are equal to  $d$ . For any degenerate representation, the kernel is easily identified from the character table. In the kernel,  $\Gamma_{md}$  reduces to  $d$  copies of  $\Gamma_0$ . In the present context, it can be seen that, if no state of self-stress becomes totally symmetric in the kernel, then all combinations of the  $d$  components of  $m_d$  are finite mechanisms. Given that the kernel is not necessarily equal to the trivial group  $C_1$ , it is possible therefore for a system to support a number of states of self-stress that cannot block a given degenerate finite mechanism.

Unlike the nondegenerate case, the symmetries accessible to a degenerate mechanism are not necessarily restricted to the kernel group. By particular choices of combination, it may be possible to retain symmetry elements additional to those in the kernel, and thus produce configurations belonging to point groups of which the kernel is a subgroup. The accessible groups are the *cokernels* of  $\Gamma_{md}$ ; McDowell [1965] discusses the identification of cokernels, and lists them for the degenerate representations of a number of spectroscopically important point-groups.

The existence of cokernels for some degenerate representations widens the scope for finite degenerate mechanisms. Even in cases where the generic mechanism is blocked in the kernel, there may be combinations of the  $d$  components that access a cokernel in which no state of self-stress is totally symmetric, and by Proposition 1, those specific combinations will remain finite.

As an example, consider a hypothetical system of  $D_{6h}$  symmetry where  $\Gamma(m) = E_{2g}$  and  $\Gamma(s) = A_{2g}$ . The relevant rows of the  $D_{6h}$  character table are shown below.

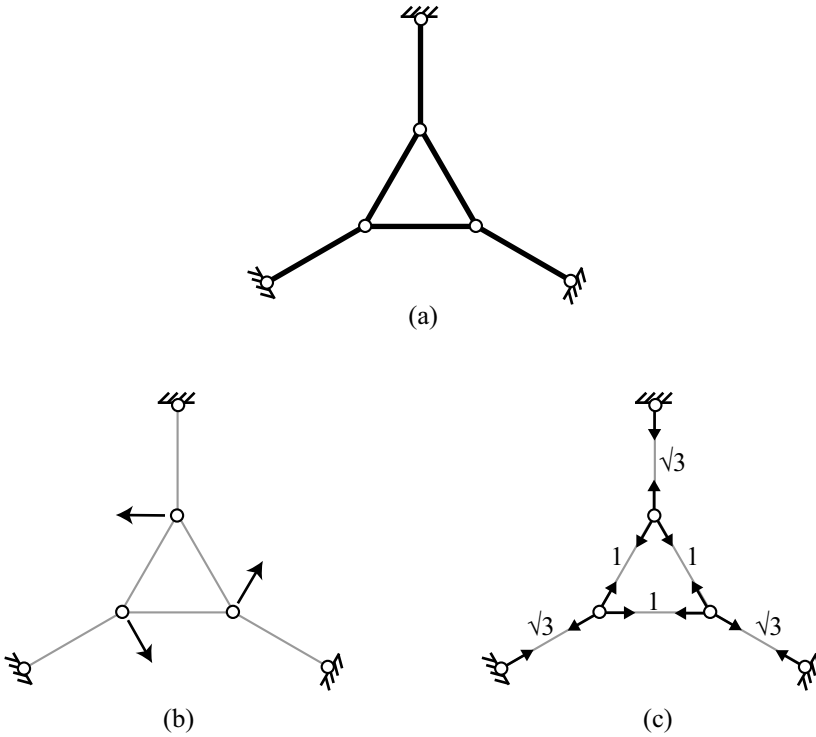
$D_{6h}$	$E$	$2C_6$	$2C_3$	$C_2$	$3C'_2$	$3C''_2$	$i$	$2S_3$	$2S_6$	$\sigma_h$	$3\sigma_d$	$3\sigma_v$
$A_{2g}$	1	1	1	1	-1	-1	1	1	1	1	-1	-1
$E_{2g}$	2	-1	-1	2	0	0	2	-1	-1	2	0	0

McDowell gives the kernel of  $E_{2g}$  as  $C_{2h}$ , and this can be confirmed by inspection of the table above, as the four columns with character  $+2$  are those for  $E$ ,  $C_2$ ,  $i$  and  $\sigma_h$ . It can also be seen by inspection that  $\Gamma(s) = A_{2g}$  becomes totally symmetric in  $C_{2h}$  and hence we cannot state that the pair of mechanisms is finite. However, the cokernel of  $E_{2g}$  is  $D_{2h}$  McDowell [1965], and as the table shows,  $A_{2g}$  is not totally symmetric in  $D_{2h}$  (four characters are  $+1$ , four characters are  $-1$  under these operations). Therefore it is guaranteed that the combination of components that lead from  $D_{6h}$  to  $D_{2h}$  is a finite mechanism.

### 4. Examples

**4.1. Structure stiffened by self-stress.** Figure 2 shows a planar pin-jointed framework that has been analysed by Kangwai and Guest [2000]. Considered in two dimensions, a structure with this connectivity





**Figure 2.** (a) A planar structure in which all mechanisms are stiffened by a state of self-stress; (b) the mechanism, showing directions of infinitesimal nodal displacement; (c) the state of self-stress, showing relative bar tensions.

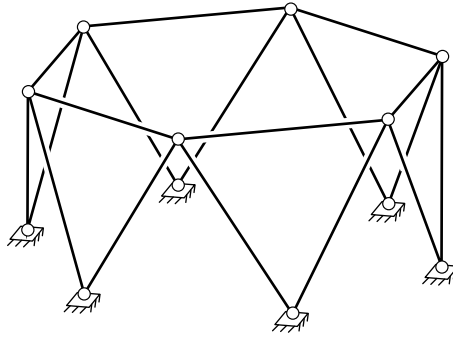
is generically both statically and kinematically determinate, but in the configuration shown has one state of self-stress and one mechanism. The planar structure has point group  $C_{3v}$ , with

$$\Gamma(m) = A_2, \Gamma(s) = A_1.$$

The single mechanism has the symmetry of an in-plane rotation of a central triangle, and the state of self-stress corresponds to a totally symmetric distribution of tensions in the bars. As the single state of self-stress is totally symmetric in  $C_{3v}$ , it can in principle stiffen any mechanism, and inspection, or a formal analysis of the tangent stiffness (see [Guest 2006], for example) shows that the mechanism is indeed stiffened.

We can also consider a structure in three dimensions that has the same set of connections. In a generic configuration, such a structure has three mechanisms, and no state of self-stress. Clearly these mechanisms must be finite. However, in the particular planar configuration shown, the structure attains  $D_{3h}$  symmetry, where it has a single state of self-stress and four mechanisms. The symmetry form of the Maxwell rule for pin-jointed frameworks [Fowler and Guest 2000] gives a full account, and yields

$$\Gamma(m) - \Gamma(s) = A'_2 + A''_2 + E''_2 - A'_1.$$



**Figure 3.** A ring structure with a finite mechanism.

As, by inspection,  $\Gamma(s) = A'_1 \equiv \Gamma_0$ , the four mechanisms span

$$\Gamma(m) = A'_2 + A''_2 + E''_2.$$

The state of self-stress is fully symmetric in this configuration, and hence can stiffen all mechanisms; analysis of the tangent stiffness shows that this stiffening is effective for all four mechanisms.

**4.2. Prestressable finite mechanism.** Figure 3 shows a classic example of a type of pin-jointed structure [Tarnai 1980] that satisfies Maxwell's rule for pin-jointed frames [Calladine 1978], but nevertheless admits a finite mechanism. The structure shown has a hexagonal ring of bars, connected in triangulated fashion to a rigid base. Its point group is  $C_{3v}$ , and as Kangwai and Guest [1999] have shown, the single mechanism has symmetry

$$\Gamma(m) = B_1$$

and the single state of self-stress has

$$\Gamma(s) = B_2.$$

It follows immediately from Proposition 2 that the mechanism is finite: there is neither an equisymmetric nor a totally symmetric state of self-stress here. The  $B_1$  mechanism leads to  $C_{3v}$  configurations where the state of self-stress has  $A_2$  symmetry.

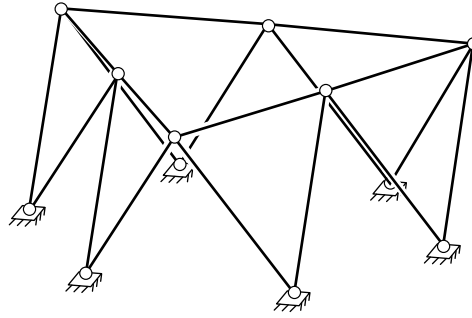
Following the finite mechanism eventually takes the structure to an interesting point of kinematic bifurcation, where the hexagon has degenerated into a triangle, as shown in Figure 4. At this point, a new pair of states of self-stress spanning the  $E$  representation emerges [Kangwai and Guest 1999], and hence  $\Gamma(m)$  becomes

$$\Gamma(m) = A_1 + E$$

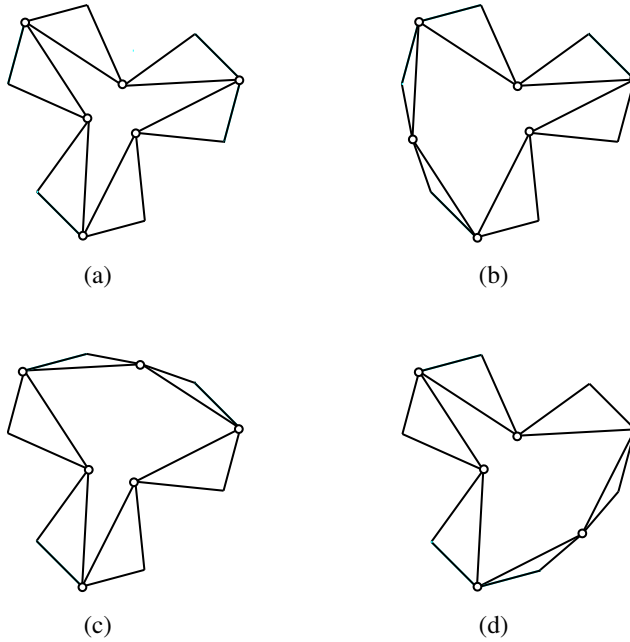
with

$$\Gamma(s) = A_2 + E.$$

The new states of self-stress do not affect the conclusion that there must be a finite  $A_1$  mechanism leading out of this configuration. However, we cannot deduce the existence of further finite mechanisms: the new states of self-stress are equisymmetric with the new mechanisms, and hence could stiffen generic combinations. In fact, in this case, there are three additional finite paths leading away from the bifurcation



**Figure 4.** The ring structure shown in Figure 3 displaced along the mechanism path until a point of kinematic bifurcation has been reached.



**Figure 5.** A plan view showing the finite paths leading out of the point of kinematic bifurcation shown in Figure 4; non-foundation joints are shown with a ring, foundation joints without a ring. The displaced structure in (a) retains  $C_{3v}$  symmetry; those in (b), (c), (d) each have  $C_s$  symmetry about one of the  $\sigma_v$  reflection planes of the  $C_{3v}$  geometry.

point, each of which retains  $C_s$  symmetry about one of the  $\sigma_v$  reflection planes of the  $C_{3v}$  geometry [Kumar and Pellegrino 2000].  $C_s$  is the cokernel of  $E$  in  $C_{3v}$ , whereas the kernel is the trivial group  $C_1$ . The paths are shown in Figure 5. Symmetry analysis shows only that stiffening of the mechanism is predicted, but not that it must occur. As always, symmetry is most powerful when showing that a phenomenon is forbidden, and hence detecting here when mechanisms *must* be finite, as blocking is not allowed, rather than when they *may* be infinitesimal, as blocking is permitted.

## References

- [Altmann and Herzig 1994] S. L. Altmann and P. Herzig, *Point-group theory tables*, Clarendon Press, Oxford, 1994.
- [Atkins et al. 1970] P. W. Atkins, M. S. Child, and C. S. G. Phillips, *Tables for group theory*, Oxford University Press, Oxford, 1970.
- [Calladine 1978] C. R. Calladine, “Buckminster Fuller’s “Tensegrity” structures and Clerk Maxwell’s rules for the construction of stiff frames”, *Int. J. Solids Struct.* **14**:2 (1978), 161–172.
- [Calladine and Pellegrino 1992] C. R. Calladine and S. Pellegrino, “Further remarks on first-order infinitesimal mechanisms”, *Int. J. Solids Struct.* **29**:17 (1992), 2119–2122.
- [Connelly and Servatius 1994] R. Connelly and H. Servatius, “Higher-order rigidity: what is the proper definition?”, *Discrete Comput. Geom.* **11**:2 (1994), 193–200. MR 94m:52027 Zbl 0793.52005
- [Fowler and Guest 2000] P. W. Fowler and S. D. Guest, “A symmetry extension of Maxwell’s rule for rigidity of frames”, *Int. J. Solids Struct.* **37**:12 (2000), 1793–1804.
- [Fowler and Guest 2005] P. W. Fowler and S. D. Guest, “A symmetry analysis of mechanisms in rotating rings of tetrahedra”, *Proc. R. Soc. A* **461**:2058 (2005), 1829–1846.
- [Garcea et al. 2005] G. Garcea, G. Formica, and R. Casciaro, “A numerical analysis of infinitesimal mechanisms”, *Int. J. Numer. Methods Eng.* **62**:8 (2005), 979–1012.
- [Guest 2006] S. D. Guest, “The stiffness of prestressed frameworks: a unifying approach”, *Int. J. Solids Struct.* **43**:3–4 (2006), 842–854.
- [Kangwai and Guest 1999] R. D. Kangwai and S. D. Guest, “Detection of finite mechanisms in symmetric structures”, *Int. J. Solids Struct.* **36**:36 (1999), 5507–5527.
- [Kangwai and Guest 2000] R. D. Kangwai and S. D. Guest, “Symmetry-adapted equilibrium matrices”, *Int. J. Solids Struct.* **37**:11 (2000), 1525–1548.
- [Kovács et al. 2004] F. Kovács, T. Tarnai, P. W. Fowler, and S. D. Guest, “Double-link expandohedra: a mechanical model for expansion of a virus”, *Proc. R. Soc. A* **460**:2051 (2004), 3191–3202.
- [Kumar and Pellegrino 2000] P. Kumar and S. Pellegrino, “Computation of kinematic paths and bifurcation points”, *Int. J. Solids Struct.* **37**:46 (2000), 7003–7027.
- [Kuznetsov 2000] E. N. Kuznetsov, “On the physical realizability of singular structural systems”, *Int. J. Solids Struct.* **37**:21 (2000), 2937–2950.
- [McDowell 1965] R. S. McDowell, “On the symmetry of normal modes”, *J. Mol. Spectrosc.* **17**:2 (1965), 365–367.
- [Salerno 1992] G. Salerno, “How to recognize the order of infinitesimal mechanisms: a numerical approach”, *Int. J. Numer. Methods Eng.* **35**:7 (1992), 1351–1395.
- [Salthouse and Ware 1972] J. A. Salthouse and M. J. Ware, *Point group character tables and related data*, Cambridge University Press, 1972.
- [Tarnai 1980] T. Tarnai, “Simultaneous static and kinematic indeterminacy of space trusses with cyclic symmetry”, *Int. J. Solids Struct.* **16**:4 (1980), 347–359.
- [Tarnai 1989] T. Tarnai, “Higher-order infinitesimal mechanisms”, *Acta Tech. Hung.* **102** (1989), 363–378.
- [Tarnai and Szabó 2000] T. Tarnai and J. Szabó, “On the exact equation of inextensional, kinematically indeterminate assemblies”, *Comput. Struct.* **75**:2 (2000), 145–155.

Received 2 Dec 2005. Accepted 21 Oct 2006.

SIMON D. GUEST: [sdg@eng.cam.ac.uk](mailto:sdg@eng.cam.ac.uk)

Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom

PATRICK W. FOWLER: [P.W.Fowler@sheffield.ac.uk](mailto:P.W.Fowler@sheffield.ac.uk)

Department of Chemistry, University of Sheffield, Sheffield S3 7HF, United Kingdom

## SINGLE MEMBER ACTUATION OF KAGOME LATTICE STRUCTURES

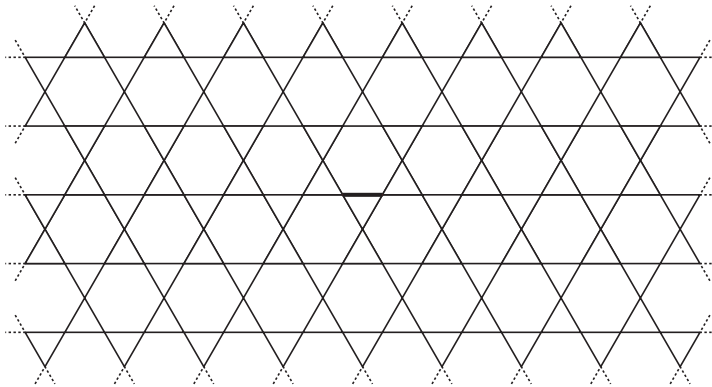
ANTHONY C. H. LEUNG AND SIMON D. GUEST

The two-dimensional kagome lattice has been shown to be a promising basis for active shape-changing structures, having both low actuation resistance and high passive stiffness. Activation of some members results in a global macroscopic shape change. Small deformation models show that the kagome lattice's properties are critically dependent on its initial geometry. This paper investigates the fundamental actuation properties of a kagome lattice subject to single-member actuation, particularly when geometric nonlinearity is introduced with large actuation strains. Actuation resistance is found to be lowered with expansive actuation; a limiting peak actuation stiffness is observed when the actuator is flexible. Conversely, actuation resistance is found to increase with contractile actuation.

### 1. Introduction

Recent studies have shown the two-dimensional kagome lattice, shown in Figure 1, to be a prime candidate to form the backbone of *active* structures [Santos e Lucato et al. 2004; Hutchinson et al. 2003; Symons et al. 2005a; 2005b]. If linear actuators replace some members of the truss, significant global macroscopic shape changes can be achieved with a relatively small number of actuators. In order to obtain a deeper understanding of the response of the kagome lattice to actuation, this paper investigates a simple case, where an infinite kagome lattice is activated by a single actuator.

Active structures, also known as *adaptive* or *smart* structures, change their configuration, shape or other properties to better carry the type and magnitude of its design loads. Shape changes in active



**Figure 1.** Part of an infinite kagome lattice. The thick line at the centre represents a single actuator.

*Keywords:* kagome lattice, morphing, actuation.

structures are achieved by actuation of some parts of the structure. To date, actuators manufactured are predominantly one-dimensional, where their capabilities are confined to extending and contracting along or rotating about one axis; *active truss structures* provide a way of incorporating linear actuators into a structure. The paper will focus on the structural response of the truss rather than methods of actuation; possible actuators are described in [Santos e Lucato et al. 2004; Symons et al. 2005a; 2005b].

Hutchinson et al. [2003] has shown that the planar kagome lattice can be actuated with minimal internal resistance while maintaining overall stiffness. The rigid-jointed planar kagome lattice exhibits high passive stiffness, while having a low actuation resistance. These properties are ‘inherited’ from the pin-jointed ‘parent’ [Deshpande et al. 2001], where the kagome truss is composed of rigid triangulated unit-cells, with mechanisms allowing relative motion of the cells; as member stockiness,  $s$ , is reduced, the properties of the rigid-jointed structure approaches the properties of the pin-jointed structure. The stockiness is a nondimensional measure of the aspect ratio of each member, defined as  $s = k/L$ , where  $k$  is the in-plane radius of gyration of each member and  $L$  is the length of the member ( $s$  is the reciprocal of the common mechanics parameter *slenderness*).

Wicks and Guest [2004] have shown, using various *linear* finite element and analytic models, that the exceptional behaviour of the kagome lattice depends on both flexure of members, and axial contraction of the bars colinear with the actuator. However, this behaviour depends crucially on the geometry of the lattice, and hence kagome lattices subjected to large actuation will be prone to *geometrically nonlinear* effects.

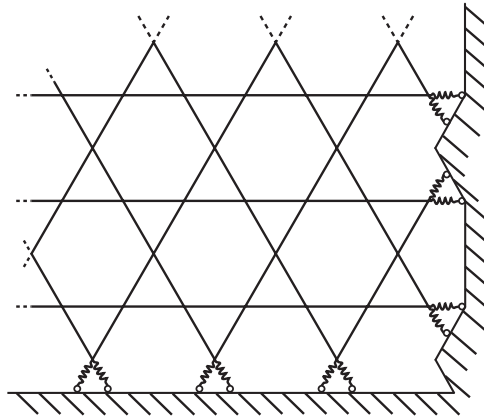
This paper investigates geometric nonlinearity in the single-bar actuation behaviour of kagome lattices with stockiness ranging from 0.001 to 0.05 (an areal density of  $\simeq 0.3\% \rightarrow 15\%$ ), covering the likely range of practical interest, with actuation strains of up to  $\pm 50\%$ . The two experimental active kagome lattice structures described in the literature have  $s = 0.007$  [Santos e Lucato et al. 2004] and  $s \simeq 0.0085$  [Symons et al. 2005a; 2005b].

The paper is structured as follows. Section 2 will describe the computational model used in the finite element analyses. Section 3 will describe the deformed mode shapes of the lattice following actuation, and Section 4 will investigate the build-up of force in the actuator. Section 5 explores the actuation limit due to material yield and member buckling. Section 6 provides a discussion of the results.

## 2. Computational model

The planar kagome lattice is a class of *repetitive truss* structure. Noor [1998] reviewed analysis methods for repetitive trusses and described the following four classes of approach, with increasing computational expense:

- (1) The substitute continuum approach which the truss considered is replaced by a continuum with resembling structural properties [Timoshenko and Gere 1961].
- (2) The periodic structure approach, which employs a transfer matrix relating a state vector on the boundaries of the unit cell [Noor and Zhang 2006].
- (3) The direct field method, which uses finite difference equations on unit cells [Dean 1976; Renton 1984].
- (4) The direct or finite element method, which discretizes the entire structure [Akin 1986].



**Figure 2.** The spring boundary condition; boundary nodes of the kagome lattice are connected to spring elements to simulate the response of an infinite grid.

or a combination of the above [Karpov et al. 2002]. This paper uses the finite element method as the basis of its analyses. Computational time has been substantially reduced by the use of symmetry and an elastic boundary conditions to mimic infinity, making the entire array of result from FEM feasible.

The finite element package ABAQUS was used to model the structure in two and three dimensions. The model used was designed to represent an infinite lattice: to do this a finite rectangular lattice, with a width 80 times the individual member length  $L$ , and a height of  $30 \times L\sqrt{3}/2 \approx 26L$ , was used together with a spring boundary condition, as shown in Figure 2. The derivation of the required spring stiffness to represent the rest of an infinite lattice is given in the Appendix by assuming linearity, and decoupling between different boundary points; the required stiffness turns out to be proportional to the stockiness,  $s$ . Computational effort is reduced by the assumption that the response is symmetric about the central vertical line.

Two extreme cases are considered for the stiffness of the actuator: either that the actuator has the same elastic properties as every other bar, which we refer to as the ‘flexible’ actuator; or the actuator is a pin-jointed infinitely stiff bar that imposes a fixed distance between its ends, which we refer to as the ‘stiff’ actuator. Apart from the actuator, all other bars in the lattice have identical properties. The actuator is either axially elongated, defined here as positive actuation, or contracted, defined here as negative actuation. Flexible actuation is achieved by raising or lowering the overall temperature in the system and using an actuator that has a coefficient of thermal expansion of unity, while the rest of the structure has a coefficient of thermal expansion of zero. Stiff actuation is achieved by increasing or decreasing the distance between the two ends of the actuator.

Each bar of the lattice except the actuator is modelled with four 3-noded Timoshenko (shear flexible) beam elements; the flexible actuator is modelled with eight such elements, while the rigid actuator is modelled by imposing a relative displacement. Test cases showed that using a higher number of elements per bar gave results that were indistinguishable from those reported here. No consideration is given to the actual size of a joint between bars, which is assumed to be at a point. For most analyses, each bar has a circular cross-section of radius  $r$ . The material properties are assumed to remain linear, with a Young’s modulus  $E$ , giving an axial stiffness  $AE$ , where  $A = \pi r^2$ , and flexural stiffness  $EI$ , where

$I = Ak^2 = Ar^2/4$ ;  $k = r/2$  for a round bar. Bars with rectangular cross-sections are used for out-of-plane buckling analyses, with area  $A = bd$ , where  $b$  and  $d$  are out-of-plane and in-plane dimensions respectively. Similarly, this gives axial and in-plane flexural stiffness  $AE$  and  $EI = EAk^2$ , but with  $k = d/2\sqrt{3}$ .

Lattices were considered with stockiness  $s$  ranging from 0.001 to 0.05, which covers the range likely to be of interest for application. We describe the actuation in terms of an *actuation strain*,  $\varepsilon_a$ : we define this strain to be the extension the bar would experience if unconstrained, divided by its original length; it is defined as being positive when the bar gets longer. The actual strain experienced in the flexible actuator,  $\varepsilon$ , is smaller than the actuation strain, as the rest of the structure imposes an axial force  $F$  to the actuator, which this gives rise to an elastic strain of opposite sign to  $\varepsilon_a$ . With a stiff actuator, the actuation strain equals the elastic strain:  $\varepsilon = \varepsilon_a$ .

### 3. Mode of deformation

This section will separately present the shape of the deformed lattice following expansive and contractile actuations. In each case, the geometrically nonlinear results for both the stiff and flexible actuator are compared with the results of a linear model. The comparisons are done for an actuation strain of  $\pm 50\%$ , and for three different values of stockiness. It is assumed that the grid has no imperfection, and only in-plane deformation are considered — Section 5.3 will revisit these assumptions.

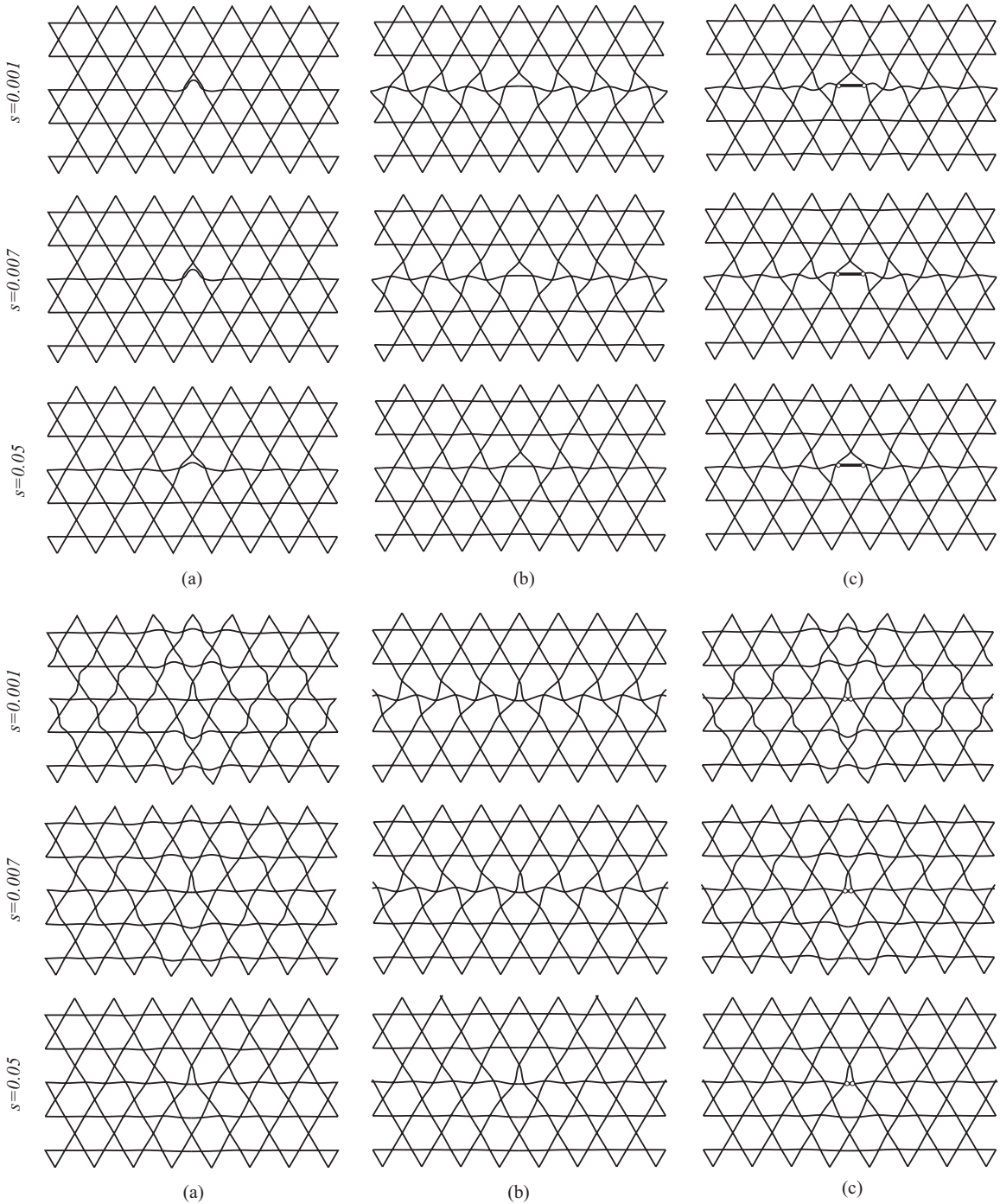
The mode of deformation for a linear calculation was characterized by Wicks and Guest [2004]; essentially identical results are presented here for the linear case. The key features of the linear deformation mode can be observed in the results shown in the middle panels of Figure 3, top and bottom. Deformation is confined largely to a corridor parallel with the actuator. The deformation dies away in an approximately exponential manner, and the rate of decay is proportional to the stockiness,  $s$ . The linear results provide a useful base to compare the geometrically nonlinear results in the left and right panels of Figure 3 (top and bottom).

**3.1. Nonlinear actuation: expansion.** A key feature of the linear response is that lengthening of the actuator can only be achieved by shortening the other members in line with the actuators, leading to the interesting combined bending and stretching responses observed in [Wicks and Guest 2004], with an exponential decay in deformation with distance from the actuator. However, once the initial straight lines of bars has been bent by actuation, further shortening can occur by bending alone.

The deformed shape of the lattice calculated using the geometrically nonlinear model, for  $\varepsilon_a = 50\%$ , is shown in Figure 3, top left, for the flexible actuator, and Figure 3, top right, for the stiff actuator. Both sets of results show increased localisation of the deformation when compared with the linear calculation shown in Figure 3, top middle, and this increased localisation is greater for smaller stockiness  $s$  (but note that the linear results themselves show *less* localisation as  $s$  is decreased). The key observation, however, is that for the flexible actuator, the deformation is almost entirely localised in the actuator itself.

**3.2. Nonlinear actuation: contraction.** As for expansion, in the linear regime, the shortening of the actuator can only be achieved by lengthening of the other members in line with the actuator. However, unlike during actuator expansion, bending of these bars due to actuation does not then make deformation of other members in line with the actuator any easier.





**Figure 3.** Deformed kagome lattices under positive actuation for  $\varepsilon_a = 0.5$  (top) and negative actuation for  $\varepsilon_a = -0.5$  (bottom), as well as different values of bar stockiness: (a) nonlinear actuation with a flexible actuator, (b) linear calculation and (c) nonlinear actuation with a stiff actuator.

The deformed shape of the lattice calculated using the geometrically nonlinear model, for  $\varepsilon_a = -50\%$ , is shown in Figure 3, bottom left, for the flexible actuator, and Figure 3, bottom right, for the stiff actuator. The linear calculation results shown in Figure 3, bottom middle, are identical to the expansion case but with a sign change for deformation.

Both sets of geometrically nonlinear results show decreasing localisation as stockiness is reduced. In the linear case, deformation is primarily confined within the corridor of bars colinear to the actuated member; but in the geometrically nonlinear cases the deformation spreads to neighbouring corridors. Unlike the expansion case, the results appear almost identical for the stiff and flexible actuators.

#### 4. Actuation forces

This section will describe the force,  $F$  (defined as compression positive), developed in the actuator as it is activated. For small deformations, when the bar is approximately straight,  $F$  is nearly constant along the length of the actuated bar; for larger deformations we define  $F$  to be an average of the varying force along the bar.

A nondimensional form of the force can be defined by comparing the force with that required to stretch a bar: this formulation was implicit in [Wicks and Guest 2004] and explicit in [Leung et al. 2004]. For consistency, we will retain this formulation, and define

$$\hat{F} = \frac{F}{EA}$$

where  $EA$  is the axial stiffness of the actuator.

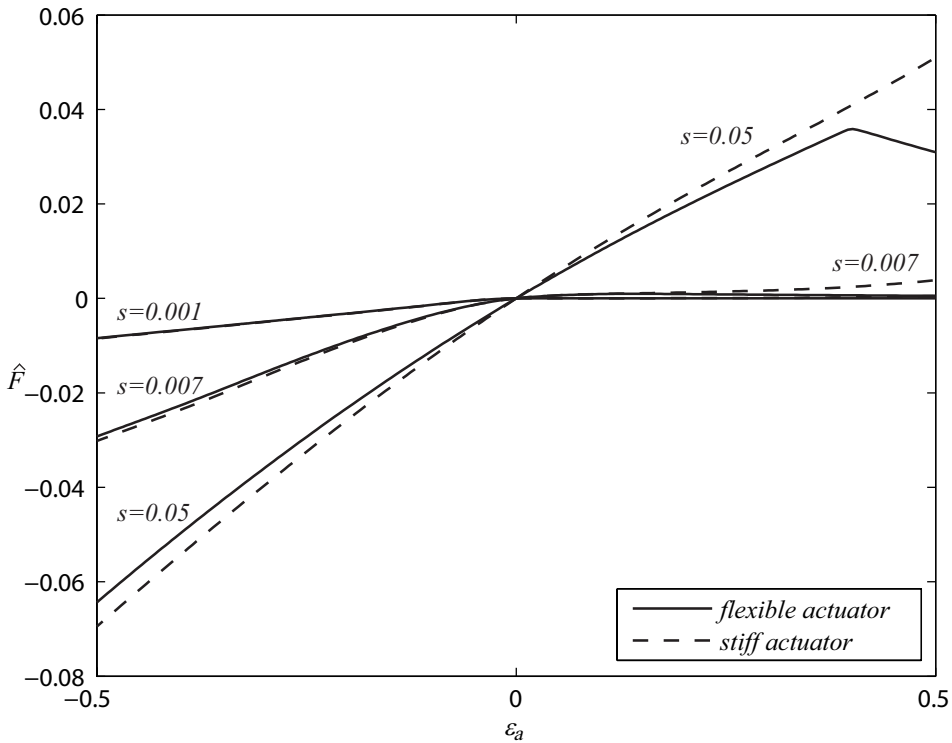
Figure 4 shows a plot of actuation force against actuation strain for three values of stockiness, for both the stiff and flexible actuator. In all cases, the plots show a softening response to expansion of the actuator (+ve  $\varepsilon_a$ ), and a stiffening response to contraction (-ve  $\varepsilon_a$ ). For expansion of the actuator the results for the flexible actuator clearly show greater softening. In fact, in all three plots the actuation force  $\hat{F}$  reaches a peak, but this can only be clearly seen in Figure 4 for  $s = 0.05$ , where  $\hat{F}$  reaches a peak at  $\varepsilon_a = 40\%$ . For the stiff actuator,  $\hat{F}$  does not reach a peak.

The results in Figure 4 are difficult to compare because of the large differences in magnitude of the results. However, an alternative nondimensional plot allows better comparison. Consider comparing  $F$  with the force required to bend a bar in the lattice. For any bending deformation, the force applied for a displacement  $\delta$  will be given by  $(cEAk^2/l^2) \times (\delta/l)$ , where  $c$  is a constant depending on the support condition. Thus the appropriate nondimensional form of  $F$  is

$$\frac{F}{EAk^2/l^2} = \frac{F}{EAs^2} = \frac{\hat{F}}{s^2}$$

and it is this form that we shall use for an alternative plotting of the results.

The values of  $\hat{F}/s^2$  during actuation are shown in Figure 5, for the flexible actuator only. Note that the abscissa used here is  $\varepsilon_a/s$  — effectively measuring actuation displacements relative to the thickness, rather than the length, of the the bar. It was found in [Wicks and Guest 2004] that, for a kagome lattice, the nondimensional energy of actuation, or equivalently the nondimensional actuation stiffness,  $d\hat{F}/d\varepsilon_a$  for small  $\varepsilon_a$ , was proportional to  $s$  [Leung et al. 2004]. Hence the choice of abscissa,  $\varepsilon_a/s$ , and the choice of ordinate,  $\hat{F}/s^2$ , ensures that the slope of the lines for  $\varepsilon_a = 0$  is approximately equal for all  $s$ . These



**Figure 4.** Nondimensional actuation force calculated using geometrically nonlinear models, for both flexible and stiff actuators, and  $s = 0.001$ ,  $s = 0.007$  and  $s = 0.05$ .

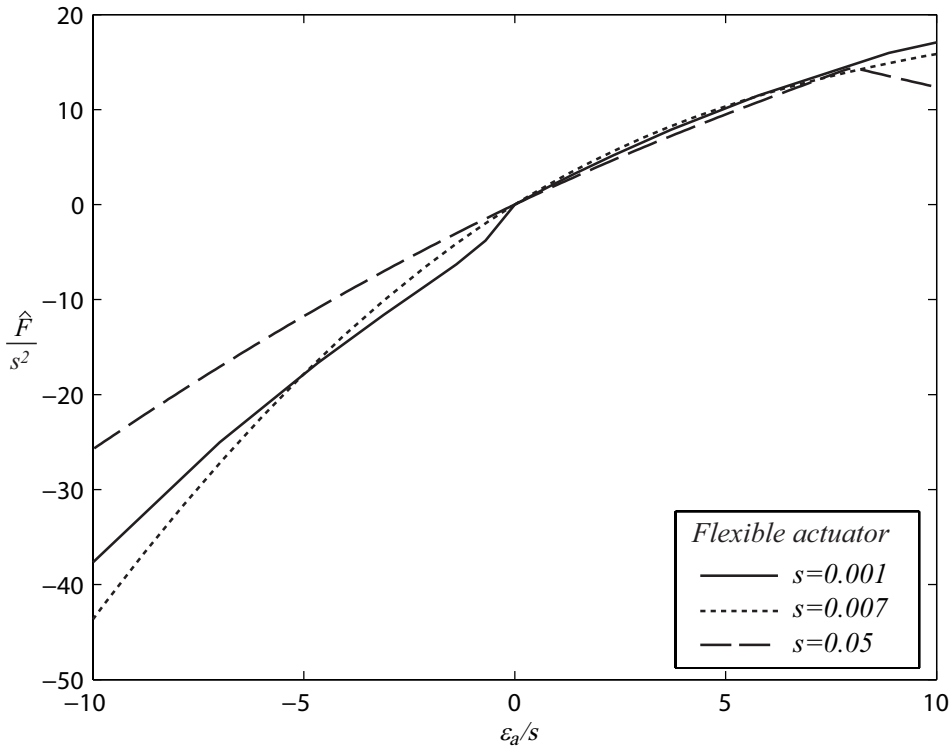
features can be seen in Figure 5. Note that in Figure 5, the maxima in  $\hat{F}$  for  $s = 0.001$  and  $s = 0.007$  now occur outside the range of data plotted. Results for the stiff actuator in this range are rather similar (apart from the absence of the peak for  $s = 0.05$ ) and are not shown for clarity.

## 5. Actuation limit

We consider two limiting cases for actuation. The first is when actuation is limited by the material properties of the lattice material: we consider the structure ineffective when a yielding strain is reached somewhere in the lattice. An alternative limit is when the actuator is not able to apply an increased force to the deform the rest of the lattice — either due to a classical buckling phenomenon, or because deformation has localised within the actuator.

**5.1. Strain limited actuation.** We consider limitation on the actuation of a structure that yielding anywhere in the structure outside of the actuator must be avoided. This section reports numerical results for the peak achievable actuation strain in infinite kagome lattices.

*Flexible actuator.* Figure 6 shows the magnitude of actuation strain,  $\epsilon_a$ , at which first yield occurs plotted against stockiness,  $s$ , for a flexible actuator. Various values of material yield strain,  $\epsilon_y$ , for both positive and negative actuation are included, covering a practical range of yield strains from material such as mild steel ( $\epsilon_y \approx 0.125\%$ ) to titanium ( $\epsilon_y \approx 2.0\%$ ). In each case, the yield strain is reached at some position in



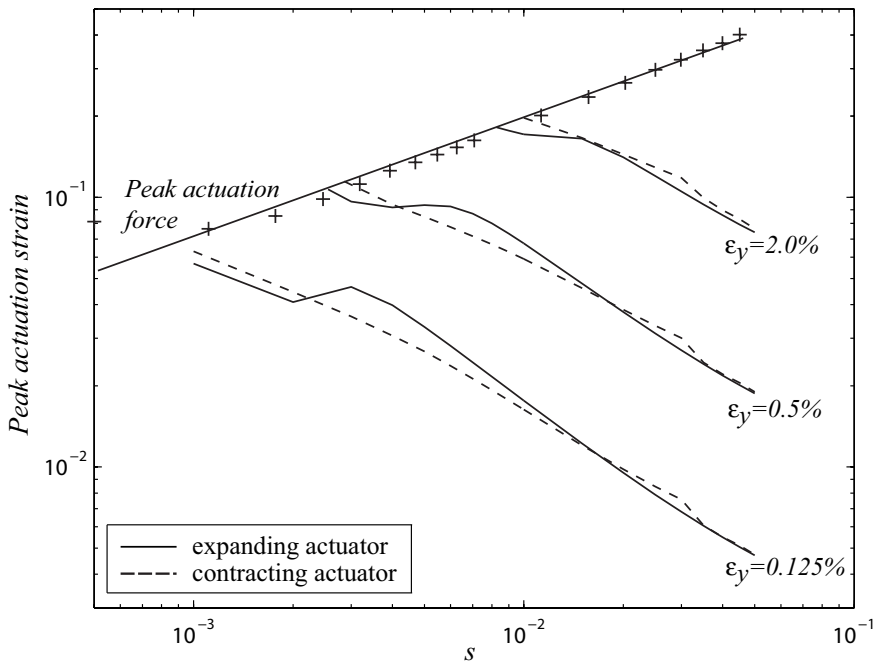
**Figure 5.** Nondimensional actuation force within the actuated bar plotted against the imposed strain, for only the flexible actuator, and stockiness  $s = 0.001$ ,  $s = 0.007$  and  $s = 0.05$ . Note this is the same data plotted in Figure 4, with different scaling.

the bar immediately adjacent to the actuator, as shown in Figure 8(a). The position of first yield varies with the stockiness of the lattice.

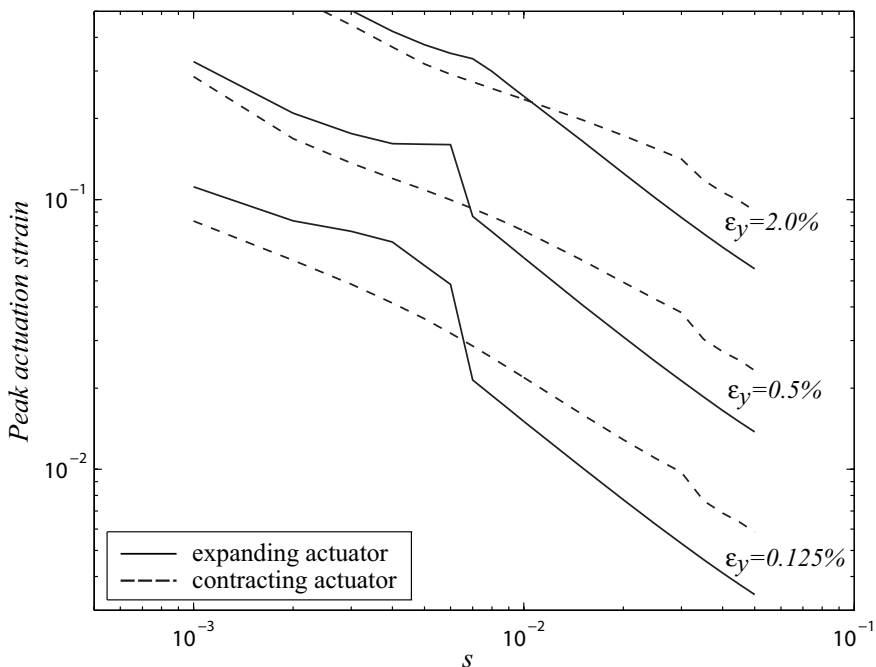
The results are shown only up to  $\epsilon_a = 0.5$  and are truncated by the line of peak actuation force, marked by crosses in Figure 6. Peak force achieved in a flexible actuator indicates a limit on the practicality of the kagome lattice as an active structure. Peak force occurs at an actuation strain proportional to the stockiness of the lattice: the results are also plotted in Figure 9 and discussed in Section 5.2.

The result from Figure 6 shows for smaller stockiness and higher yield strains, geometric nonlinearity affects maximum actuation strain most severely. For the smallest yield strain case  $\epsilon_y = 0.125\%$ , it is observed that when stockiness decreases in the positive actuation case, peak actuation is suddenly lowered from the gradually increasing trend. The same effect is observed for other  $\epsilon_y$ . In the negative actuation cases, no sudden drop in achievable actuation strain is observed.

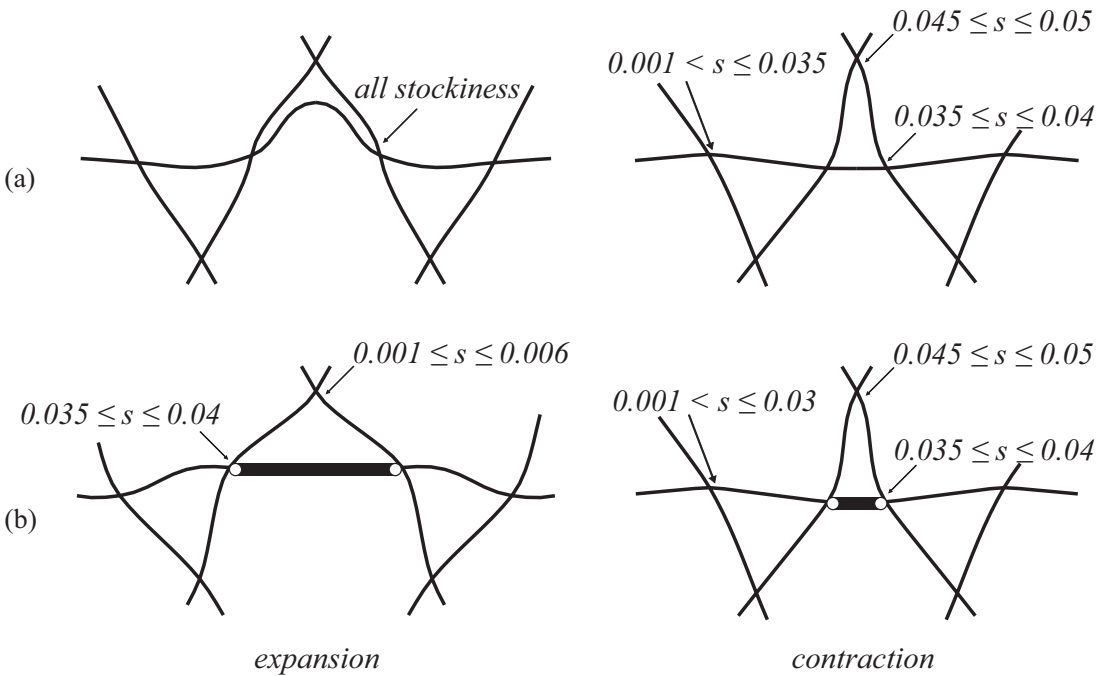
*Stiff actuator.* Figure 7 shows the magnitude of actuation strain,  $\epsilon_a$ , at which first yield occurs plotted against stockiness,  $s$ , for a stiff actuator. Three material yield strains,  $\epsilon_y$ , from 0.125% to 2.0% are included, for both positive and negative actuation. In each case, the yield strain is reached at some position in the bar immediately adjacent the actuator, as shown in Figure 8(b). The position of first yield varies with the stockiness of the lattice.



**Figure 6.** Peak actuation strain achievable for the flexible actuator before either the material yields, or a peak force is reached in the actuator.



**Figure 7.** Peak actuation strain achievable for the stiff actuator before either the material yields, or a peak force is reached in the actuator.

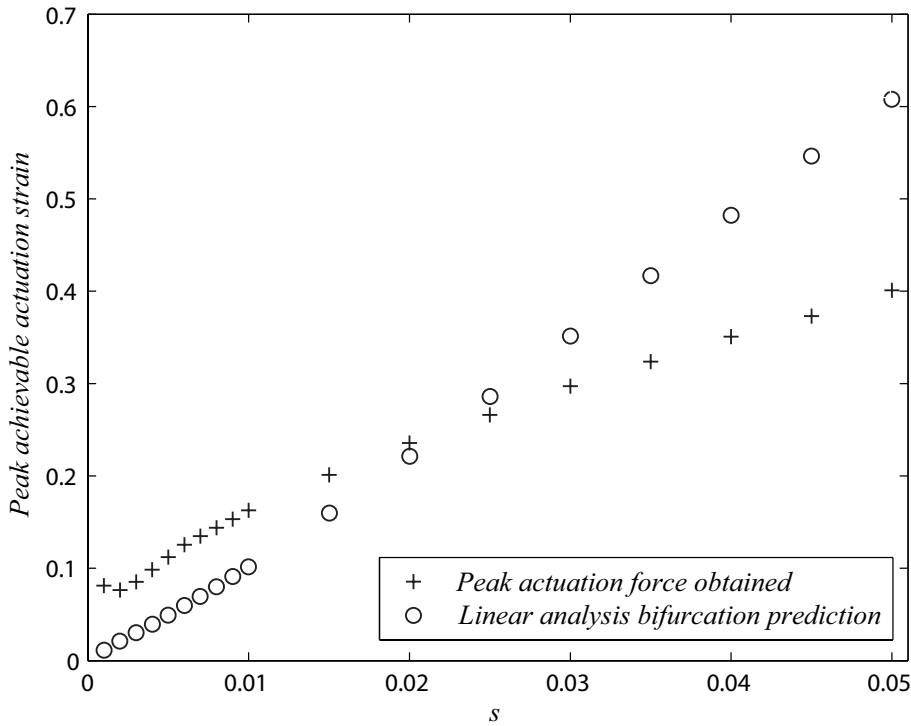


**Figure 8.** The position of the peak strain in the kagome lattice due to the actuation of a single bar, indicated by arrows, with (a) flexible, and (b) stiff actuation.

**5.2. Peak actuation force and buckling.** In Figure 6, the straight line marked by crosses shows the actuation strain at which peak actuation force in the actuator is achieved. This peaking effect can be seen in Figure 4 for  $s = 0.05$ : the solid line representing flexible positive actuation peaks at approximately  $\epsilon_a = 0.4$ . This critical actuation strain exists for each of the models with a flexible actuator, and its value increases approximately linearly with the stockiness of the lattice. These data is also shown in Figure 9. When actuation force ceases to increase with actuation strain, the maximum activation capability of the lattice is reached. This phenomenon is only observed in the positive flexible actuation case, and we consider it as a limiting case of actuation of the kagome lattice.

One way to predict the peak actuation force would be to consider a bifurcation analysis based on a linear model. However, for this case, this is not found to give good results, as is shown in Figure 9: the structure undergoes large deformations before the peak force is obtained, and this renders a prediction based on a linear model inaccurate. Another possible failure mode is out-of-plane buckling, which is investigated in the following section.

**5.3. Buckling of an imperfect structure .** The results described thus far are for a perfect planar structure. For most practical uses of a kagome lattice, deformation out of plane is likely to be restrained [Santos e Lucato et al. 2004; Symons et al. 2005a; 2005b]. However, we do here consider the possible out-of-plane buckling by considering a full three-dimensional structure with imperfections. This has the useful side-effect of showing that the imperfections barely affect the in-plane response.



**Figure 9.** Comparison of limiting actuation strain at peak actuation force and buckling load.

Random imperfection has been implemented into the system to invoke any possible buckling modes. The magnitude of imperfection,  $\delta$ , in a particular direction, is a function of radius of gyration, and is defined as

$$\delta_x, \delta_y, \delta_z = P * R * 2k,$$

where  $P$  is the imperfection factor: values of 0.01%, 0.1% and 1.0% were used.  $R$  is a normally distributed random number between 0 and  $\pm 1.0$ ; and  $k$  is the radius of gyration of the model. These imperfections are applied on all the member nodes of a kagome lattice model, except the actuator. Results show that no in-plane buckling response occurs. The in-plane response is indistinguishable from the ‘perfect’ results described in Section 4. Thus the following will focus on the effect of out-of-plane imperfections,  $\delta_z$ , on the out-of-plane buckling response.

The response of the imperfect structure to actuation is plotted in Figure 10 for models with circular bars of  $s = 0.01$  and flexible actuator, together with results for the perfect structure for comparison. The response of the imperfect structure initially matches that of the perfect structure, but at some point, depending on the imperfection factor,  $P$ , there is a sudden drop in the actuator load, and the structure deforms into a plate buckling like deformed mode. If the actuation is reversed for a buckled structure, the structure follows this post-buckling path, and does not return to the initial curve until the actuation reaches zero.

For cases when there is no out-of-plane restraint, the out-of-plane buckling can still be suppressed by increasing the out-of-plane bending stiffness of the bars. Models with rectangular cross-section bars have been analyzed, with constant imperfection factor and a flexible actuator. To maintain a constant  $s$ , the in-plane dimension,  $d$ , of the members must remain constant, while the out-of-plane dimension,  $b$ , is altered. Figure 11 shows results from changing  $b$ , while keeping,  $d$ ,  $s$  and the imperfection factor constant;  $b$  ranges from  $0.5d$  to  $4d$ . By increasing one dimension of the bar cross-section, the actuation response approaches that of the perfect model.

The results above have all been for a flexible actuator. Essentially similar results were obtained for a stiff actuator.

## 6. Discussion

It is found that positive actuation causes a decrease in actuation stiffness. This effect is especially prominent at small stockiness because of the smaller member bending stiffness. For flexible actuators only,  $\hat{F}$  reaches a peak value.

Negative actuation results in an increased actuation stiffness, for both flexible and stiff actuation, indicating a stiffening response. This is due to a more stretching dominated response invoked by a contracting actuator.

The mode of deformation can provide explanation for the softening and stiffening effects under single-member actuation. The softening response observed with positive flexible actuation is caused by localised deformation within the actuator itself. Figure 3, top left, shows severe bending deformation in the flexible actuator when stockiness is low. It shows that deflection attenuation is exceptionally rapid for the smallest stockiness  $s = 0.001$ , where almost all deformation is found within the actuator and the colinear bars immediately adjacent. For positive stiff actuation, Figure 3, top right, shows a similar response where local deformation is concentrated near the actuator. As these bars deform quickly and locally with increasing actuation, less effective actuation can be transmitted to the rest of the structure.

Negative actuation causes deflection to propagate in a direction perpendicular to the actuator, especially when stockiness is small. In Figure 3, bottom middle, where actuation is small and linear, deflection is largely confined within the actuator's corridor. With large actuation, Figure 3, bottom left and right, shows that more diagonal bars are affected, and this contributes to the stiffening effect observed.

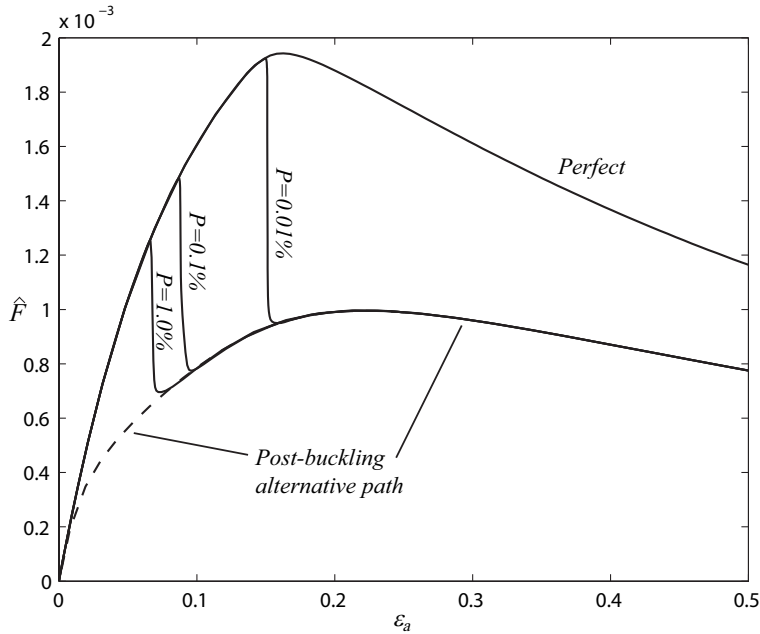
Figure 4 shows the discrepancies between flexible and stiff actuation models of the same stockiness. The sudden decrease in stiffness observed in flexible actuation models are not observed in the stiff actuation models. This shows that the peak in actuation force, described in the above discussion, is the result of the properties of the actuator deforming in bending. Kagome lattices can be made more effective for positive actuation if the actuator is stiffer.

Finally, actuation is limited only by material strength and the limitation on actuation force. In-plane buckling proves not to be an important consideration.

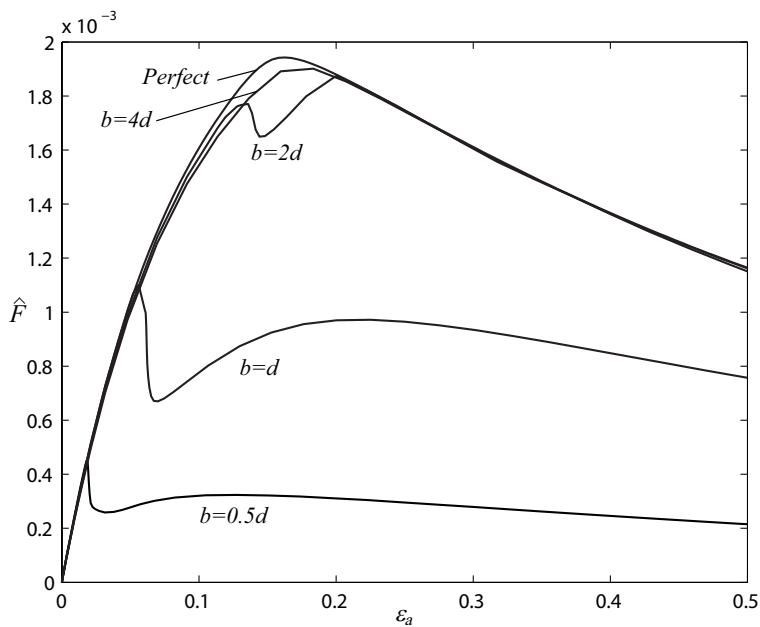
### Appendix: Spring boundary stiffness derivation

The springs at the boundary of the modelled lattice should represent an infinite continuation of the kagome lattice. As long as the modelled lattice is large enough, the deformations at the boundary will be small, and hence the result of a linear model gives an appropriate boundary spring stiffness.

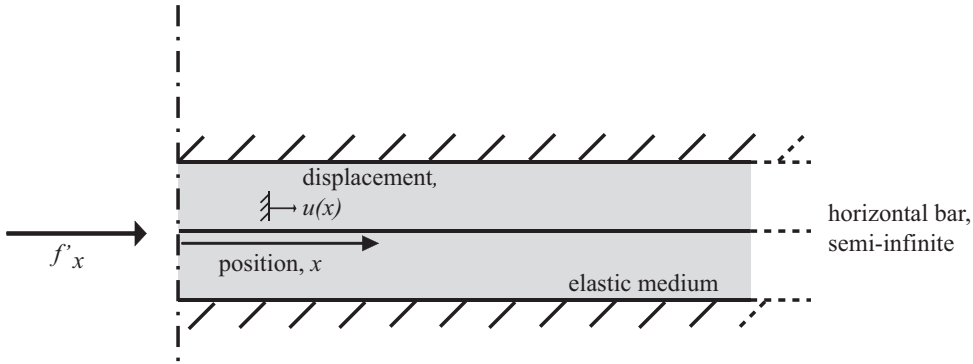




**Figure 10.** Actuation force versus actuation strain for kagome lattice models of  $s = 0.01$  with a flexible actuator and circular bars. Results are shown for three imperfection factors,  $P$ . The dashed line indicates reverse loading. The ‘perfect’ result from 2D modeling is included for comparison.



**Figure 11.** Actuation force versus actuation strain for kagome lattice models of  $s = 0.01$  with a flexible actuator and rectangular bars. The in-plane dimension  $d$  is kept constant, and the out-of-plane dimension  $b$  varies. The ‘perfect’ result is included for comparison.



**Figure 12.** Smeared stiffness model, where a bar is suspended in an elastic material; adapted from [Wicks and Guest 2004].

Here, we adapt the smeared stiffness model shown in Figure 12 to find the boundary stiffness. Now the force  $f'_x$  represents the force at the edge of the modelled lattice. Wicks and Guest [2004] show that

$$u(x) = u_0 e^{-cx}, \quad (1)$$

where  $u(x)$  is the horizontal nodal displacement of colinear bar members,  $u_0$  is the displacement at  $x = 0$ , and a simple estimate gives  $c = 4\sqrt{3}s/L$ . If we consider that the central bar remains elastic, with axial stiffness  $AE$  and strain  $du/dx$ , its tension  $t(x)$  is given by

$$t(x) = AE \frac{du}{dx} = -cAEu_0 e^{-cx},$$

so the central force  $f'_x$  is obtain from the tension at the  $x = 0$  position:

$$f'_x = -t(0) = cAEu_0. \quad (2)$$

This response can be represented by a spring of stiffness  $k$ ; combining (1) and (2) we get

$$k = \frac{f'_x}{u_0} = cAE = 4\sqrt{3}s \frac{AE}{L}.$$

In fact, this estimate can be improved by considering linear finite element models with a large range of stockiness, to give a more suitable accurate relationship, where  $k_{\text{empirical}} = 0.6134k$ . It was found that using boundary springs with this stiffness gave essentially identical results to those described in this paper with a modelled lattice that was either double or half the original model's width and height, showing that the boundary was correctly modelled.

## References

- [Akin 1986] J. E. Akin, *Finite element analysis for undergraduates*, Academic Press, London, 1986.
- [Dean 1976] D. L. Dean, *Discrete field analysis of structural systems*, Springer, New York, 1976.
- [Deshpande et al. 2001] V. S. Deshpande, M. F. Ashby, and N. A. Fleck, "Foam topology bending versus stretching dominated architectures", *Acta Mater.* **49**:6 (2001), 1035–1040.

- [Hutchinson et al. 2003] R. G. Hutchinson, N. Wicks, A. G. Evans, N. A. Fleck, and J. W. Hutchinson, “Kagome plate structures for actuation”, *Int. J. Solids Struct.* **40**:25 (2003), 6969–6980.
- [Karpov et al. 2002] E. G. Karpov, D. L. Dorofeev, and N. G. Stephen, “Characteristic solutions for the statics of repetitive beam-like trusses”, *Int. J. Mech. Sci.* **44**:7 (2002), 1363–1379.
- [Leung et al. 2004] A. C. H. Leung, D. D. Symons, and S. D. Guest, “Actuation of kagome lattice structures”, in *Proceedings of the 45th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference* (Palm Springs, CA), AIAA, Reston, VA, 2004. Paper #2004-1525.
- [Noor 1998] A. K. Noor, “Continuum modelling for repetitive lattice structures”, *Appl. Mech. Rev.* **41** (1998), 285–296.
- [Noor and Zhang 2006] N. G. Noor and Y. Zhang, “Eigenanalysis and continuum modelling of pre-twisted repetitive beam-like structures”, *Int. J. Solids Struct.* **43**:13 (2006), 3832–3855.
- [Renton 1984] J. D. Renton, “The beam-like behaviour of space trusses”, *AIAA J.* **22**:2 (1984), 273–280.
- [Santos e Lucato et al. 2004] S. L. dos Santos e Lucato, J. Wang, P. Maxwell, R. M. McMeeking, and A. G. Evans, “Design and demonstration of a high authority shape morphing structure”, *Int. J. Solids Struct.* **41**:13 (2004), 3521–3543.
- [Symons et al. 2005a] D. D. Symons, R. G. Hutchinson, and N. A. Fleck, “Actuation of the kagome double-layer grid, 1: Prediction of performance of the perfect structure”, *J. Mech. Phys. Solids* **53**:8 (2005), 1855–1874.
- [Symons et al. 2005b] D. D. Symons, R. G. Hutchinson, and N. A. Fleck, “Actuation of the kagome double-layer grid, 2: Effect of imperfections on the measured and predicted actuation stiffness”, *J. Mech. Phys. Solids* **53**:8 (2005), 1875–1891.
- [Timoshenko and Gere 1961] S. P. Timoshenko and J. M. Gere, *Theory of elastic stability*, 2nd ed., McGraw-Hill, New York, 1961.
- [Wicks and Guest 2004] N. Wicks and S. D. Guest, “Single member actuation in large repetitive truss structures”, *Int. J. Solids Struct.* **41**:3–4 (2004), 965–978.

Received 28 Apr 2006. Accepted 5 Dec 2006.

ANTHONY C. H. LEUNG: [anthony.leung@cantab.net](mailto:anthony.leung@cantab.net)

*Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom*

SIMON D. GUEST: [sdg@eng.cam.ac.uk](mailto:sdg@eng.cam.ac.uk)

*Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom*

<http://www.eng.cam.ac.uk/~sdg>

## AN ESHELBY-TYPE APPROACH FOR DEFECT ENERGETICS IN CARBON NANOTUBES

LIANG ZHANG AND JIA LU

This paper presents a method for characterizing the formation energy of Stone–Wales defect transformation in deforming carbon nanotubes. A formula is derived to show that the structural energy variation consists of the change of atomic potential due to bond reconfiguration in a local defective region, and an elastic correction that represents the influence of the remaining system. The advantage of the method lies in its computational efficiency. Specifically, global optimization for the configuration of the defective tube is eliminated. We use the method to investigate the formation energy distribution in nonuniformly deforming nanotubes, and to study the energetic interaction between multiple defects. Accuracy of the method is also assessed through numerical experiments.

### 1. Introduction

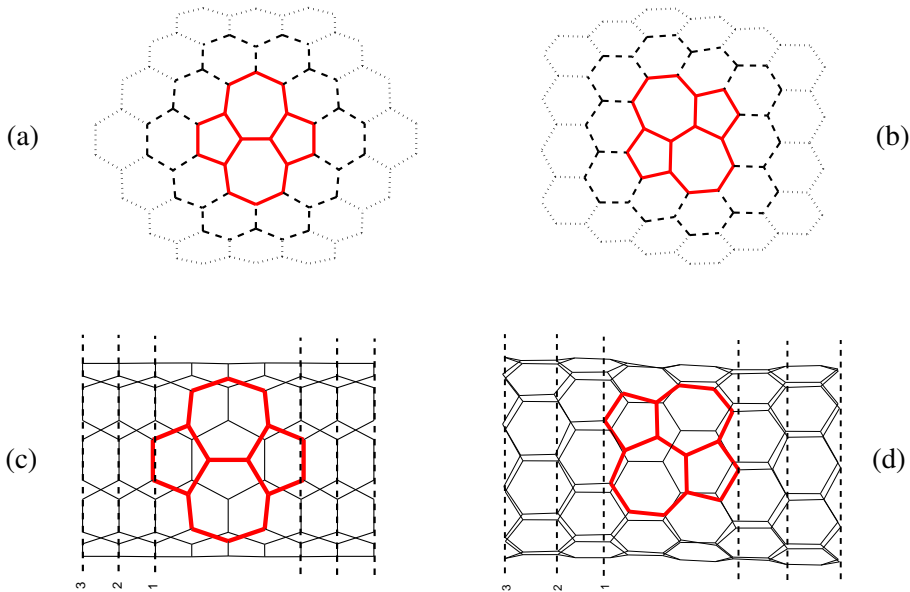
When loaded beyond a certain limit, carbon nanotubes (CNTs) respond to mechanical load via the formation of topological defects [Buongiorno Nardelli et al. 1998a; 1998b]. In a stretched CNT, a fundamental defect appears via a  $90^\circ$  rotation of a C-C bond from the predominantly transverse direction to the predominantly axial direction. This transition, known as the Stone–Wales (SW) transformation [1986], preserves the threefold  $sp^2$  bond structure but replaces the hexagon rings with two pentagon-heptagon (5-7-7-5) pairs as indicated in Figure 1. At large stretches the bond rotation releases the excessive strain [Buongiorno Nardelli et al. 1998a; 1998b] and results in an energetically favorable state. The SW defect is active at high temperatures, and further relaxation through successive rotations at neighboring bonds becomes energetically favorable [Samsonidze et al. 2002a]. In this case, the SW transition furnishes a mechanism of “intramolecular plasticity” [Yakobson 1998; Buongiorno Nardelli et al. 1998a; 1998b]. At low temperatures, SW defects are not as active but their presence weakens the local stiffness [Chandra et al. 2004] and presents a potential site for crack nucleation. Recent studies reported that a SW defect can significantly reduce the limit strain of a carbon nanotube [Zhang et al. 1998; Zhang and Crespi 2000; Zhao et al. 2002; Troya et al. 2003; Mielke et al. 2004], and also affect the electrical properties of CNTs [Choi et al. 2000; Liu et al. 2004b; Nordlund and Hakonen 2005].

The formation energy, namely the energy difference between the defective and the maiden states, is a significant factor in defect kinetics. For a transformation to be energetically favorable, the defective state should rest at a lower energy level. Hence, a negative formation energy is often taken as the necessary condition for defect formation. In order to trigger the transformation, however, the system must receive sufficient energy to overcome the activation energy barrier [Samsonidze et al. 2002b; Buongiorno Nardelli et al. 1998a; Zhao et al. 2002]. It has been reported that a SW defect in a tensile CNT

---

*Keywords:* Stone–Wales defect, carbon nanotube, Eshelby method.

This work was partially supported by a Scientific Research Initiative Grant from the University of Iowa.



**Figure 1.** Illustration of the defective patches: (a) and (b) are layer models for large diameter tubes; (c) and (d) are ring models for small diameter tubes.

becomes energetically favorable at about 5–6% strain. However, at lower temperatures the activation barrier for this formation remains high and thus hinders the defect creation [Zhao et al. 2002]. The formation energy nevertheless remains an important indicator as it reflects the tendency of defect formation and the thermodynamical stability of the defect. In [Samsonidze et al. 2002a], for example, the energy of multiple SW defects as a function of the relative position is used to evaluate possible patterns of defect propagation. Conceivably, the formation energy distribution in a nonhomogeneously deforming CNT may indicate the possible weak spot where a defect is most likely to occur if the environment permits its nucleation.

Many researchers have addressed the problem of computing the formation energy of the SW defects in carbon nanotubes, as well as in boron-nitride tubes. Some papers focused on quantifying the change of chemical bond energy due to defect reconfiguration [Zhang et al. 1998; Zhang and Crespi 2000; Pan et al. 2000; Zhou and Shi 2003; Bettinger et al. 2002; Dumitrica et al. 2003; Piquini et al. 2005]. Typically, for this purpose, small isolated patches of atoms were considered. Nevertheless, the formation energy includes the contribution from elastic distortion. Strictly speaking, the energy variation should be considered at the structural level. At present, various levels of theories from *ab initio* calculations to molecular dynamics (MD) simulations have been applied to structural energetics computation. The MD approach as in [Moon and Hwang 2004a; 2004b; Marques et al. 2005] can handle relatively large systems. It should be noted that different theories do lead to different predictions [Grossman et al. 1995; Bettinger et al. 2002].

In this paper, we present a method for computing the structural energetic variation of SW transformation in pre-loaded CNTs. The method rests on the observation that a 5-7-7-5 ring is a local defect,

in the sense that the elastic distortion is confined to a local region and diminishes rapidly after a short distance. If long range interactions such as the Coulomb force can be neglected, the energetic variation would result primarily from defective bond reconfiguration and the local elastic distortion. In Section 2, a formula for the relative Gibbs free energy is derived, in which the relative energy is expressed as a sum of the change of potential in the local defective region and an elastic correction that approximates the influences of the remaining system. The formula suggests an efficient way to compute the formation energy in relatively large scale atomic systems. Instead of computing the global energetics of the defective system, one only considers an embedded local defective cluster, and takes into account the remaining system through the correction term. The reduction allows one to readily perform structural level energetic analysis. For example, one can employ the method to compute the formation energy map, that is, the formation energy distribution in nonhomogeneously deforming systems. Likewise, one can study the energetic interaction between multiple defects by placing new defects in a pre-defective system. With the proposed method, global configuration optimization is performed only once. This local scheme works for classical potentials with finite range interaction. In Section 4, the method is applied to carbon nanotubes where the bond energy is modeled by the Brenner potential [Brenner 1990].

## 2. Structural formation energy of Stone–Wales defect

We consider CNTs at low temperatures where the thermal fluctuation is negligibly small and the atomic system remains approximately in a quasiequilibrium state. The formation of a defect causes a change of the total energy in the system. The difference

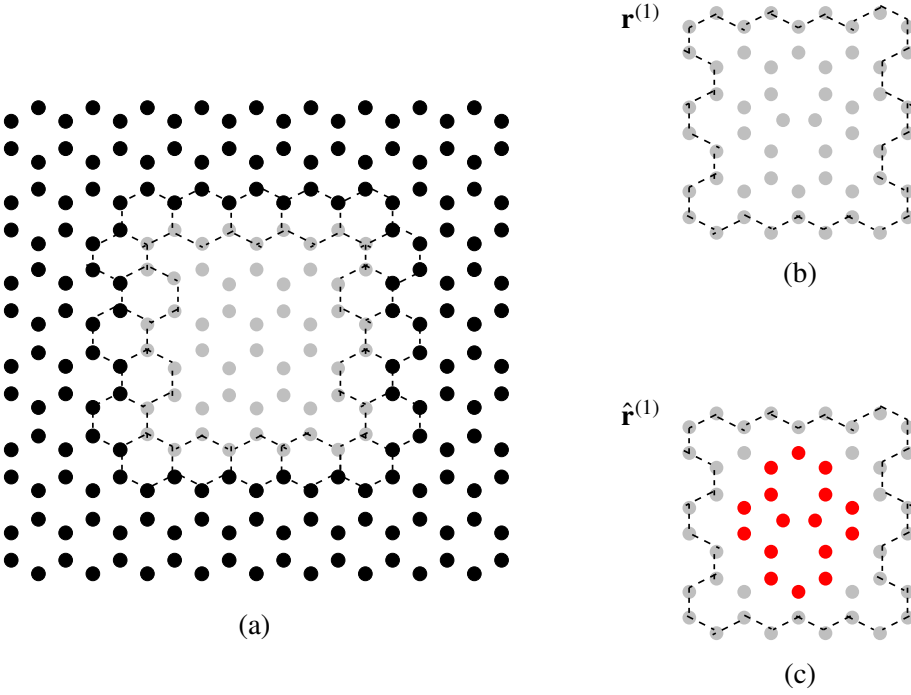
$$E_f = \Pi^{(2)} - \Pi^{(1)} \quad (1)$$

defines the formation energy, where  $\Pi^{(1)}$  and  $\Pi^{(2)}$  are the Gibbs free energies of the reference and the defective systems. As alluded to earlier, the energy difference should primarily result from the reconfiguration of  $sp^2$  bonds in the defect and its distortion to the surrounding lattices. Inspired by Eshelby's method for the inclusion problem [Eshelby 1957] and his approach to defect energetics [Eshelby 1975], we propose a local approach for computing the energy difference of the entire system.

Figure 2 illustrates the notions used in the localization method. The local cluster  $D$  is a small set of atoms in which a 5-7-7-5 dislocation is to be embedded. The set of atoms that is the complement of  $D$  in the pristine domain is denoted by  $P$ . The interface set, denoted by  $I$ , includes the atoms in  $D$  that have energetic interaction with any atom in  $P$  and vice versa. For pairwise interaction, the interface set contains the bonding pairs along the internal boundary. For many-body potentials the interface may extend to a thin layer of atoms. We employ the Brenner bond order potential [1990] to model the bond energy. The interface set consists of the atoms on the layer of hexagon rings directly connecting  $D$  and  $P$  (see Figure 2). In addition, we identify the atoms in  $P$  that are subject to external load as the boundary set  $B$ , which can be further divided into disjoint subsets  $B = B_r \cup B_t$  corresponding to displacement and force boundary points.

Consider a reference tube subject to the load

$$\begin{aligned} \mathbf{r}_i^{(1)} &= \bar{\mathbf{r}}_i, & \text{for atom } i \in B_r, \\ \mathbf{t}_i^{(1)} &= \bar{\mathbf{t}}_i, & \text{for atom } i \in B_t. \end{aligned}$$



**Figure 2.** Schematics of the method. (a) Gray dots represent the atoms of the local cluster  $D$  and black dots represent pristine atoms. The interface is marked by the layer of dashed hexagon rings; (b) The deformed local region with the atom positioned at  $\mathbf{r}^{(1)}$ ; (c) The defective replica  $\widehat{D}$ , whose boundary is fixed, but interior atoms are allowed to relax.

Let  $\mathbf{r}_j^{(1)}$  be the coordinates of atom  $j$  at equilibrium. Under the same external load, a defective tube with a SW dipole embedded in  $D$  produces a different equilibrium state  $\mathbf{r}^{(2)}$ . The Gibbs free energies of the reference and the defective tubes are

$$\begin{aligned}\Pi^{(1)} &= E_{P+D}(\mathbf{r}^{(1)}) - \sum_{b \in B_t} \bar{\mathbf{t}}_b \cdot \mathbf{r}_b^{(1)}, \\ \Pi^{(2)} &= E_{P+\widehat{D}}(\mathbf{r}^{(2)}) - \sum_{b \in B_t} \bar{\mathbf{t}}_b \cdot \mathbf{r}_b^{(2)}.\end{aligned}\tag{2}$$

The total interatomic potential consists of the sum of bond energies  $V_{ij}$  and can be regrouped into the sum of site potentials  $E_i$ . We write

$$E_{P+D} = \sum_{i < j} V_{ij} = \sum_{i \in P+D} E_i.\tag{3}$$

With Brenner potential and the threefold bond structure, the energy of site  $i$  is

$$E_i = \frac{1}{2} \sum_j V_{ij},\tag{4}$$

where the sum runs over the three  $sp^2$  bonds connecting to atom  $i$ . The potential of the whole tube can be naturally split into  $E_{P+D} = E_P + E_D$ , where each term consists of the sum of site energies in the respective region.

For the succeeding development, we introduce an auxiliary configuration  $\hat{\mathbf{r}}^{(1)}$  of the defective tube using an imaginary fixing and replacing process: (1) Select a region  $D$  into which a 5-7-7-5 ring is to be inserted; (2) Fix all the interface atoms; (3) Replace the interior atoms in  $D$  with a defective replica. The interior atoms in  $D$  do not directly interact with the interior atoms in  $P$  and hence, once the interface atoms are constrained, any reconfiguration in  $D$  will not induce atomic displacements in  $P$ . However, a set of force (called mismatch force) is needed on the interface atoms to keep the them from moving. This imaginary process leads to a configuration

$$\hat{\mathbf{r}}_i^{(1)} = \begin{cases} \mathbf{r}_i^{(1)} & \text{if } i \in P \cup I, \\ \hat{\mathbf{r}}_i^{(1)} & \text{if } i \in \hat{D} \setminus I. \end{cases}$$

Imagine that one isolates the local cluster  $D$  and keeps it in the deformed state  $\mathbf{r}^{(1)}$ . To sustain the deformation, one must apply a set of forces, denoted by  $\mathbf{t}^{(1)}$ , on the interface atoms. This set of forces corresponds the internal force that interface atoms experience in the reference state  $\mathbf{r}^{(1)}$ . Likewise, to deform the defective replica  $\hat{D}$  to the auxiliary configuration, one must apply another set of forces  $\hat{\mathbf{t}}^{(1)}$  on the interface atoms. By equilibrium, the difference  $\hat{\mathbf{t}}^{(1)} - \mathbf{t}^{(1)}$  is the interface mismatch force that is necessary to sustain the configuration  $\hat{\mathbf{r}}^{(1)}$  in addition to the original boundary forces that are left intact.

We now return to the computation of formation energy. The difference in free energy between the auxiliary and the reference structure is

$$E_{\hat{D}}(\hat{\mathbf{r}}^{(1)}) - E_D(\mathbf{r}^{(1)}). \quad (5)$$

If the mismatch force is gradually released, the defective tube will relax to the final defective configuration  $\mathbf{r}^{(2)}$ . If the tube is subject to all displacement boundary conditions, the work done by gradually releasing the mismatch force can be approximated as

$$-\frac{1}{2} \sum_{s,t \in I} (\hat{\mathbf{t}}_s^{(1)} - \mathbf{t}_s^{(1)}) \cdot \mathbb{C}_{st} (\hat{\mathbf{t}}_t^{(1)} - \mathbf{t}_t^{(1)}), \quad (6)$$

where  $\mathbb{C}_{st}$  is the compliance matrix associated with the interface atoms evaluated at the auxiliary configuration. Since the relaxation process is elastic, the work reflects the change of potential from the auxiliary to the final configuration. Adding (5) and (6), the change of free energy from the reference to the final defective state can be approximated as

$$E_f \approx E_{\hat{D}}(\hat{\mathbf{r}}^{(1)}) - E_D(\mathbf{r}^{(1)}) - \frac{1}{2} \sum_{s,t \in I} (\hat{\mathbf{t}}_s^{(1)} - \mathbf{t}_s^{(1)}) \cdot \mathbb{C}_{st} (\hat{\mathbf{t}}_t^{(1)} - \mathbf{t}_t^{(1)}). \quad (7)$$

In the Appendix, we show that the same formula applies to the force boundary condition as well. Thus, the formula furnishes a unified local scheme for the computation of the formation energy. The leading term  $E_{\hat{D}}(\hat{\mathbf{r}}^{(1)}) - E_D(\mathbf{r}^{(1)})$  gives a first order approximation to the structural energy variation. The remaining term represents structural influence. If the mismatch force is small, this term provides a second order correction to the formation energy.



Our experience indicates that the compliance matrix can be replaced by the corresponding compliance of the reference tube without causing a noticeable difference in the formation energy. This leads to further simplification in the computation, since the compliance matrix of the reference system can be readily obtained during configuration optimization.

### 3. Computational procedure

The molecular mechanics (MM) method with the implementation proposed in [Liu et al. 2004a] is employed to obtain the equilibrium state of the CNT under loads. The MM calculation is based on the energy minimization with respect to discrete atomic positions. The free energy in the system is

$$\Pi = E(\mathbf{r}) - \sum_{k \in B_f} \bar{\mathbf{t}}_k \cdot \mathbf{r}_k \quad (8)$$

The first generation Brenner potential [1990; 1992] is used to model CNT bond energy. The parameters listed as potential I in [Brenner 1990] are used in the computation.

The equilibrium of the atomic system requires that atomic positions cause the total potential energy to reach its minimum

$$\frac{\partial \Pi}{\partial \mathbf{r}_i} = 0, \quad (9)$$

subjected to  $\mathbf{r}_i = \bar{\mathbf{r}}$  on  $B_r$ . This yields a nonlinear system of  $3(N - N_{B_r})$  equations where  $N$  is the total number of atoms and  $N_{B_r}$  the number of atoms subject to boundary displacements. To find equilibrium atomic positions, the Newton–Raphson method is employed to solve the equations of (9). At the  $k$ -th iteration (configuration), the increments of the atomic coordinates are given by

$$\left[ \frac{\partial^2 \Pi}{\partial \mathbf{d} \partial \mathbf{d}} \right]_{\mathbf{d}^{(k)}} \Delta \mathbf{d}^{(k+1)} = - \frac{\partial \Pi}{\partial \mathbf{d}} \Big|_{\mathbf{d}^{(k)}}, \quad (10)$$

where

$$\mathbb{K} := \frac{\partial^2 \Pi}{\partial \mathbf{d} \partial \mathbf{d}}$$

is the tangent stiffness matrix (the Hessian), and  $\mathbf{d}$  is the vector of atomic displacements. The system (10) is solved iteratively until convergence is achieved.

At convergence we partition the stiffness matrix  $\mathbb{K}$  into

$$\mathbb{K} = \begin{bmatrix} \mathbb{K}_{ss} & \mathbb{K}_{sd} \\ \mathbb{K}_{ds} & \mathbb{K}_{dd} \end{bmatrix}, \quad (11)$$

where  $s$  stands for the degrees of freedom of the interface atoms  $I$ , and  $d$  denotes the remaining degrees of freedom. The compliance matrix  $\mathbb{C}$  in (7) is computed according to

$$\mathbb{C} = (\mathbb{K}_{ss} - \mathbb{K}_{sd} \mathbb{K}_{dd}^{-1} \mathbb{K}_{ds})^{-1}.$$

The procedure for computing the energy difference is summarized as follows:

1. Compute the equilibrium configuration of the reference tube.
2. At the position where a SW ring is to be inserted, select a local cluster, identify the interface, and record the coordinates of the interface atoms.

3. Replace the cluster with a defective replica, solve a local equilibrium problem to compute the atomic coordinates in the replica keeping interface atoms fixed from step 2.
4. Compute the relative energy according to Equation (7).

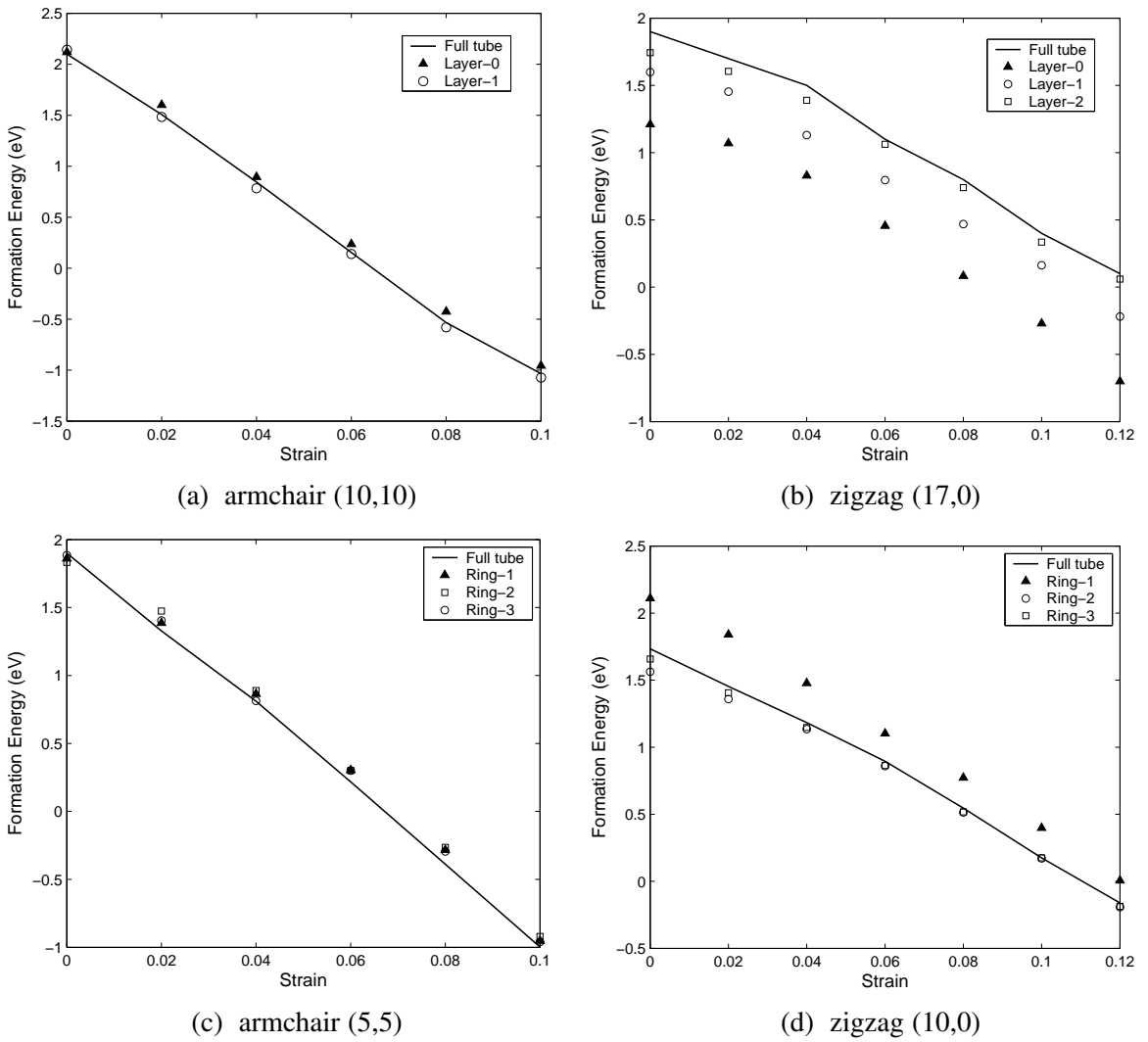
Global configuration optimization is performed only once, on the reference tube. This reduction allows one to efficiently compute the defect energetics in a relatively large system, and also to compute the relative energy as the defect position varies. If the structure has unconstrained rigid body modes (typical in MD simulations), the tangent stiffness  $\mathbb{K}$  is singular and hence the compliance cannot be directly computed. In this case, one can eliminate the rigid body singularity by augmenting  $\mathbb{K}$  with several rank-one updates  $\sum_{\alpha} a_{\alpha} \mathbf{R}_{\alpha} \otimes \mathbf{R}_{\alpha}$ , where  $\alpha$  ( $\alpha \leq 6$ ) is the number of rigid body modes,  $\mathbf{R}_{\alpha}$  are the modes vectors,  $\mathbb{K} \mathbf{R}_{\alpha} = \mathbf{0}$ , and  $a_{\alpha}$  are positive scalars. This will not affect the energy computation since the matching force  $\hat{\mathbf{t}}^{(1)} - \mathbf{t}^{(1)}$  is self-equilibrium. The proof of this fact is straightforward.

#### 4. Results and discussion

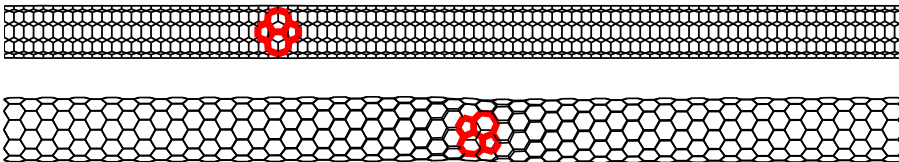
**Cluster size.** The proposed method is based on the premise that if the size of the local cluster is sufficiently large, the mismatch force should be small, so Equation (7) should provide a second-order accurate estimation for the formation energy. For practical purposes, it is of foremost importance to investigate the influence of cluster size on the energy computation. We used two types of clusters. For large diameter tubes we introduced the layer patch as shown in Figures 1(a) and (b). The pentagon-heptagon ring itself (bold solid line), which has 16 atoms, is defined as the layer-0 patch. The region enclosed by dashed lines, with 42 atoms, is termed layer-1; the layer-2 region is enveloped by the dotted line. These types of clusters are consistent with the defect-shell model of [Zhou and Shi 2003]. For small diameter tubes, the ring model was introduced. A defective ring is a tube segment that contains a SW defect. Figure 1(c) shows the family of hierarchically defined defective rings in a (5,5) armchair tube. The ring-1 corresponds to the segment that just contains a 5-7-7-5 dipole, the ring-2 extends to two additional layers of atoms, and so on. The corresponding ring models in a (10,0) zigzag appear in Figure 1(d). Note that a 5-7-7-5 dipole has three possible orientations in each chirality [Zhou and Shi 2003]. Here we consider the one in which the pentagon-pentagon axis points predominantly along the tube axis.

We considered tubes under axial tension by prescribed end-displacement. A SW defect is inserted in a 5 nm long (10,10) armchair and a 5 nm long (17,0) zigzag. Figures 3(a) and 3(b) show the values of formation energy obtained using clusters of various sizes. It is seen that, for the (10,10) armchair, the local cluster as small as layer-0 can yield results almost identical to those by the full tube. The high accuracy may be attributed to the relatively minor distortion the SW defect produces in the armchair configuration. In contrast, Figure 3(b) shows that the layer-2 cluster is necessary to achieve a comparable accuracy in the (17,0) zigzag owing to the relatively large distortion the SW defect produces in a zigzag tube.

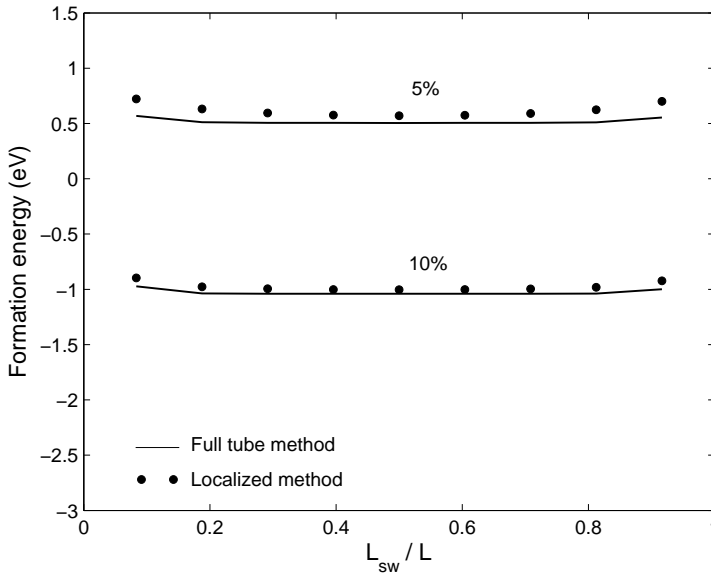
The analyses were repeated for the 12 nm long (5,5) armchair and the (10,0) zigzag shown in Figure 4 using the defective ring. Figures 3(a) and 3(b) present the formation energy results. For the armchair, the ring-1 and the ring-2 results are almost identical and both are very close to the full tube result. On the contrary, for zigzag tube the ring-1 results show a moderately large deviation from the full tube result. The ring-2 cluster seems necessary for an accurate estimation. The formation energies of the SW defect are found to approach zero when the strain is around 6% for armchair tubes and around 12% in zigzag tubes. These values agree nicely with existing reports [Buongiorno Nardelli et al. 1998a].



**Figure 3.** Comparative formation energies versus stretch.



**Figure 4.** Slender tubes (12nm) used in the computation. Top: (5,5) armchair. Bottom: (10,0) zigzag.

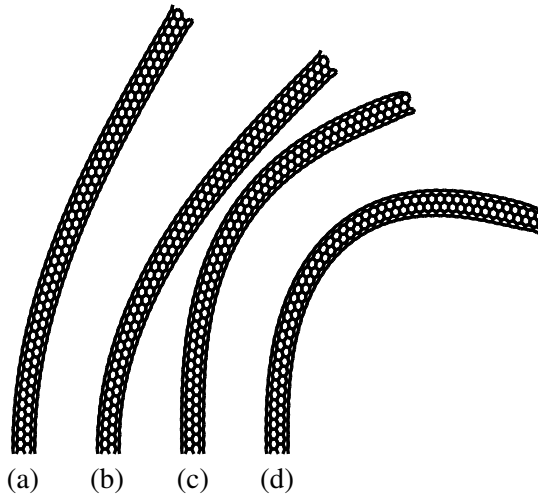


**Figure 5.** Distribution of formation energy in the uniformly stretched nanotube.

**Formation energy map.** In a nonuniformly deforming CNT, the defect formation energy depends on the local strain (stress) level and factors such as the distance to the tube boundary. Hence, the spatial distribution of formation energy provides an indicator for the potential weak spot where a defect transformation is most likely to occur. We applied the method to study the formation energy distribution. We first considered a 12 nm long (5,5) armchair under axial tension. Nine locations uniformly placed along the tube length were chosen where a SW dipole was to be inserted. The formation energies were calculated for these nine locations respectively, assuming a single SW defect appears in the tube for each time. The ring-1 model was employed; see Figure 1(c). Figure 5 presents the formation energy distribution under three levels of strain.

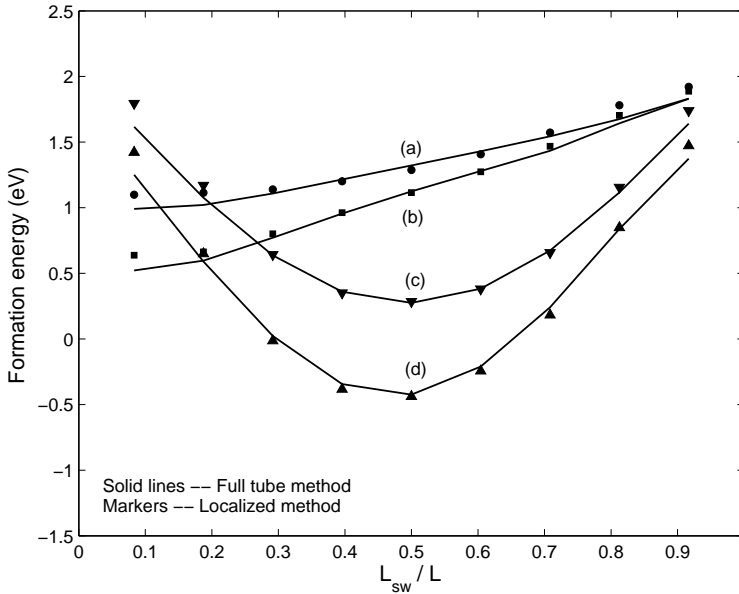
Notably, the formation energy possesses a symmetric distribution and attains its lowest value in the middle, implying that the defect transformation is most likely to initiate near the middle. Although the reference state is uniform, the formation energy shows a small variation across the tube length due to boundary effects. For a further assessment of accuracy, the full tube results are also computed. The approximate results agree well with the full tube computations in most locations, except near the ends of the tube, where the solutions deviate slightly.

Bending is a common mode of deformation that slender CNTs are likely to experience. We calculated the formation energies of the 12nm (5,5) armchair in four bent configurations shown in Figure 6. Configurations (a) and (b) were obtained by applying a transverse force at the tip; and configurations (c) and (d) were obtained by prescribing end-displacement. Since we were interested in stretch-induced SW rotation, the defect was inserted in the stretched side of the tube. The same nine locations were employed. In configurations (a) and (b), the maximum bending moment occurs at the fixed end while in (c) and (d) the maximum bending occurs in the middle of the tube. Figure 7 shows the formation energy distribution. For (a) and (b), where maximum bending appears at the end, the lowest formation

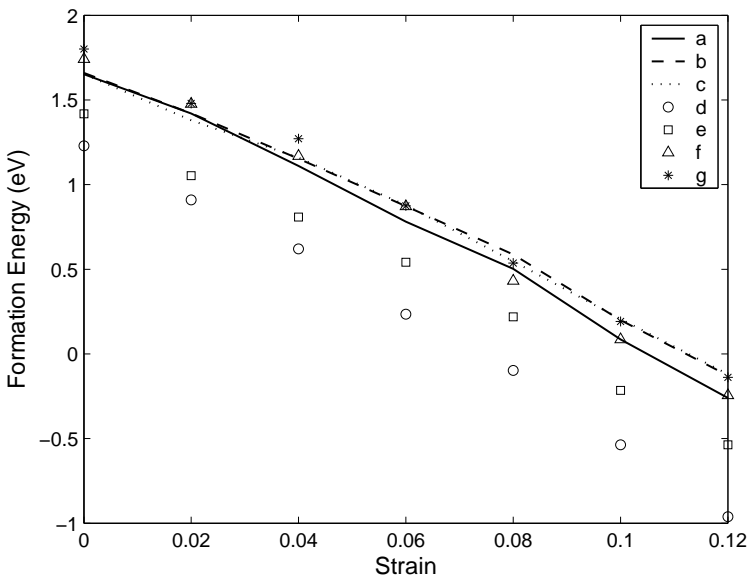


**Figure 6.** Four bent configurations of the (5,5) carbon nanotube.

energy is found near the end. In contrast, for (c) and (d), formation energy is lowest in the middle. The distribution indicates that the formation energy correlates directly with the bending moment, which is expected since the latter determines the local stretch in the tube. Figure 7 also includes the results from full tube computation. The approximate results are found to be close to the full tube solution in most locations. Only near the boundary does the solution show minor deviations.

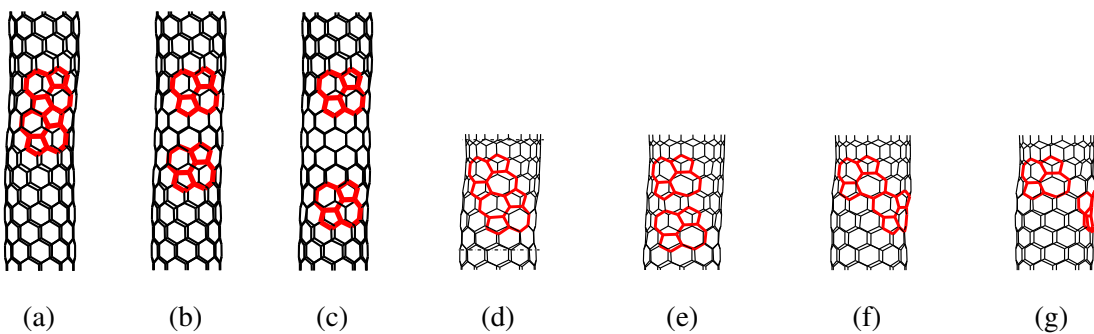


**Figure 7.** The distribution of formation energy on bent (5,5) nanotubes. Curves (a) through (d) correspond to the configurations in Figure 6.



**Figure 8.** Interaction energy between two defects. The cases (a) through (f) correspond to the relative positions illustrated in Figure 9.

**Defect interaction.** Samsonidze et al. [2002a] reported that when two SW defects are placed close to each other, the formation energy varies with their relative position and orientation, indicating that two defects may repulse or attract each other depending on their relative position. We studied the interaction of SW defects in stretched tubes using the proposed method. Figure 9 shows a defective (10,0) tube having a primary SW defect and a secondary defect inserted at various positions. We considered the following cases: axially adjacent (a); axially distanced (b) and (c); diagonally adjacent (d); axially skewed (e); and diagonally distanced (f) and (g). For locations (b) and (c), we employed the ring-1 model for each defect. A larger ring enclosing two SW defects (see part (e) of Figure 9) was used for other cases. The interaction is characterized by the difference of the Gibbs free energies with and without the second defect.



**Figure 9.** Relatively positioned multiple defects.

The energy curves are presented in Figure 8. The diagonally adjacent case (d) shows the lowest interaction energy across the strain range, indicating that such a configuration is most favorable. The axially skewed location (e) has the second lowest energy across the strain range. Although the defects are closely packed in (a), the energy is only slightly lower than other cases where the defects are further separated, suggesting that axially adjacent defects are less likely to appear in a zigzag tube. It is seen that if the defects are sufficiently distanced from each other, the interaction energies appear to be insensitive to the relative position. While the pattern of orientation dependence appears to agree with the early report in [Samsonidze et al. 2002a], the lack of interaction for distanced defects may be a consequence of the Brenner potential which vanishes after a short cut-off distance.

## 5. Summary

We have proposed a method to compute the structural energetic variation induced by the formation of a SW defect. The method exploits the local nature of a SW defect and yet properly takes into account the energetic influence from the remaining system. The advantage lies in the computational efficiency in comparison to other methods routinely used in computation of formation energy. Within the framework of MD (or MM) computation, the method provides an efficient way to study SW energetics at the structural level. The method applies to systems where the interaction is of finite range. In such systems the influence of bond reconfiguration can be localized. If long range interactions, such as the Coulomb force, are significant, the nonlocal effect needs to be carefully considered.

### Appendix: Derivation of Equation (7)

Starting from Equations (1) and (2), adding and subtracting the boundary term  $\sum_{s \in I} \mathbf{t}_s^{(1)} \cdot (\mathbf{r}_s^{(2)} - \mathbf{r}_s^{(1)})$  gives

$$E_f \equiv \Pi^{(2)} - \Pi^{(1)} = \left( E_P(\mathbf{r}^{(2)}) - E_P(\mathbf{r}^{(1)}) - \sum_{b \in B_I} \bar{\mathbf{t}}_b \cdot (\mathbf{r}_b^{(2)} - \mathbf{r}_b^{(1)}) - \sum_{s \in I} (-\mathbf{t}_s^{(1)}) \cdot (\mathbf{r}_s^{(2)} - \mathbf{r}_s^{(1)}) \right) + \left( E_{\hat{D}}(\mathbf{r}^{(2)}) - E_D(\mathbf{r}^{(1)}) - \sum_{s \in I} \mathbf{t}_s^{(1)} \cdot (\mathbf{r}_s^{(2)} - \mathbf{r}_s^{(1)}) \right), \quad (12)$$

where  $\mathbf{t}_s^{(1)}$  is the internal force exerted on the interface atoms to maintain  $D$  in the reference state  $\mathbf{r}^{(1)}$ . Adding and subtracting terms further, the energy difference can be written as

$$E_f = \left( E_P(\mathbf{r}^{(2)}) - E_P(\mathbf{r}^{(1)}) - \sum_{b \in B_I} \bar{\mathbf{t}}_b \cdot (\mathbf{r}_b^{(2)} - \mathbf{r}_b^{(1)}) - \sum_{s \in I} (-\mathbf{t}_s^{(1)}) \cdot (\mathbf{r}_s^{(2)} - \mathbf{r}_s^{(1)}) \right) + \left( E_{\hat{D}}(\mathbf{r}^{(2)}) - E_{\hat{D}}(\hat{\mathbf{r}}^{(1)}) - \sum_{s \in I} \hat{\mathbf{t}}_s^{(1)} \cdot (\mathbf{r}_s^{(2)} - \mathbf{r}_s^{(1)}) \right) + E_{\hat{D}}(\hat{\mathbf{r}}^{(1)}) - E_D(\mathbf{r}^{(1)}) + \sum_{s \in I} (\hat{\mathbf{t}}_s^{(1)} - \mathbf{t}_s^{(1)}) \cdot (\mathbf{r}_s^{(2)} - \mathbf{r}_s^{(1)}). \quad (13)$$

In what follows the two terms in large parentheses will be simplified. The first term corresponds to the energy variation in region  $P$  under the variation of atomic positions  $\mathbf{r}^{(1)} \rightarrow \mathbf{r}^{(2)}$ . By Taylor expansion,

$$E_P(\mathbf{r}^{(2)}) - E_P(\mathbf{r}^{(1)}) = \sum_{i \in P} \left. \frac{\partial E}{\partial \mathbf{r}_i} \right|_{\mathbf{r}^{(1)}} \cdot \Delta \mathbf{r}_i + \frac{1}{2} \sum_{i, j \in P} \Delta \mathbf{r}_i \cdot \left[ \left. \frac{\partial^2 E}{\partial \mathbf{r}_i \partial \mathbf{r}_j} \right|_{\mathbf{r}^{(1)}} \Delta \mathbf{r}_j \right] + O(|\Delta \mathbf{r}|^3), \quad (14)$$

where  $\Delta \mathbf{r}_i = \mathbf{r}_i^{(2)} - \mathbf{r}_i^{(1)}$ . Invoking the virtual work principle, we have

$$\sum_{i \in P} \left. \frac{\partial E}{\partial \mathbf{r}_i} \right|_{\mathbf{r}^{(1)}} \cdot \Delta \mathbf{r}_i = \sum_{b \in B_i} \bar{\mathbf{t}}_b \cdot \Delta \mathbf{r}_b + \sum_{s \in I} (-\mathbf{t}_s^{(1)}) \cdot \Delta \mathbf{r}_s \quad (15)$$

Substituting (15) and (14) into the first bracket in (13), we find that the first order terms vanish, leaving

$$\frac{1}{2} \sum_{i, j \in P} \Delta \mathbf{r}_i \cdot \left[ \left. \frac{\partial^2 E}{\partial \mathbf{r}_i \partial \mathbf{r}_j} \right|_{\mathbf{r}^{(1)}} \Delta \mathbf{r}_j \right] + O(|\Delta \mathbf{r}|^3). \quad (16)$$

The domain sum term can be converted into boundary sum. With repeated use of the virtual work principle we can deduce

$$\sum_{i \in P} \Delta \mathbf{r}_i \cdot \left[ \left. \frac{\partial^2 E}{\partial \mathbf{r}_i \partial \mathbf{r}_j} \right|_{\mathbf{r}^{(1)}} \Delta \mathbf{r}_j \right] = - \sum_{s \in I} (\mathbf{t}_s^{(2)} - \mathbf{t}_s^{(1)}) \cdot (\mathbf{r}_s^{(2)} - \mathbf{r}_s^{(1)}) + O(|\Delta \mathbf{r}|^3). \quad (17)$$

Combining (16) and (17), the expression inside the first bracket in (13) reduces to

$$-\frac{1}{2} \sum_{s \in I} (\mathbf{t}_s^{(2)} - \mathbf{t}_s^{(1)}) \cdot (\mathbf{r}_s^{(2)} - \mathbf{r}_s^{(1)}) + O(|\Delta \mathbf{r}|^3). \quad (18)$$

Applying the same procedure to the second bracket yields

$$E_{\hat{D}}(\mathbf{r}^{(2)}) - E_{\hat{D}}(\hat{\mathbf{r}}^{(1)}) - \sum_{s \in I} \hat{\mathbf{t}}_s^{(1)} \cdot (\mathbf{r}_s^{(2)} - \hat{\mathbf{r}}^{(1)}) = \frac{1}{2} \sum_{s \in I} (\mathbf{t}_s^{(2)} - \hat{\mathbf{t}}_s^{(1)}) \cdot (\mathbf{r}_s^{(2)} - \mathbf{r}_s^{(1)}) + O(|\Delta \mathbf{r}|^3). \quad (19)$$

Substituting (18) and (19) into (13), we conclude

$$E_f = E_{\hat{D}}(\hat{\mathbf{r}}^{(1)}) - E_D(\mathbf{r}^{(1)}) + \frac{1}{2} \sum_{s \in I} (\hat{\mathbf{t}}_s^{(1)} - \mathbf{t}_s^{(1)}) \cdot (\mathbf{r}_s^{(2)} - \mathbf{r}_s^{(1)}) + O(|\Delta \mathbf{r}|^3), \quad (20)$$

where  $\hat{\mathbf{t}}_s^{(1)} - \mathbf{t}_s^{(1)}$  is the mismatch force. The final position  $\mathbf{r}^{(2)}$  can be obtained by relaxing the mismatch force, or equivalently by applying a set of equal and opposite forces on the interface. To the first order, the interface atoms incremental displacements induced by the relaxation can be computed by

$$\mathbf{r}_s^{(2)} - \mathbf{r}_s^{(1)} = - \sum_{s, t \in I} \mathbb{C}_{st} (\hat{\mathbf{t}}_t^{(1)} - \mathbf{t}_t^{(1)}), \quad (21)$$

where  $\mathbb{C}$  is the current compliance associated with the interface atoms. In light of this, we can further write

$$E_f = E_{\hat{D}}(\hat{\mathbf{r}}^{(1)}) - E_D(\mathbf{r}^{(1)}) - \frac{1}{2} \sum_{s, t \in I} (\hat{\mathbf{t}}_s^{(1)} - \mathbf{t}_s^{(1)}) \cdot \mathbb{C}_{st} (\hat{\mathbf{t}}_t^{(1)} - \mathbf{t}_t^{(1)}) + O(|\Delta \mathbf{r}|^3). \quad (22)$$



It is evident that Equation (7) is second-order accurate with respect to the magnitude of the mismatch force. If the potential energy is quadratic, the formula becomes exact.

## References

- [Bettinger et al. 2002] H. F. Bettinger, T. Dumitrică, G. E. Scuseria, and B. I. Yakobson, “Mechanically induced defects and strength of BN nanotubes”, *Phys. Rev. B* **65**:4 (2002), 041406.
- [Brenner 1990] D. W. Brenner, “Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films”, *Phys. Rev. B* **42**:15 (1990), 9458–9471.
- [Brenner 1992] D. W. Brenner, “Erratum: Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films”, *Phys. Rev. B* **46**:3 (1992), 1948–1948.
- [Buongiorno Nardelli et al. 1998a] M. Buongiorno Nardelli, B. I. Yakobson, and J. Bernholc, “Mechanism of strain release in carbon nanotubes”, *Phys. Rev. B* **57**:8 (1998), R4277–R4280.
- [Buongiorno Nardelli et al. 1998b] M. Buongiorno Nardelli, B. I. Yakobson, and J. Bernholc, “Brittle and ductile behavior in carbon nanotubes”, *Phys. Rev. Lett.* **81**:21 (1998), 4656–4659.
- [Chandra et al. 2004] N. Chandra, S. Namilaie, and C. Shet, “Local elastic properties of carbon nanotubes in the presence of Stone–Wales defects”, *Phys. Rev. B* **69**:9 (2004), 094101.
- [Choi et al. 2000] H. J. Choi, J. Ihm, S. G. Louie, and M. L. Cohen, “Defects, quasibound states, and quantum conductance in metallic carbon nanotubes”, *Phys. Rev. Lett.* **84**:13 (2000), 2917–2920.
- [Dumitrica et al. 2003] T. Dumitrica, H. F. Bettinger, G. E. Scuseria, and B. I. Yakobson, “Thermodynamics of yield in boron nitride nanotubes”, *Phys. Rev. B* **68**:8 (2003), 085412.
- [Eshelby 1957] J. D. Eshelby, “The determination of the elastic field of an ellipsoidal inclusion, and related problems”, *Proc. R. Soc. Lond. A* **241**:1226 (1957), 376–396.
- [Eshelby 1975] J. D. Eshelby, “The elastic energy-momentum tensor”, *J. Elasticity* **5**:3–4 (1975), 321–335.
- [Grossman et al. 1995] J. D. Grossman, L. Mitas, and K. Raghavachari, “Structure and stability of molecular carbon: importance of electron correlation”, *Phys. Rev. Lett.* **75**:21 (1995), 3870–3873.
- [Liu et al. 2004a] B. Liu, Y. Huang, H. Jiang, S. Qu, and K. C. Hwang, “The atomic-scale finite element method”, *Comput. Methods Appl. Mech. Eng.* **193**:17–20 (2004), 1849–1864.
- [Liu et al. 2004b] B. Liu, H. Jiang, H. T. Johnson, and Y. Huang, “The influence of mechanical deformation on the electrical properties of single wall-carbon nanotubes”, *J. Mech. Phys. Solids* **52**:1 (2004), 1–26.
- [Marques et al. 2005] L. A. Marques, L. Pelaz, M. Aboy, P. Lopez, and J. Barbolla, “A novel technique for the structural and energetic characterization of lattice defects in the molecular dynamics framework”, *Comput. Mater. Sci.* **33**:1–3 (2005), 112–117.
- [Mielke et al. 2004] S. L. Mielke, D. Troya, S. L. Zhang, J. L. Li, S. Xiao, R. Car, R. S. Ruoff, G. C. Schatz, and T. Belytschko, “The role of vacancy defects and holes in the fracture of carbon nanotubes”, *Chem. Phys. Lett.* **390**:4–6 (2004), 413–420.
- [Moon and Hwang 2004a] W. H. Moon and H. J. Hwang, “A force field approach of structure and formation energy of defects of boron nitride nanotubes with tetragon-octagon pairs”, *Mater. Lett.* **58**:17–18 (2004), 2331–2334.
- [Moon and Hwang 2004b] W. H. Moon and H. J. Hwang, “Molecular-dynamics simulation of defect formation energy in boron nitride nanotubes”, *Phys. Lett. A* **320**:5 (2004), 446–451.
- [Nordlund and Hakonen 2005] K. Nordlund and P. Hakonen, “Nanotube: controlling conductance”, *Nat. Mater.* **4**:7 (2005), 514–515.
- [Pan et al. 2000] B. C. Pan, W. S. Yang, and J. Yang, “Formation energies of topological defects in carbon nanotubes”, *Phys. Rev. B* **62**:19 (2000), 12652–12655.
- [Piquini et al. 2005] P. Piquini, R. J. Baierle, T. M. Schmidt, and A. Fazzio, “Formation energy of native defects in BN nanotubes: an ab initio study”, *Nanotechnology* **16**:6 (2005), 827–831.
- [Samsonidze et al. 2002a] G. G. Samsonidze, G. G. Samsonidze, and B. I. Yakobson, “Energetics of Stone–Wales defects in deformations of monoatomic hexagonal layers”, *Comput. Mater. Sci.* **23**:1–4 (2002), 62–72.

- [Samsonidze et al. 2002b] G. G. Samsonidze, G. G. Samsonidze, and B. I. Yakobson, “Kinetic theory of symmetry-dependent strength in carbon nanotubes”, *Phys. Rev. Lett.* **88**:6 (2002), 065501.
- [Stone and Wales 1986] A. J. Stone and D. J. Wales, “Theoretical studies of icosahedral  $C_{60}$  and some related species”, *Chem. Phys. Lett.* **128**:5–6 (1986), 501–503.
- [Troya et al. 2003] D. Troya, S. L. Mielke, and G. C. Schatz, “Carbon nanotube fracture: differences between quantum mechanical mechanisms and those of empirical potentials”, *Chem. Phys. Lett.* **382**:1–2 (2003), 133–141.
- [Yakobson 1998] B. I. Yakobson, “Mechanical relaxation and intramolecular plasticity in carbon nanotubes”, *Appl. Phys. Lett.* **72**:8 (1998), 918–920.
- [Zhang and Crespi 2000] P. Zhang and V. H. Crespi, “Plastic deformations of boron-nitride nanotubes: an unexpected weakness”, *Phys. Rev. B* **62**:16 (2000), 11050–11053.
- [Zhang et al. 1998] P. Zhang, P. E. Lammert, and V. H. Crespi, “Plastic deformations of carbon nanotubes”, *Phys. Rev. Lett.* **81**:24 (1998), 5346–5349.
- [Zhao et al. 2002] Q. Zhao, M. Buongiorno Nardelli, and J. Bernholc, “Ultimate strength of carbon nanotubes: a theoretical study”, *Phys. Rev. B* **65**:14 (2002), 144105.
- [Zhou and Shi 2003] L. G. Zhou and S.-Q. Shi, “Formation energy of Stone–Wales defects in carbon nanotubes”, *Appl. Phys. Lett.* **83**:6 (2003), 1222–1224.

Received 27 Feb 2006. Accepted 27 Jun 2006.

LIANG ZHANG: [lianzhan@engineering.uiowa.edu](mailto:lianzhan@engineering.uiowa.edu)

*Department of Mechanical and Industrial Engineering, Center for Computer Aided Design, The University of Iowa, Iowa City, IA 52242-1527, United States*

JIA LU: [jia-lu@uiowa.edu](mailto:jia-lu@uiowa.edu)

*Department of Mechanical and Industrial Engineering, Center for Computer Aided Design, The University of Iowa, Iowa City, IA 52242-1527, United States*

## FRACTURE AND FATIGUE CRACK GROWTH ANALYSIS OF RAIL STEELS

HESHMAT A. AGLAN AND MAHMOOD FATEH

Low carbon bainitic steel shows promising potential, especially in critical components such as frogs and switches. Microstructural analysis of J6 bainitic rail steel was performed and compared with the microstructure of premium pearlitic rail steel. The bainitic microstructure revealed a mixture of tempered martensite and ferrite associated with intralath carbides. Typical pearlitic microstructure with a fine lamellar aggregate of very soft and ductile ferrite and very hard carbide cementite was observed. The mechanical properties, plane stress fracture toughness,  $K_{Ic}$ , and the fatigue crack growth behavior of the two steels were evaluated. Test specimens were machined from railheads of each material using electrical discharge machining (EDM). Rectangular unnotched and notched specimens were used for the mechanical properties and fatigue evaluation respectively.  $1/2T$  compact tension specimens were used for the  $K_{Ic}$  evaluation according to ASTM E399. The J6 bainitic steel has ultimate strength, yield strength, and elongation to failure of about 1500 MPa, 1100 MPa, and 13% respectively. These values are higher than those for pearlitic steel. It was found that the average  $K_{Ic}$  for the bainitic rail steel is 52 MPa  $\sqrt{m}$ , while that of the premium pearlitic steel is 41 MPa  $\sqrt{m}$ . Fatigue studies showed that the crack speed for the bainitic steel is lower than that for the pearlitic steel over the entire range of the energy release rate. The bainitic steel exhibits a higher rate of crack deceleration in the second stage, as indicated by the lower slope of the fatigue crack propagation kinetics curve in comparison with the pearlitic steel. This attests to the superior fatigue damage tolerance of the bainitic rail steel and provides evidence to support the superior rolling fatigue damage tolerance of the bainitic rail steel reported in the literature.

### 1. Introduction

Pearlitic steel accounts for most of the steel tonnage produced today [Smith 2004, 427–522]. Bainitic steel can offer some advantages compared to standard pearlitic steel in special track applications, for example, high angle crossing diamonds [Davis et al. 2002]. Recent studies reported in [Kristan 2005] have shown that J6 bainitic rail steel in a revenue service trial has shown performance superior to standard head-hardened pearlitic steel. A 50% reduction in rolling contact fatigue (RCF) damage was noted for the bainitic steel compared to head-hardened pearlitic steel. Kristen also stated that in the J6 bainitic low rail of the curve, cracks and shelling were almost completely absent, while for the pearlitic steel, widespread RCF damage, extensive cracks and shelling were observed. Although bainitic steels have shown superior performance, their application as premium rails has been hindered due to manufacturing and welding difficulties. Nevertheless, fundamental research into bainitic steels in comparison with premium

---

*Keywords:* bainitic steel, fracture toughness, fatigue, pearlitic steel.

This work was sponsored by the Federal Railroad Administration, Department of Transportation, under grant No. DTFR53-02-G-00021. The fracture toughness study was sponsored by the Y-12 security complex, through the PDRD program under subcontract No. 430003430.

pearlitic steels can elucidate the microstructural origin of mechanical properties, fracture toughness, and fatigue damage tolerance of these materials. This will lead to the development of rail steels with superior performance.

Pearlitic steels obtain their strength from fine grains of pearlite. However, there is a limit to the production of very fine grains in manufacturing and post-heat treatment processes. In contrast, bainitic steels derive their strength from ultra-fine structure with many dislocations that are harmless and confer high strength [Sawley 1997]. The microstructure of bainitic steel is a metastable aggregate of ferrite and cementite produced from the transformation of austenite at temperatures below the pearlite range and above the martensite starting temperature. Unlike the microstructure of pearlitic steel, the ferrite in bainitic steel has an acicular morphology and the carbides are discrete particles. The microstructure of bainitic steel is more complex than that of pearlitic steel and is largely dependent on the compositions and processing conditions. Bainite can be formed from austenite by isothermal transformation or by continuous cooling transformation (CCT). Usually, bainite in isothermally transformed steel is well characterized and distinguished as *upper bainite* and *lower bainite* depending on whether the carbides are distributed between individual ferrite regions or within them, respectively. The microstructure of upper bainite contains parallel lath-shape units of ferrite, which produce the so-called "feathery" appearance in optical microscopy and are formed at temperatures above 350° C. Lower bainite, on the other hand, has an acicular appearance similar to tempered martensite, and is often produced at temperatures below 350° C. These microstructural differences between upper and lower bainitic steel have led to differences in their mechanical properties. Usually, lower bainitic steel has higher strength and toughness than upper bainitic steel [Davis 1998].

A limited amount of research has so far been conducted on bainitic steels. Su and Clayton [1996] gave a review of experimental research on wheel-rail contact, pointing out that the study of bainitic rail steel remained a fertile area for research. They found that rolling contact fatigue cracks are associated with plastic deformation and embryo cracks even under dry running conditions. Jeong et al. [1998] have estimated rail wear limits based on rail strength investigations. Surface modification has been proposed to mitigate the wear and prevent the formation of surface cracks on the rails [Dimelfi et al. 1998; Yun et al. 1996]. Yokoyama et al. [2002] studied the effect of angle of attack ( $1^\circ - 5^\circ$ ) on the RCF damage resistance of pearlitic and bainitic rail steels. They found that bainitic rail steel displayed better RCF damage resistance than pearlitic steel for all angles of attack tested. The RCF performance of low strength bainitic steel was higher than the head-hardened pearlitic rail steel. Very recently Li et al. [2006] studied the RCF of a medium carbon bainitic steel (0.43% carbon). Three stages of RCF short-crack kinetics were observed, similar to those for conventional long-crack fatigue. Vertical short cracks parallel to the axis of the roller were observed in the initial stage, followed by a stage of un-propagating behavior, which consumed a larger number of RCF cycles. In the third stage the vertical short cracks re-propagated with a high acceleration in the direction parallel to the contact surface. It was suggested that the number of RCF cycles at which the re-propagation of the vertical short cracks occur could be used as the crack initiation life of the RCF. At this point, typically, the RCF cracks are very small and the initiation life may not be a dominating factor and hence the crack growth life ought to be considered. Kang et al. [2006] studied the ratcheting behavior of the same bainitic steel. They found that the ratcheting strain strongly depends on the stress level. At a low stress level a constant ratcheting strain rate was reached. At higher stress

levels the evolution of ratcheting strain with the number of RCF cycles displayed three stages, in manner similar to those obtained for vertical short crack kinetics.

Orringer [1997; 1988] studied the fatigue crack propagation life of detail fracture in rails. More recently, Glowacki and Kuziak [1997] have investigated the effect of coupled thermal-mechanical processes on the evolution of microstructure in rails. Head hardening behavior and rolling contact fatigue have also been studied [Wong et al. 1996; Muster et al. 1996]. Fracture toughness and fatigue strength have been applied as criteria for characterization of the fatigue damage resistance of railway rails [Vitez 1997]. The limitations of fracture toughness and fatigue strength criteria are obvious due to the diversified composition of rails and the complicated live bending stress and shear stress, and residual stress conditions. Another factor that must be addressed is that data on rail flaw growth for lifetime estimates is insufficient, which makes it difficult to further understand the fatigue behavior and assess the fracture resistance of railway rail alloys under cyclic loading conditions. Thus, it is necessary to acquire more fatigue data, to study the related fracture and failure mechanism and to propose new criteria of damage tolerance for rail steel systems.

In the current study, the microstructure-properties relationships, plane strain fracture toughness, and fatigue crack growth of a J6 bainitic and a premium pearlitic rail steel are studied to gain fundamental understanding of the underlying mechanisms between these performance-related properties and the microstructure of the materials. ASTM standard E399A [ASTM 1983] is used to determine and compare the plane strain fracture toughness of the two materials. The fatigue crack growth behavior of the two steels generated by the current authors is correlated with plane strain fracture toughness data for the two materials.

## 2. Materials and experimental approach

**Materials.** The materials used in this study were a J6 bainitic rail steel and a premium pearlitic rail steel provided by the Transportation Technology Center, Inc., Pueblo, Colorado. Both rails were new and had not been placed in service. The premium pearlitic rail steel was manufactured by Rocky Mountain Steel Mills, Pueblo, Colorado in 1998. Although the exact composition of the rail used in the current investigation is not known, it is assumed to be within the composition range for pearlitic steel as controlled by the American Railway Engineering Manufacturing Association (AREMA). This range is presented in [AREMA 1997, Table 1]. No reported details on the manufacture of the J6 bainitic rail are available, but it is believed that no heat treatment was done. The composition of bainitic rail steel as determined by CMS Colorado Metallurgical Services (J. Kristan, private communication, 2003) is given in Table 1. According to Sawley and Kristan [2003], J6 bainitic steel has a carbide-free structure with lath ferrite and interlath austenite in the as-rolled condition.

Element	C	Mn	P	S	Si	Ni	Cr	Cu	Mo	V
Bainitic steel	0.23	1.93	0.012	0.008	1.96	0.14	1.84	0.13	0.43	0.007
Pearlitic steel	0.72~0.78	0.60~1.25	0.035	0.037	0.1~0.6	0.25	0.25~0.5		0.1	0.03~0.05

**Table 1.** Composition of the rail steels in percent by weight.

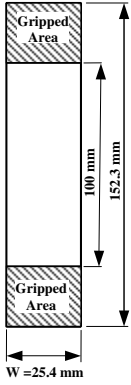
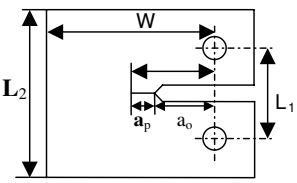
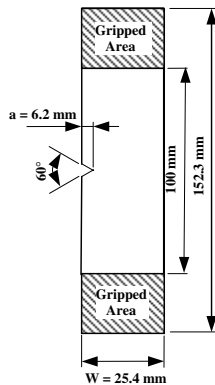
Mechanical Properties Rectangular	Fracture Toughness $\frac{1}{2}$ T compact tension	Fatigue Crack Growth Single edge notch
 <p>Thickness, <math>B = 2</math> mm</p>	 <p> <math>W = 22.86</math> mm  <math>a_0 = 8.89</math> mm  <math>a_p = \text{pre-crack length}</math>  <math>L_2 = 27.43</math> mm  <math>L_1 = 13.97</math> mm            Pin diameter = 6.35 mm  <math>B = 9.14</math> mm         </p>	 <p><math>B = 2</math> mm</p>

Figure 1. Specimen geometry.

**Specimen geometry.** All test specimens were cut from railheads along the direction of the rail using EDM (electric discharge machining). Samples were taken from the middle of the railheads in order to avoid top surface effects. This is especially important for the pearlitic railhead which has undergone head hardening by air quenching. The specimen geometries for each loading configuration are shown in Figure 1. The compact tension specimens were subjected to fatigue pre-cracking in order to introduce a sharp notch for the fracture toughness testing. An intended pre-crack length of about 2.5 mm was introduced into each specimen to have a target geometrical correction factor ( $a/W$ ) of 0.5. Difficulties were encountered initially when attempting to pre-crack the bainitic samples and therefore they were stress relieved at  $454^\circ\text{C}$  for 1/2 hour prior to pre-cracking. This observation raises a question about the inherent residual stress in the bainitic rail steel. Further research on understanding the residual stress level in the bainitic rail steel is needed. All of the pearlitic steel samples pre-cracked in the normal fashion with relatively straight pre-cracks without additional treatment. For the purposes of calculation the actual crack length was measured from the fracture surface following the tests.

**Experimental approach.** For mechanical properties and fracture toughness studies, static tensile experiments were performed using an MTS 810 materials testing system equipped with a 100 kN load cell. The MTS is equipped with TestStar II software. Load was applied at a rate of 20 kN/min. The fracture toughness data were analyzed according to ASTM standard E399.

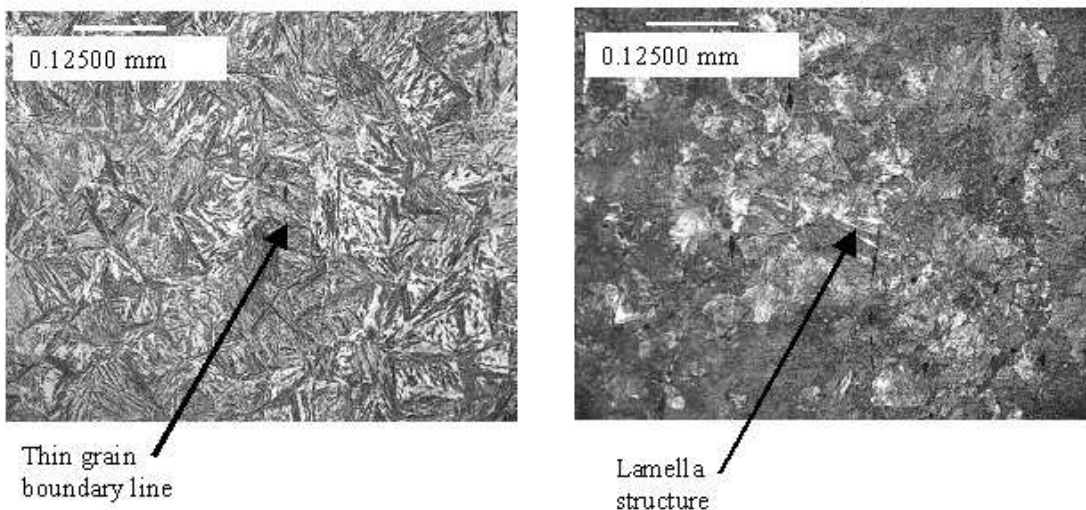
The fatigue tests were conducted at an ambient temperature of  $25^\circ\text{C}$  under load control conditions using a sinusoidal waveform and a frequency of 1 Hz. The maximum stress was 200 MPa, and the ratio of minimum stress to maximum stress was 0.1. The crack length at various intervals of number of cycles was recorded during the tests. A video camera with a zoom lens was used to view the crack tip region, measure the crack length and capture the damage associated with the crack growth. A total of three

samples were tested from each material, and data from each sample were used for the fatigue crack propagation analysis and the fatigue damage species examination.

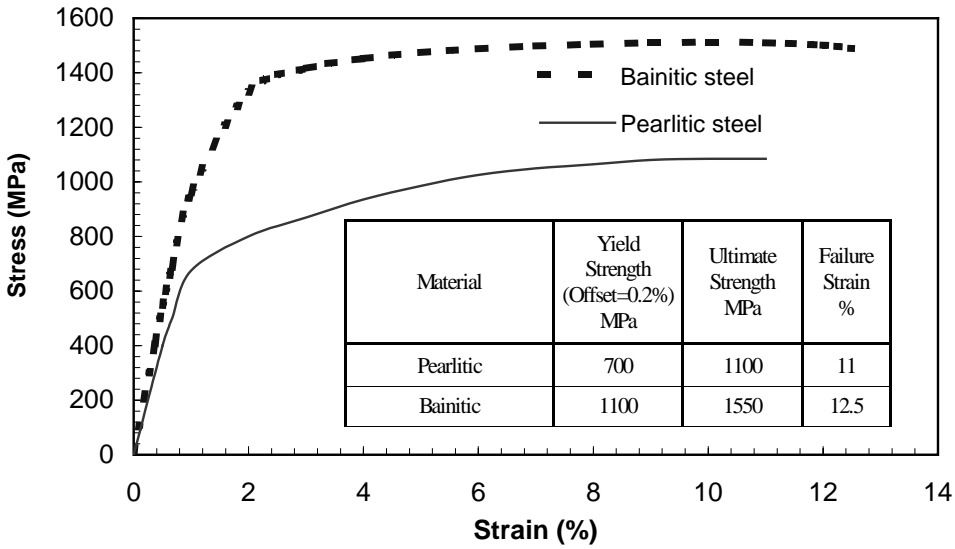
The microstructures of the bainitic and the pearlitic rail steels were studied using an optical microscope. Various microstructural features of each rail steel were identified. The fracture surfaces for all specimens were examined using a Hitachi S-2150 Scanning Electron Microscope operated at acceleration voltage of 20 kV. Typical micrographs revealing the fracture surface morphology were taken using a PCI-Image Management System.

### 3. Results and discussion

**Microstructural analysis.** Figure 2 shows the optical microstructures of bainitic and pearlitic steels are shown in The bainitic microstructures (left) reveals a mixture of tempered martensite and ferrite associated with intralath carbides. The autotempered martensite is distinguished from the bainite by the precipitation of multiple, rather than single, cementite structures within each lath, very similar to those reported in [Bramfitt and Speer 1990]. Typical pearlitic microstructure with fine lamellar aggregate of very soft and ductile ferrite and very hard carbide cementite is shown in the right half of Figure 2. The lamellae are aligned in the same orientation in one grain. The properties of pearlitic steels are mainly determined by the spacing between the ferrite-cementite lamellae. The grains of bainitic steels are easy to see and the grain boundary is very thin. Optical measurements of the grain size of pearlitic and bainitic steels were taken using OLYMPUS GX51 Inverted System Metallurgical Microscope and software-PAX-IT5.2. Twenty grains were sampled to determine the average grain size. Three diagonal measurements were taken for each grain and an average grain size was obtained. A total average was obtained based on the twenty grains. The average grain size of bainitic steel is about  $90\ \mu\text{m}$ , while the average grain size of pearlitic steel is about  $50\ \mu\text{m}$ . The microstructure of bainite is more complex than that of pearlite because of the wide range of cooling rates.



**Figure 2.** Optical microstructures of rail steels: (left) bainitic, (right) pearlitic.



**Figure 3.** Stress-strain relationships of pearlitic and bainitic rail steels.

**Stress-strain relationships.** The stress-strain relationships of bainitic and pearlitic rail steels based on unnotched specimens are shown in Figure 3. It can be seen that the ultimate strength, yield strength and strain to failure of bainitic steel are higher than that of pearlitic steel. The stress is calculated based on the original cross-sectional area before testing. A summary of these mechanical properties for J6 bainitic rail steel and a premium pearlitic rail steel are also given in Figure 3. There is an increase of about 36% in the ultimate strength, 77% in the yield strength and 15% in the strain to failure from the pearlitic to the bainitic rail steel. The higher strength of the J6 bainitic steel can be related to the distribution of carbide particles. Strength increases with increasing number of carbide particles per unit area in the plane of the section. By decreasing the bainite transformation temperature, the carbides will be dispersed, the size of carbide particles will be smaller, and the number of carbide particles in the plane of section will be large. Thus, the strength will increase greatly.

**Fracture toughness evaluation.** The general expression of the plane strain fracture toughness for a compact tension specimen used to compare the two rail steels is

$$K_1 = \frac{P}{B\sqrt{W}} f\left(\frac{a}{W}\right), \tag{1}$$

where  $P$  is the load,  $B$  is the thickness,  $W$  is the distance from the center of the loading holes to the edge of the specimen,  $a$  is the total crack length (initial plus pre-crack), and  $f(a/W)$  is a geometrical correction factor. The geometrical correction factor,  $(a/W)$ , in Equation (1) can be expressed as a function as follows [Srawley 1976; Newman 1974]:

$$f\left(\frac{a}{W}\right) = \frac{\left(2 + \frac{a}{W}\right) \left[0.886 + 4.64\left(\frac{a}{W}\right) - 13.32\left(\frac{a}{W}\right)^2 + 14.72\left(\frac{a}{W}\right)^3 - 5.6\left(\frac{a}{W}\right)^4\right]}{\left(1 - \frac{a}{W}\right)^{3/2}}$$



The bainitic steel demonstrated elastic behavior and cleavage fracture. Therefore ASTM Standard E399 was used for the analysis. The data for all three samples tested along with their calculated values of  $K_{Ic}$  based on Equation (1) are presented in Table 2. To validate the calculated value of  $K$  as a true  $K_{Ic}$  fracture toughness, the following conditions must be met:

$$a \text{ and } B \geq 2.5(Kq/\sigma_y)^2, \quad P_{\max}/P_q < 1.1. \tag{2}$$

For the bainitic steel, the value of  $2.5(Kq/\sigma_y)^2 = 2.5(52/1100)^2 * 1000 = 2.24$  mm. This value is less than  $B$ , which is 9.14 mm, and less than  $a$  which is about 12 mm. Therefore, the first condition in Equation (2) is met. In addition  $P_{\max}$  equals  $P_q$ , thus also meeting the second condition. Therefore these fracture tests for the bainitic steel yield a valid value of  $K_{Ic}$  according to ASTM E399. The average value of  $K_{Ic}$  for the bainitic steel was found to be 52 MPa  $\sqrt{m}$ .

The pearlitic steel also demonstrated elastic behavior and cleavage fracture. Similarly ASTM Standard E399 was used for the analysis of the plane strain fracture toughness,  $K_{Ic}$ . The data for all three premium pearlitic steels samples tested along with their calculated values of  $K_{Ic}$  based on Equation (1) are also presented in Table 2. For the pearlitic steel we have  $2.5(Kq/\sigma_y)^2 = 2.5(41/700)^2 * 1000 = 8.6$  mm. This value is less than  $B$ , which is 9.14 mm and less than  $a$ , which is about 12 mm; this meets the first condition in Equation (2). In addition  $P_{\max}$  equals  $P_q$ , so the second condition is met. Therefore these fracture tests for the pearlitic steel yield a valid value of  $K_{Ic}$  according to ASTM E399. The average value of  $K_{Ic}$  was found to be 41 MPa  $\sqrt{m}$  for the premium pearlitic steel.

Based on the plane strain fracture toughness analysis using compact tension specimens the fracture toughness of the bainitic steel was found to be about 27% higher than that of the premium pearlitic steel.

Sample	$W$ (mm)	$a_o$ (mm)	$a_p$ (mm)	$a$ (mm)	$(a/W)$	$f(a/W)$	$P_m$ (kN)	$P_q$ (kN)	$B$ (mm)	$Kq$ (MPa $\sqrt{m}$ )	$K_{Ic}$ (MPa $\sqrt{m}$ )
Bainitic Rail Steel											
#1	22.4	8.89	3.0	11.89	.531	10.66	6.6	6.6	9.14	51.43	51.43
#2	22.4	8.89	3.5	12.39	.553	11.48	6.3	6.3	9.14	52.87	52.87
#3	22.6	8.89	3.4	12.29	.544	11.13	6.4	6.4	9.14	51.44	51.44
											(average) 52
Pearlitic Rail Steel											
#1	22.4	8.89	3.2	12.09	.540	10.98	5.7	5.7	9.14	45.75	45.75
#2	22.3	8.89	3.3	12.19	.547	11.25	4.7	4.7	9.14	38.74	38.74
#3	22.4	8.89	3.7	12.19	.562	11.85	4.3	4.3	9.14	37.25	37.25
											(average) 41

**Table 2.** Compact tension specimen geometry and test results for bainitic and pearlitic steels. Notation:  $W$  is the distance from the center of the loading holes to the edge of the specimen,  $a_o$  the initial machined notch length,  $a_p$  the fatigue pre-crack,  $a$  the total crack length,  $P_m$  the maximum load,  $P_q$  the load at which  $K_q$  is calculated,  $B$  the thickness,  $K_q$  the stress intensity factor calculated from Equation (1), and  $K_{Ic}$  the mode I fracture toughness.

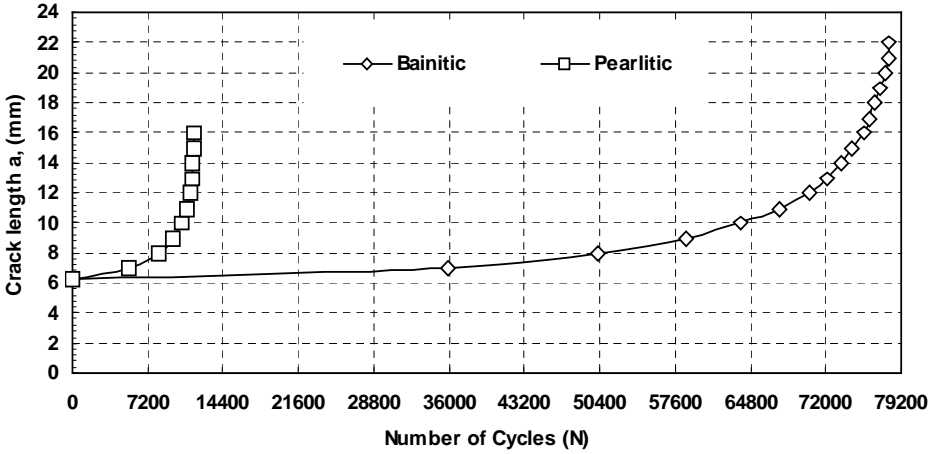


Figure 4. Fatigue crack length,  $a$ , versus the number of cycles,  $N$ .

**Fatigue crack growth analysis.** Figure 4 shows the average crack length  $a$  versus the number of cycles  $N$ , for both bainitic and pearlitic rail steel. We see that the total fatigue lifetime of bainitic steel is much higher than that of pearlitic steel. This is based on the average of three macroscopically identical specimens from each material. The total average fatigue lifetime for the bainitic and pearlitic steel samples is about 78000 and 11000 cycles, respectively. We also see that both initiation lifetime and propagation lifetime for the bainitic steel is higher than that for the pearlitic. The slopes of the curves in Figure 4 are taken as the average crack speed at each crack length.

The potential energy  $P$ , was calculated from the hysteresis loops recorded at intervals of number of cycles. It is the area above the unloading curve at each crack length. On this basis, the relationship

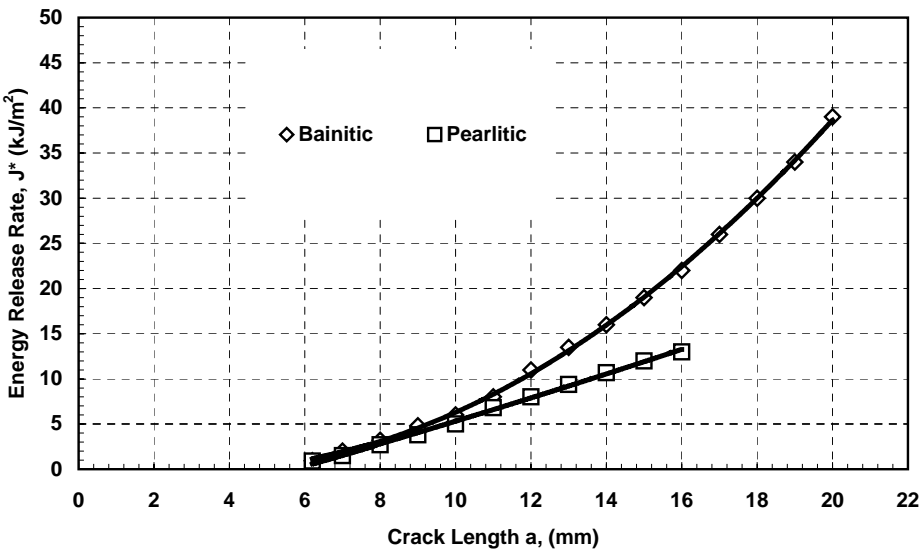


Figure 5. Energy release rate,  $J^*$ , versus the fatigue crack length,  $a$ .

between the potential energy and the fatigue crack length  $a$ , was established. The energy release rate  $J^*$  from the potential energy curve is based on

$$J^* = \frac{1}{B} \frac{\partial P}{\partial a},$$

where  $P$  is the potential energy,  $a$  is the crack length and  $B$  is the specimen thickness. Figure 5 illustrates the average energy release rate,  $J^*$ , as a function of the crack length  $a$ , for the bainitic and pearlitic rail steels. The value of  $J^*$  increases with the increase of the crack length  $a$ . The critical value of  $J^*$  for the bainitic steel is about 40 kJ/m<sup>2</sup>, while that for the pearlitic steel is about 13 kJ/m<sup>2</sup>. This is the point where the onset of rapid crack growth is observed. The ratio between the critical crack length and specimen width ( $\frac{a_c}{W}$ ) is higher for bainitic steel than that for pearlitic steel. The values of the ratio  $a_c/W$  are 0.8 and 0.63 for bainitic and pearlitic steels respectively.

The crack speed versus the energy release rate  $J^*$  for both bainitic and pearlitic rail steel are shown in Figure 6. The crack deceleration in the case of the bainitic steel started after a value of  $J^*$  of about 10 kJ/m<sup>2</sup>. This is indicative of material damage ahead of the crack tip. It can also be seen from the figure that the first stage, or initiation stage, of fatigue crack propagation (FCP) kinetics is well developed in the bainitic steel, while that for the pearlitic is less pronounced. Under the same  $J^*$ , the crack growth rate of pearlitic steel is higher than that of bainitic steel, which means that bainitic rail steel has higher resistance to FCP.

Fracture surface morphology examination was performed on typical fatigue-failed specimens of each material to identify the fatigue damage species. Micrographs (1000X) taken from the stable crack propagation region, ahead of the notch tip, are shown in Figure 7 for the bainitic and pearlitic rail steels respectively. Ductile tearing, and extensive ridge formation are associated with the bainitic steel (left); pulled-up pearlite lamella, microcracks and micro-void coalescence can be found in the pearlitic steel

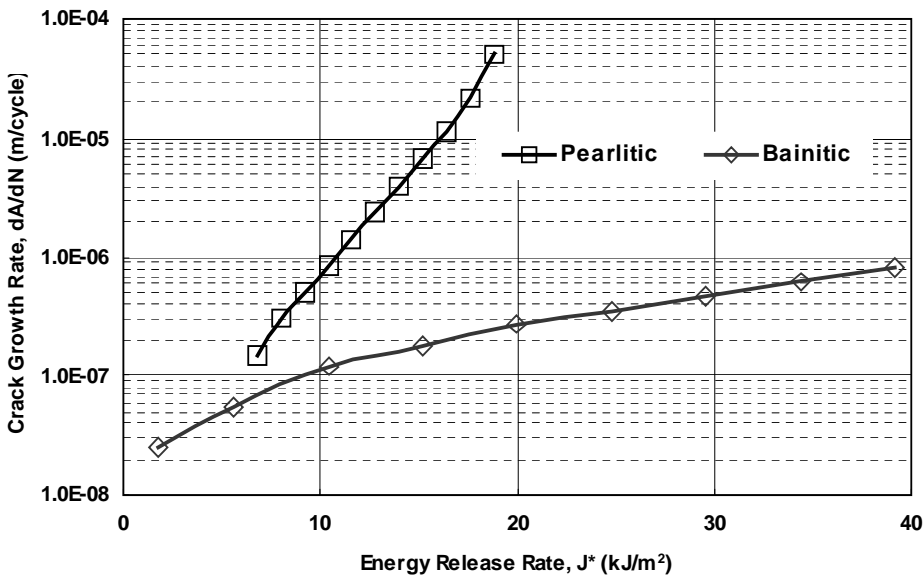
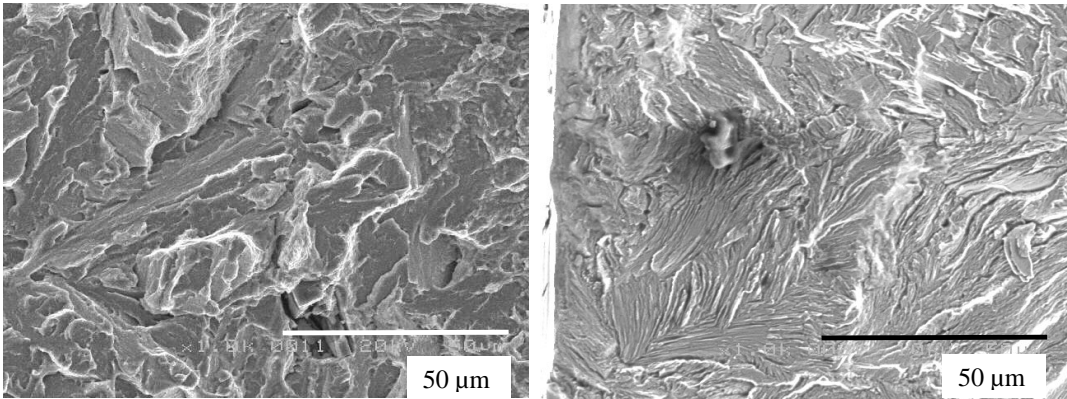


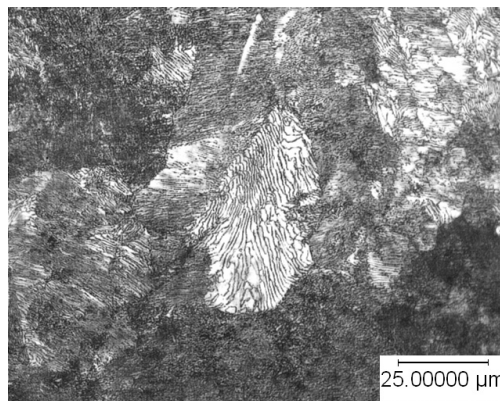
Figure 6. Crack growth rate,  $da/DN$ , versus energy release rate  $J^*$ .



**Figure 7.** SEM micrographs at 1000X taken from the beginning of the stable crack propagation region: (left) bainitic and (right) pearlitic steels.

(right). In general, these features reflect the crack deceleration and indicate a considerably high energy consuming process associated with the crack propagation. It appears to be the case that the bainitic rail steel displays more ductile fracture features in the stable region than the pearlitic steel.

To identify the microstructural origin of higher fracture toughness and fatigue damage tolerance of the bainitic rail steel as compared to pearlitic steel, we examine the optical micrograph of the pearlitic rail steel in Figure 8, which displays pearlitic colonies. The lamella structure is clearly visible. The properties of pearlitic steel are mainly determined by the spacing between the ferrite-cementite lamella shown in this grain [Aglan et al. 2004]. It has been stated [Yokoyama et al. 2002] that the lamella structure causes anisotropy in the physical and mechanical properties of the pearlitic steel, depending on lamella orientation. Under plastic deformation, high stress concentration can occur at the colony boundaries and cause cracking along these boundaries. The bainitic microstructure as seen in Figure 2 (left) does not



**Figure 8.** Optical micrograph of pearlitic railsteel showing the lamella structure in a pearlitic colony.

show strong microstructural anisotropy, reducing the crack initiation and growth sites. It is believed that this behavior is responsible for the higher fracture toughness and fatigue damage tolerance of the bainitic steel in comparison with pearlitic steels. This can also explain the better RCF damage resistance of the bainitic steel in comparison with pearlitic steel.

#### 4. Conclusions

Based on the testing of the rail steels and the related data analyses, the following conclusions are drawn:

1. The microstructure of J6 bainitic rail steel is more complex than that of pearlitic steel. The average grain size of bainitic rail steel is about  $90\ \mu\text{m}$  while that for pearlitic rail steel is about  $50\ \mu\text{m}$ . The J6 bainitic steel has ultimate strength, yield strength and elongation to failure of about 1500 MPa, 1100 MPa and 13% respectively. These values are higher than those for pearlitic steel.
2. Fracture toughness evaluation was conducted using  $(1/2T)$  compact tension specimens according to ASTM standard E399. It was found that the average  $K_{Ic}$  for the bainitic rail steel is  $52\ \text{MPa}\sqrt{m}$ , while that of the premium pearlitic steel is  $41\ \text{MPa}\sqrt{m}$ .
3. Crack speed for the bainitic steel is lower than that for the pearlitic steel over the entire range of the energy release rate. The bainitic steel exhibits a higher rate of crack deceleration in the second stage, as indicated by the lower slope of the FCP curve in comparison with pearlitic steel. This indicates the superior resistance of bainitic rail steel to FCP, that is, higher fatigue damage tolerance. This is in agreement with the plane strain fracture toughness data for the two materials.
4. Ductile tearing and extensive ridge formation are associated with the stable crack propagation region of the bainitic steel. Pulled-up pearlitic lamella, limited microcracks and micro-void coalescence can be found in the pearlitic steel. The more ductile features of the bainitic steel reflect the crack deceleration and indicate a considerably higher energy-consuming process giving it its superior fracture and fatigue crack propagation resistance.

#### References

- [Aglan et al. 2004] H. A. Aglan, Z. Y. Liu, H. M. F., and M. Fateh, "Mechanical and fracture behavior of bainitic rail steel", *J. Mater. Process. Technol.* **151**:1–3 (2004), 268–274.
- [AREMA 1997] *AREMA manual for railway engineering*, 4, American Railway Engineering Manufacturing Association, 1997.
- [ASTM 1983] ASTM, "Standard test method for fracture toughness of metallic materials", Technical Report E 399-90, American Society for Testing and Materials, Philadelphia, 1983.
- [Bramfitt and Speer 1990] B. L. Bramfitt and J. G. Speer, "Perspective on the morphology of bainite", *Metall. Trans. A* **21** (1990), 817–829.
- [Davis 1998] J. R. Davis (editor), *Metals handbook, desk edition*, 2nd ed., ASM International, Materials Park, OH, 1998.
- [Davis et al. 2002] D. Davis, C. Sasaoka, S. Singh, and D. Guillen, "Advanced design of bainitic steel rail crossing diamond under HAL service", Technical Report TD-02-024, Transportation Technology Center, Pueblo, CO, 2002.
- [Dimelfi et al. 1998] R. J. Dimelfi, P. G. Sanders, B. Hunter, J. A. Eastman, K. J. Sawley, K. Leong, and K. J., "Mitigation of subsurface crack propagation in railroad rails by laser surface modification", *Surf. Coat. Technol.* **106**:1 (1998), 30–43.
- [Glowacki and Kuziak 1997] M. Glowacki and R. Kuziak, "Application of coupled thermal-mechanical model to the numerical analysis of the microstructure development during rolling and air-cooling of rails", pp. 1313–1316 in *Fifth International Conference on Computational Plasticity* (Barcelona, 1997), edited by D. R. J. Owen et al., Pineridge, Swansea, 1997.

- [Jeong et al. 1998] D. Jeong, Y. Tang, and O. Orringer, "Estimation of rail wear limits based on rail strength investigations", Technical Report DOT/FRA/ORD-98/07, Federal Railroad Administration, Office of Research and Development, Washington, DC, 1998.
- [Kang et al. 2006] G. Z. Kang, Y. Li, Y. Gao, Q. Kan, and J. Zhang, "Uniaxial ratcheting in steels with different cyclic softening/hardening behaviours", *Fatigue Fract. Eng. Mater. Struct.* **29**:2 (2006), 93–103.
- [Kristan 2005] J. Kristan, "R&T results of AAR developed J6 bainitic type rail steel", TTCI Technology Digest TD-05-014, Transportation Technology Center, Pueblo, CO, 2005.
- [Li et al. 2006] Y. Li, G. Kang, C. Wang, P. Dou, and J. Wang, "Vertical short-crack behavior and its application in rolling contact fatigue", *Int. J. Fatigue* **28**:7 (2006), 804–811.
- [Muster et al. 1996] H. Muster, H. Schmedders, K. Wick, and H. Pradier, "Rail rolling contact fatigue: the performance of naturally hard and head-hardened rails in track", *Wear* **191**:1–2 (1996), 54–64.
- [Newman 1974] J. E. Newman, "Stress analysis of compact specimens including the effects of pin loading", pp. 105–121 in *Fracture analysis: proceedings of the 1973 National Symposium on Fracture Mechanics, II* (College Park, MD, 1973), edited by G. R. Irwin, ASTM Special Technical Publication **560**, ASTM, Philadelphia, 1974. Paper ID STP33136S.
- [Orringer 1988] O. Orringer, "Crack propagation life of detail crack fractures in rails", Technical report, Volpe Transportation Systems Center Report, 1988.
- [Orringer et al. 1997] O. Orringer et al., "Risk/benefit assessment of delayed action concept for rail inspection", Technical Report DOT/FRA/ORD-99/03 DOT-VNTSC-FRA-99-7, Federal Railroad Administration, Office of Research and Development, Washington, DC, 1997.
- [Sawley 1997] K. J. Sawley, "Bainitic steels for rails", Technical Digest TD97-001, Transportation Technology Center, Pueblo, CO, 1997.
- [Sawley and Kristan 2003] K. Sawley and J. Kristan, "Development of bainitic rail steels with potential resistance to rolling contact fatigue", *Fatigue Fract. Eng. Mater. Struct.* **26**:10 (2003), 1019–1029.
- [Smith 2004] W. F. Smith, *Foundations of materials science and engineering*, 3rd ed., McGraw-Hill, New York, 2004.
- [Srawley 1976] J. E. Srawley, "Wide range stress intensity factor expressions for ASTM E399 standard fracture toughness specimens", *Int. J. Fract. Mech.* **12**:3 (1976), 475–476.
- [Su and Clayton 1996] X. Su and P. Clayton, "Surface-initiated rolling contact fatigue of pearlitic and low carbon bainitic steels", *Wear* **197**:1–2 (1996), 137–144.
- [Vitez 1997] I. Vitez, "The fracture toughness and the fatigue strength of railway rails", *Int. J. Fatigue* **19**:10 (1997), 734–734.
- [Wong et al. 1996] S. L. Wong, P. E. Bold, M. W. Brown, and R. J. Allen, "A branch criterion for shallow angled rolling contact fatigue cracks in rails", *Wear* **191**:1–2 (1996), 45–53.
- [Yokoyama et al. 2002] H. Yokoyama, S. Mitao, S. Yamamoto, and M. Fujikake, "Effect of the angle of attack on flaking behavior in pearlitic and bainitic steel rails", *Wear* **253**:1–2 (2002), 60–66.
- [Yun et al. 1996] X. Yun, D. B. Bogy, and C. S. Bhatia, "Wear of hydrogenated carbon coated disks by carbon coated and uncoated Al<sub>2</sub>O<sub>3</sub>/TiC sliders in ultra high vacuum", *IEEE Trans. Magn.* **32**:5 (1996), 3669–3671.

Received 18 Jun 2006. Accepted 15 Sep 2006.

HESHMAT A. AGLAN: [aglanh@tuskegee.edu](mailto:aglanh@tuskegee.edu)

College of Engineering, Architecture, and Physical Sciences, 218 Foster Hall, Tuskegee University, Tuskegee, AL 36088, United States

MAHMOOD FATEH: [mahmood.fateh@dot.gov](mailto:mahmood.fateh@dot.gov)

Federal Railroad Administration, 1120 Vermont Ave., Washington, DC 20590, United States

## **BISTABLE STRUCTURES FOR ENERGY ABSORPTION I. METALLIC STRUCTURES UNDER TENSION**

ZACHARY WHITMAN AND VALERIA LA SAPONARA

This paper presents proof-of-concept experiments on metallic bistable structures, and is followed by a companion paper about experiments on composite bistable structures. A bistable structure is characterized by a stress/strain curve with stable branches separated by unstable branches. The authors were interested in a particular bistable structure, one that once activated, has a second stronger state which has the ability to sustain higher loads. This allows for a better distribution of damage. In addition, the structure keeps its integrity for a longer time, leading to a fail safe design. Results on metallic configurations under tensile loading are shown in this paper. In particular, chains with one, two, and three bistable elements of 5052-H32 aluminum were designed, manufactured, tested, and compared to their corresponding baselines. A strain energy increase from 11% to about 30% is shown with respect to the baselines. Moreover, a comparative study with A36 annealed mild steel and C10100 copper shows the effect of the different ductility and stiffness on energy absorption.

### **1. Introduction**

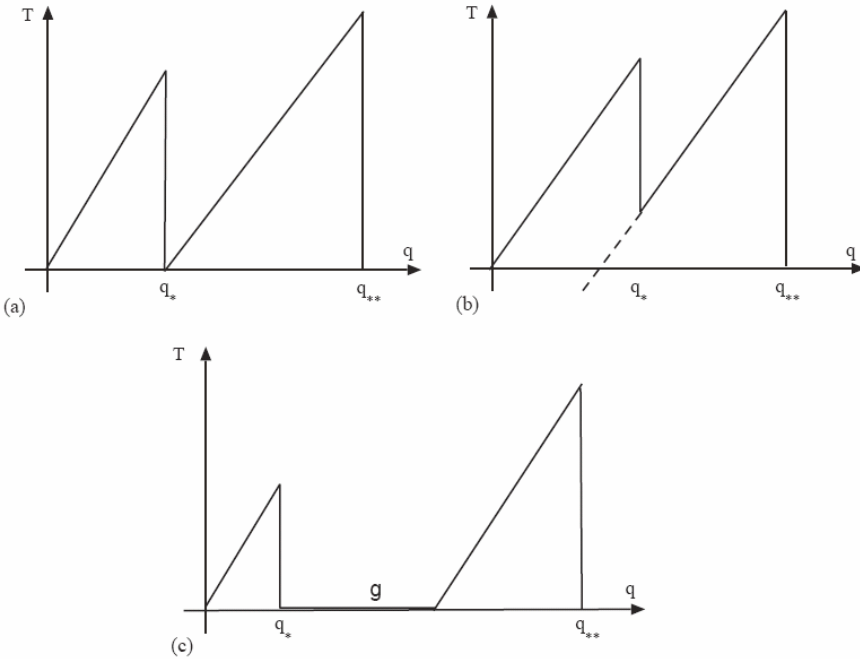
In traditional metallic structures under tensile loading, failure occurs when a weak location of the structure begins to yield, neck, and ultimately fracture. In such structures, only the energy absorption associated with yielding and fracture of a relatively small volume of the material is realized. Thus, the total energy absorbed is less than what would be available if the entire structure experienced necking and fracture. The metallic bistable structures discussed in this paper are designed to use necking and ductility properties of the material to absorb more energy, while maintaining the integrity of the structure. Results are shown for metallic chains under quasi-static tension. Chains with one to three bistable links of 5052-H32 aluminum alloy were examined, as well as the effect of ductility and strength of three different materials: A36 annealed mild steel, 5052-H32 aluminum alloy, and annealed C10100 copper.

In addition to ductility-driven metallic bistable structures, it is possible to design bistable structures made of composite materials, which are typically brittle. The design in this case is based on different mechanisms: the loading and unloading of the elastic structure, with tailoring of the strain capacity of the different subparts of the structure. Our companion paper discusses results for composite chains under quasi-static tension. The use of composites allows greater flexibility in tailoring the structure to the needed load carrying capability. The work in these two papers constitutes a first proof-of-concept of the ideas introduced by Cherkaev and Slepyan [1995]; see Figure 1.

---

*Keywords:* tension, metal, energy absorption, bistable.

This paper is based on work supported by the Army Research Office (grant ARO No. 41363-MA) and the National Science Foundation (grant CMMI-0621696, formerly CMS-0409514). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.



**Figure 1.** Bistable force-elongation diagrams are characterized by stable branches connected by unstable branches, taken in this illustration as piecewise linear functions. Reproduced from [Cherkaev et al. 2005], courtesy of Andrej and Elena Cherkaev.

We have been collaborating with A. Cherkaev and E. Cherkaev since the summer of 2004, and preliminary work was presented at two conferences [Whitman et al. 2005a; 2005b]. A. Cherkaev and L. Slepyan introduced the concepts of *main links* and *waiting links*. Main links are designed to break first, by having, for example, a smaller cross-sectional area, or by being made of a weaker material. Waiting links become active as the attached main link breaks. They provide a redundant load path and continue deforming under loading while preventing overall fracture. Moreover, the partially damaged state (the load/displacement of the overall structure after the first main link has broken and the attached waiting link becomes active) has a higher load-carrying capability than the undamaged state (the load/displacement of the overall structure before the first main link breaks). In these papers, chains made of main links and waiting links are defined as *bistable chains*. Further research and numerical simulations by Cherkaev, Slepyan and their collaborators showed that bistable chains are able to sustain a higher dynamic loading (in particular, impact and explosion) without breaking, compared to conventional chain structures, since they spread damage across a larger part of the volume [Cherkaev and Zhornitskaya 2004; Slepyan and Ayzenberg 1997; Cherkaev et al. 2005; Slepyan et al. 2005].

A literature survey showed that, while the idea of bistable structures is not new, its design, driven on material-based mechanisms and with the objective of higher energy absorption, is indeed novel. Bistable structures have been used for robotics, semiconductor research, smart morphing structures [Schultz 2005], deployable structures, and shells [Kebadze et al. 2004; Santer and Pellegrino 2004]. Dancila



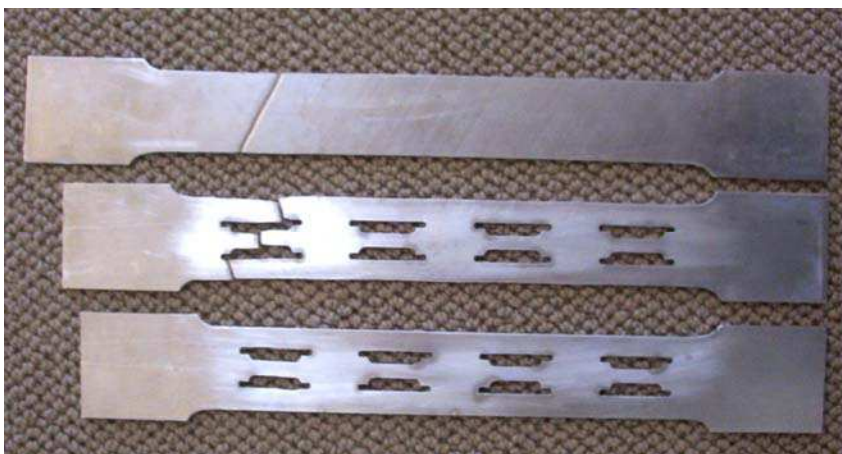
and Armanios [2001] did not use the word ‘bistable’, but showed a one-dimensional composite member with a progressive failure and yield-type response similar to the stress/strain curve of a bistable structure. Dancila [1998], in a Ph.D. thesis, discussed a flexible one-dimensional structure made of prepreg, Scotch 898 High Performance Filament Tape made by the 3M Corporation, embedded with E-glass fibers. No progress of the idea past these 1D specimens seems to have appeared in the literature.

Further work is currently being carried out on these concepts by the authors of this paper, in collaboration with A. Cherkaev, E. Cherkaev and L. Slepyan. Long-term applications are expected for design of crashworthy structures in mechanical, aerospace and civil engineering. Examples are the design of structural components for aircraft, floors, ceilings, crumple zones in cars, guard-rails, ship hulls, as well as design of redundant, ductile, and impact-resistant parts of civil infrastructures in earthquake-prone or hazard-prone areas.

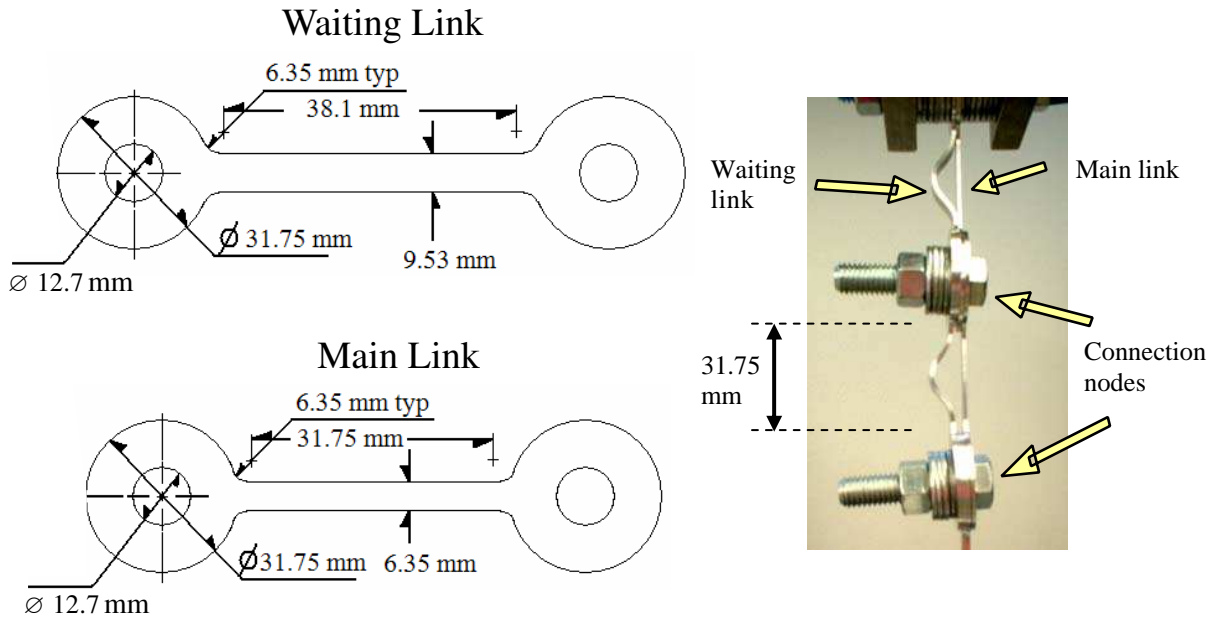
## 2. Design of a single bistable link

The ductility and strength of main links and waiting links need to be tailored appropriately to have the desired bistable behavior and synergistic mechanism with a controlled, high energy failure. The first studies were carried out on a mildly strain-hardened 5000 series aluminum, namely 5052-H32 aluminum. While the property of thermal hardening of 2000-, 6000- and 7000-series aluminum alloys would prove more useful for engineering applications, tests using a 5000-series configuration proved adequate for the scope of this study. The chosen alloy has a published elongation to failure of 12% [Oberg et al. 2000]. Figure 2 shows one of the first iterations, which failed.

By trial and error, a design with dog-bone shaped specimens was found to give the desired response. A dog-bone shape is also consistent with standard tensile testing (per ASTM standard E8). The material used was 1.59 mm thick. Dimensions and shapes of main links and waiting links are shown in Figure 3, left. The waiting link was curved to shape, and is 20% longer and 50% wider than the main link, as shown in Figure 3, right. The connection nodes, that is, the fasteners connecting the main link to a waiting link, had rather significant weight and stiffness with respect to the thin specimens.



**Figure 2.** Some first, unsuccessful designs. Specimens are made of 5052-H32 aluminum.



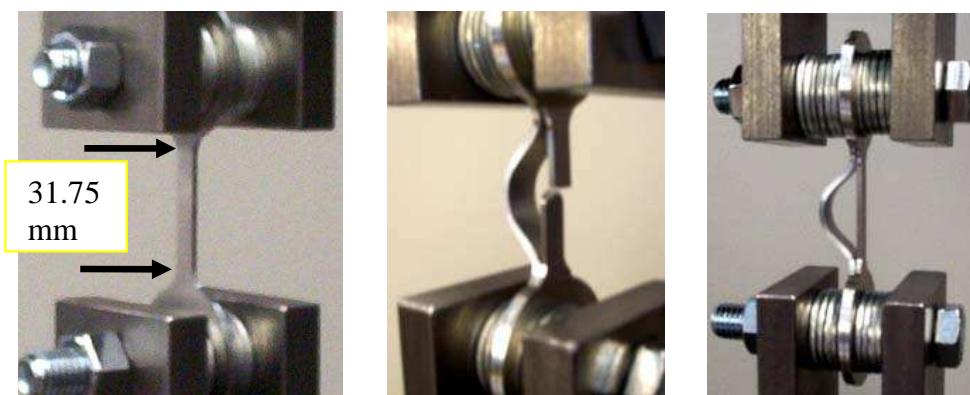
**Figure 3.** Left: Dimension of the main link and waiting link (sketch not to scale). The waiting link is 20% longer than the main link, that is, the percent wait is 20%. Right: Detail of main links, waiting links and connection nodes in an intact three-link chain of 5052-H32 aluminum, whose behavior under tension is shown in Figure 6.

Extensometers were not available which could span the gauge lengths of the specimens, so the cross-head displacement had to be used to assess the strain of the specimens. Any deformation outside the gauge lengths was to be avoided, which required the use of heavy and stiff connection nodes. The focus of this research was the behavior of the bistable links and not node or connection optimization. It was necessary to maximize the deformation of main and waiting links with respect to the connections.

Tests were conducted using either a 222 kN screw-driven machine or a 44.5 kN screw-driven machine (based on availability), at a rate of 1.5 mm/minute. The first machine was preferred, due to its simpler load train.

In this paper we introduce the notion of *percent wait*, the percent difference of the waiting link length relative to the attached main link length. The main link of a given material will fail at a particular elongation. Cherkhev and Slepian predicted that the waiting link would be disabled until the elongation to failure of the attached main link is reached. Using our definition of percent wait, this would mean that a percent wait equivalent to the elongation to failure would be needed for the mechanism to work.

A number of different materials, thicknesses and shapes were tested, which showed that the straightening of the curved waiting link could not be ignored in the process. It was observed through testing (discussed below) that an aluminum main link with a nominal elongation to failure equal to 15% needed a waiting link with percent wait equal to 20%, that is, an increase of 5 percent wait over that expected by the theory. This difference is due to the nonlinear response of the waiting link as it straightens. The waiting link not only had to be able to sustain the maximum load of the attached main link, but also a



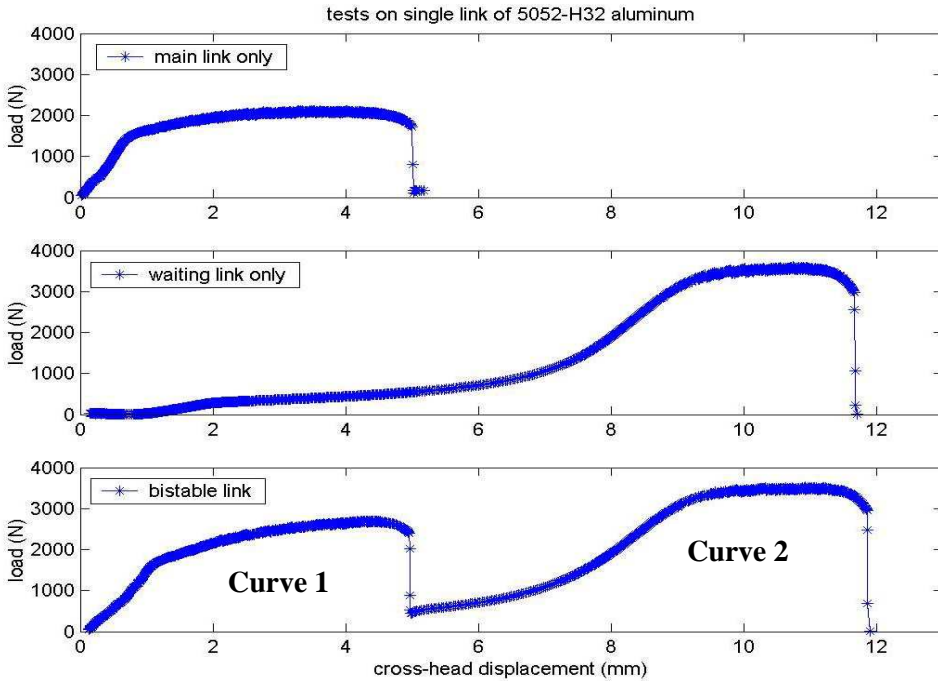
**Figure 4.** Tests on 5052-H32 aluminum. Left: a main link (Test A). Middle: a waiting link, where the main link has been cut (Test B). Right: full link (Test C). The load/displacement curves are shown in Figure 5.

sufficient additional load to allow the fracture of another main link in the chain (for chains having more than one main link and one waiting link). This is why the waiting link in Figure 3 is 20% longer than the main link.

The three tests that led us to the conclusions above were: Test A, a tensile test on a main link by itself, Test B, a tensile test to identify the behavior of a waiting link, and Test C, a tensile test on a main link connected to a waiting link. Figure 4 shows photographs of these tests. Test A (left pane) determined the overall elongation and maximum load of one single main link of 5052-H32 aluminum, following the ASTM standard E8. Test B evaluated the waiting link's response (middle pane). This required that the waiting link had the same node-to-node length as the main link. During the initial part of the test, a waiting link by itself would elongate past this node-to-node distance as the load increased to the 44.5 N preload. To avoid this, a main link was added and connected to the waiting link, forming the full link. After the preload was established — that is, when there was no slack in the load train — the main link was cut across its section, ensuring further testing of a waiting link of the correct dimensions. A nonlinear load/cross-head displacement curve was obtained as a result of the waiting link's straightening and strain hardening effects caused by the initial curved shape of the link. The acceptable waiting link dimensions and response were thus determined experimentally. Finally, Test C was carried out on the full bistable link, that is, the main link plus waiting link (Figure 4, right). Figure 5 shows the load/displacement curves corresponding to these tests.

**Test A.** The main link had length of 31.75 mm. The maximum cross-head displacement recorded was 5.15 mm, thus resulting in a 16.2% elongation to failure ( $5.15/31.75 \cdot 100$ ). The maximum load was reached at an elongation of 10.3% ( $3.27/31.75 \cdot 100$ ). Thus, the final 5.9% elongation was due to necking.

Had the main link been twice as long, the elongation to failure would have been different: the main link would presumably have reached the same maximum load at the same elongation, 10.3%, but the strain energy involved in the link's further extension and necking (the additional 1.09 mm) would have been associated with twice as much length, with a resulting elongation of  $1.09/(31.75 \cdot 2) \cdot 100 = 1.72\%$ .



**Figure 5.** Load/displacement curves for tests A, B, C of the 5052-H32 aluminum specimens shown in Figure 4. Responses of main link (top), waiting link (center), and full main link plus waiting link (bottom).

This, added to the elongation at maximum load, 10.3%, would result in a 12.0% elongation to failure. These considerations show that testing on the main link is needed, because the elongation to failure changes with the length of the specimen.

**Test B.** As discussed above, the elongation to failure of the main link was 16.2%. The waiting link's elongation to failure was a smaller 14%. Some of the strain energy involved in the process was used to straighten the waiting link, which had been initially bent to shape, and was 6.36 mm longer than the main link. Hence, the elongation to failure was calculated by subtracting 6.36 mm from the overall 11.7 mm. Also, the maximum load for the waiting link was 3602 N, a 70% increase with respect to the maximum load of the main link. Recall that the waiting link is 50% wider than the corresponding main link, hence a 50% increase should have been obtained on account of this width difference. The additional 20% increase of maximum load should be attributed to the waiting link's strain hardening process, since it was curved to shape and then straightened during testing. This also illustrates the difficulty of predicting response of the structure without conducting experiments.

**Test C.** The full link was tested next. As mentioned earlier, the waiting link associated with a failed main link had to carry not only the maximum load of another intact main link but also additional load due to the straightening of the latter's waiting link. This allowed for greater energy absorption of the chain. Curve 1 in Figure 5 corresponds to the bistable link, which is undamaged until the main link

breaks, indicated by the drop in Curve 1. Peak load is 27.2% higher than the maximum load for Test A (the tensile behavior of the main link by itself). Curve 2 corresponds to the now partially damaged link (following Curve 1, the main link is now broken, and the load is carried by the waiting link still attached to the main link). The peak load of Curve 2 is 2.23% lower than the maximum load for Test B (the tensile behavior of the waiting link by itself), but it is also 30.6% higher than the peak of Curve 1 (the tensile behavior of the bistable link before the main link's failure). Therefore, the curve not only exhibits bistable behavior, but has a partially damaged state with higher load carrying capability than its corresponding undamaged state.

The strain energies associated with these tests were calculated by numerical integration using the trapezoid rule (through the MATLAB routine 'trapz'), and were found to be 8.89J for Test A (the main link), 16.0J for Test B (the waiting link), and 25.1J for Test C (the full link), which is equivalent to the sum of the first two tests.

The following section will show that this bistable structure can absorb more energy than its baseline, depending on the number of links in the chain.

### 3. Proof of increased energy absorption of bistable links

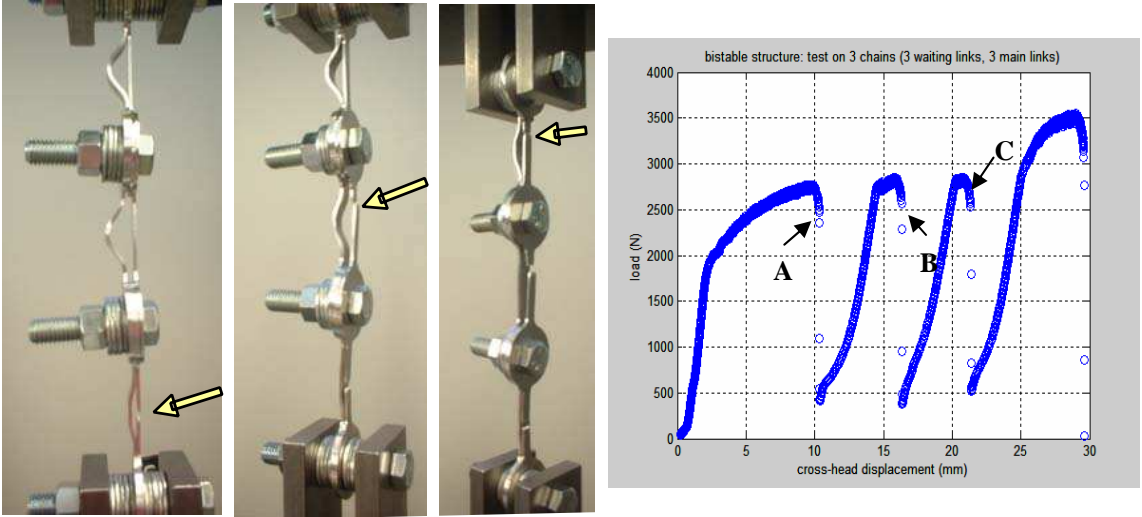
Three different chains of bistable links, manufactured with 5052-H32 aluminum, were tested to investigate the effect of additional necking on the overall energy absorption of the specimens, and to identify the actual increase of energy absorption of these configurations. Also, three corresponding baseline specimens were tested, where the formation of a single necking region caused the failure of the entire specimen. These latter specimens had the same cross-section as each full bistable link. Recall that the main link and the waiting link were made of 1.59 mm thick sheet material, and were 6.35 mm and 9.52 mm wide, respectively. The baseline had the same thickness and a width of 15.9 mm, the sum of the two widths. The baseline length was one, two and three times the length of the single bistable link plus an additional 12%. This 12% accounts for the additional material present in the waiting link, which is 20% longer than its corresponding main link. Normalizing with respect to the width, one obtains  $0.20 \cdot 9.52/15.9 = 12\%$ . The material used in the nodes was not accounted for in the baseline.

Figure 6 shows a tensile experiment on a three-link chain with its corresponding load/displacement curve. Figure 7 shows the load/displacement curves for all bistable and baseline specimens. The absorbed energy was calculated, and is reported in Table 1. These data indicate that at least three bistable links would be needed for a consistent increase of energy. For three bistable elements, the energy increase is about 30% with respect to the baseline specimens.

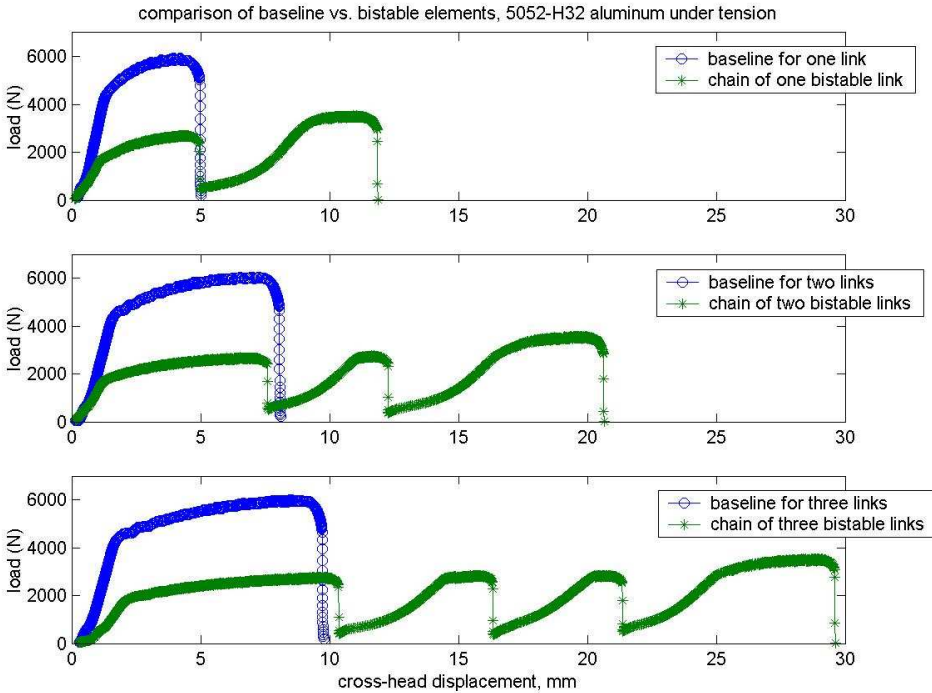
Future work will include uncertainty analysis. The goal of the current study was to identify a metallic bistable structure with the desired behavior. The authors plan to optimize test conditions, manufacturing, or any other elements of the process that would reduce data scatter for metallic bistable configurations.

### 4. Comparison of materials with different ductility

An investigation was carried out using two ductile materials in addition to the 5052-H32 aluminum: A36 annealed mild steel, and C10100 annealed copper. These materials were selected because of their published high ductility. This study aimed to assess how materials with different ductility and strength compared to each other. The dimensions of the main links and waiting links were the same for all



**Figure 6.** A chain of 5052-H32 aluminum with three bistable links. Left: Experiment. Arrows highlight the fracture and straightening of the waiting links in time. Right: Load/displacement curve. Points A, B, and C correspond to fracture of the first, second, and third main links.



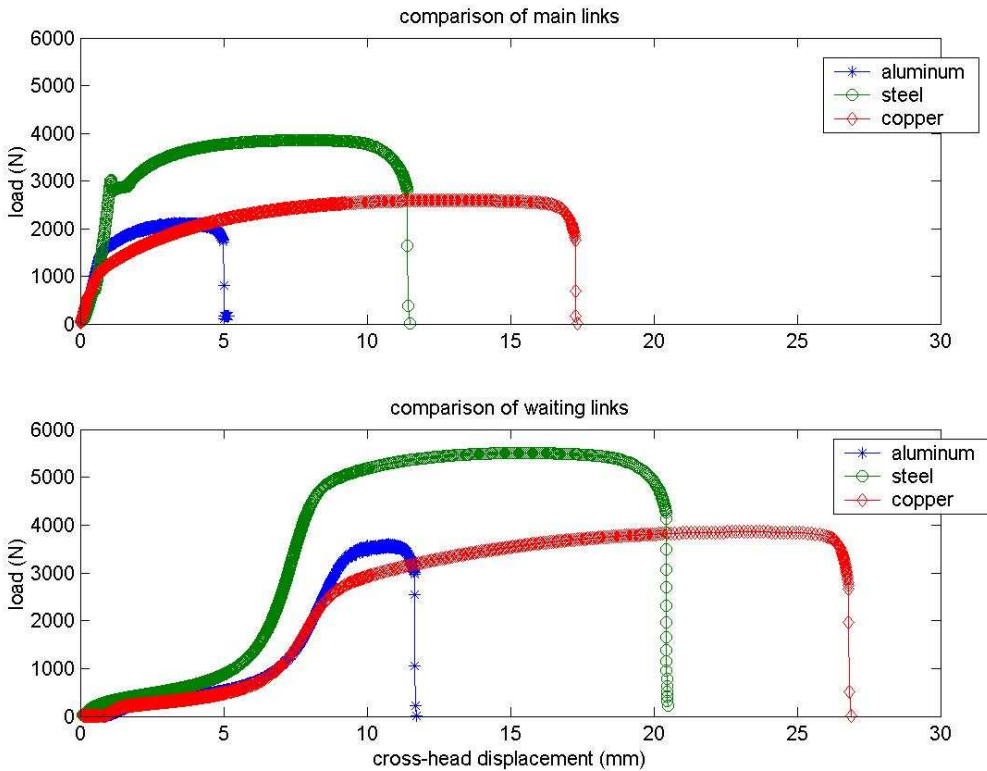
**Figure 7.** Comparison of bistable links and baseline links. From top to bottom, displacement/load curves for one, two, and three links of 5052-H32 aluminum.

<i>Number of links</i>	<i>Energy of bistable specimens (J)</i>	<i>Energy of baseline specimens (J)</i>	<i>Bistable/baseline energy ratio (%)</i>
1	25.1	22.3	112.4
2	43.1	38.6	111.6
3	61.3	46.7	131.3

**Table 1.** Energy absorbed in the yielding and fracture process of one to three links of bistable elements (5052-H32 aluminum), compared to corresponding baselines.

materials, that is, those shown in Figure 3. Tensile experiments were run on the main links alone and on the waiting links alone. Tensile experiments on the bistable link were omitted since the percent wait, 20%, was less than the elongation to failure for steel and copper, which prevented the desired behavior in the resulting main link + waiting link (that is, a second phase stronger than the first).

Figure 8 shows the load/displacement curves for the three materials. Steel was considerably stronger than the other materials, while copper had the highest elongation.



**Figure 8.** Comparison of main links and waiting links for three materials: 5052-H32 aluminum, A36 mild steel, and C10100 annealed copper.

Table 2 reports data on elongation to failure, elongation at maximum load, maximum load, energy absorbed, and its normalized value with respect to the density. Note the differences in the behavior of main links and waiting links regarding elongation, maximum load and absorbed energy.

Assume that necking starts at the maximum load. The aluminum main link, for example, will spend  $(16.2 - 10.3)\% = 5.9\%$  of its elongation in necking and fracturing, while its associated waiting link will spend  $(14 - 11.5)\% = 2.5\%$  in the same physical operation.

Inspecting maximum load, one can expect the maximum load of waiting links to be 1.5 times higher than for the corresponding main links, because the waiting link has a 50% wider section. This turned out to be a relatively good estimate (within 5%) for steel and copper, but not for aluminum, where the waiting link had a maximum load 13% higher than the predicted value. Therefore, it seems that the strain hardening in aluminum due to the bending to shape and straightening of the waiting link indeed does increase the maximum load. Steel and copper, however, had a small decrease of maximum load (within 5%).

Considering energies absorbed, it was expected that waiting links, which are 20% longer and 50% wider than the main link, would absorb  $(1 + 0.50) \cdot (1 + 0.20) = 1.8$  times more energy than the corresponding main link. This turned to be a correct estimate for aluminum and copper (within 1%). However, steel absorbed 7.6% more energy than expected.

Finally, one can calculate the absorbed energy per density (using  $2680 \text{ kg/m}^3$  for 5052-H32 aluminum,  $7850 \text{ kg/m}^3$  for A36 steel, and  $8940 \text{ kg/m}^3$  for C10100 copper). In this latter case, the steel main link and waiting link are found to have superior energy absorption/density with respect to the other materials investigated.

We see from our comparative study of the three materials that an estimate for maximum load and absorbed energy may be a reasonable approximation for some materials, but not others. It may not capture different behavior as the waiting links are first bent to shape and then straightened under loading. It is important to test waiting links in order to correctly design a bistable structure.

	<i>Elongation to failure (%)</i>	<i>Elongation at maximum load (%)</i>	<i>Maximum load (N)</i>	<i>Absorbed energy (J)</i>	<i>Absorbed energy / density (J m<sup>3</sup>/kg)</i>
<b>Main link</b>					
<i>Aluminum</i>	16.2	10.3	2122	8.9	0.0033
<i>Steel</i>	36.0	24.8	3854	38.6	0.0049
<i>Copper</i>	54.4	40.8	2594	38.6	0.0043
<b>Waiting link</b>					
<i>Aluminum</i>	14.0	11.5	3603	16.0	0.0060
<i>Steel</i>	37.0	23.9	5506	74.8	0.0095
<i>Copper</i>	53.7	43.3	3855	69.9	0.0078

**Table 2.** Data for main links and waiting links made with three materials: 5052-H32 aluminum, A36 annealed mild steel, and C10100 annealed copper.



Further studies on metallic bistable structures will adopt steel for both main links and waiting links, because of the higher performance per density of this material compared to aluminum and copper.

## 5. Summary and conclusions

This paper presents an experimental investigation of metallic bistable structures loaded in quasi-static tension. The concepts of bistable structures for energy absorption were introduced by Cherkaev, Slepian and their collaborators starting in 1995. However, this paper is the first actual experimental study which backs their theories. Bistable structures are material-based mechanisms that may absorb more energy than their corresponding baselines because of a more efficient way of using the material's ductility and multiple necking regions. Moreover, the two main components of these structures, *main links* and *waiting links*, are designed to ensure a redundant load path and a delayed and controlled failure. Results were shown and discussed as follows for the different investigations:

***Design of a main link and waiting link such that the corresponding bistable link has a partially damaged state which can absorb a higher load than the undamaged state.*** Trial-and-error experimental testing on 5052-H32 aluminum (Tests A, B, and C in this paper) allowed us to design appropriate dimensions of main links and waiting links to reach this goal. It is observed that the response of the waiting link cannot be predicted by conventional tensile testing, and that the so-called *percent wait* (the percent difference in waiting link length with respect to the attached main link length) may need to be higher than the elongation to failure predicted by Cherkaev, Slepian, and their collaborators.

***Study of the energy absorption with respect to the number of bistable links.*** The behavior of chains with one, two and three bistable links of 5052-H32 aluminum was compared to the corresponding baselines of the same material. Results show that a three-link bistable chain absorbs 30% more energy than its baseline.

***Study of main links and waiting links of three different materials: 5052-H32 aluminum, A36 mild, annealed steel and C10100 annealed copper.*** Main links and waiting links had dimensions as in Figure 3. This work shows that predictions of maximum loads and absorbed energy may not capture correctly the behavior of waiting links for all materials, thus testing is recommended. Moreover, main links and waiting links made of steel turned out to have a better performance (in terms of maximum loads and energy absorbed per unit density) than for the other materials.

In conclusion, this paper discusses initial experimental work on metallic bistable structures for energy absorption. Further investigation will concentrate on identifying the role of the connection nodes, collecting more data for statistical significance, optimizing the manufacturing and testing process to reduce data scatter, using steel for main links and waiting links, and extending these concepts to different types of loading (such as bending, shear and impact). At this stage, it is not possible to give guidelines for designing such structures, but future work will include optimization and numerical simulations, and will allow us eventually to reach this goal. The authors strongly believe that bistable structures have the potential to offer revolutionary energy absorption and crash worthiness compared to traditional structures, and that they would be very useful in applications where fail-safe behavior is critical.

## Acknowledgments

The authors thank Dr. Andrej Cherkaev and Dr. Elena Cherkaev, Department of Mathematics, University of Utah, for their collaboration. The experimental work was conducted in the Strength of Materials Lab and in the Composite Manufacturing Lab, Department of Mechanical Engineering, University of Utah, the former institution of the two authors.

## References

- [Cherkaev and Slepyan 1995] A. Cherkaev and L. Slepyan, “Waiting element structures and stability under extension”, *Int. J. Damage Mec.* **4**:1 (1995), 58–82.
- [Cherkaev and Zhornitskaya 2004] A. Cherkaev and L. Zhornitskaya, “Dynamics of damage in two-dimensional structures with waiting links”, pp. 273–284 in *Asymptotics, singularities and homogenisation in problems of mechanics* (Liverpool, 2002), edited by A. B. Movchan, *Solid Mech. Appl.* **113**, Kluwer, Dordrecht, 2004.
- [Cherkaev et al. 2005] A. Cherkaev, E. Cherkaev, and L. Slepyan, “Transition waves in bistable structures. I. Delocalization of damage”, *J. Mech. Phys. Solids* **53**:2 (2005), 383–405.
- [Dancila 1998] D. S. Dancila, *Energy-dissipating composite members with progressive failure*, Ph.D. thesis, School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, 1998.
- [Dancila and Armanios 2001] D. S. Dancila and E. A. Armanios, “Energy-dissipating composite members with progressive failure: concept development and analytical modeling”, *AIAA Journal* **40**:10 (2001), 2096–2104.
- [Kebadze et al. 2004] E. Kebadze, S. D. Guest, and S. Pellegrino, “Bistable prestressed shell structures”, *Int. J. Solids Struct.* **41**:11–12 (2004), 2801–2820.
- [Oberg et al. 2000] E. Oberg, F. D. Jones, H. L. Horton, and H. H. Ryffel, *Machinery’s handbook*, 26th, Industrial Press, New York, 2000.
- [Santer and Pellegrino 2004] M. J. Santer and S. Pellegrino, “An asymmetrically-bistable monolithic energy-storing structure”, in *45th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference* (Palm Springs, CA), AIAA, Reston, VA, 2004. Paper 2004-1527.
- [Schultz 2005] M. R. Schultz, “A new concept for active bistable twisting structures”, pp. 244–252 in *Smart structures and materials: Smart structures and integrated systems* (San Diego), edited by A. B. Flatau, *Proceedings of SPIE* **5764**, Internat. Soc. Optical Eng., Bellingham, WA, 2005.
- [Slepyan and Ayzenberg 1997] L. Slepyan and M. V. Ayzenberg, “Superplastic protective structures”, pp. 252–259 in *Progress in industrial mathematics*, edited by M. Brons et al., Teubner, Stuttgart, 1997.
- [Slepyan et al. 2005] L. Slepyan, A. Cherkaev, and E. Cherkaev, “Transition waves in bistable structures, II: Analytical solution: wave speed and energy dissipation”, *J. Mech. Phys. Solids* **53**:2 (2005), 407–436.
- [Whitman et al. 2005a] Z. Whitman, V. La Saponara, D. Adams, S. Leelavanichkul, A. Cherkaev, E. Cherkaev, and V. Vinogradov, “Advancements in fail safe response with bistable composite structures”, in *Proceedings of the ASME Winter conference* (Orlando, FL), New York, 2005. Paper #IMECE 2005-80627.
- [Whitman et al. 2005b] Z. Whitman, V. La Saponara, D. Adams, S. Leelavanichkul, A. Cherkaev, E. Cherkaev, and V. Vinogradov, “Improvement in energy absorption through use of bistable structures”, in *Proceedings of the SAMPE 2005 International Symposium and Exhibition*, Long Beach, CA, 2005.

Received 30 May 2006. Revised 31 May 2006. Accepted 26 Sep 2006.

ZACHARY WHITMAN: zachary.whitman@swri.org  
Southwest Research Institute

VALERIA LA SAPONARA: vlasaponara@ucdavis.edu  
Department of Mechanical and Aeronautical Engineering, One Shields Ave, University of California, Davis, CA 95616-5294  
<http://mae.ucdavis.edu/vlasaponara/>

## BISTABLE STRUCTURES FOR ENERGY ABSORPTION II. COMPOSITE STRUCTURES UNDER TENSION

ZACHARY WHITMAN AND VALERIA LA SAPONARA

This paper discusses proof-of-concept experiments of composite bistable structures and is preceded by a companion paper about metallic bistable structures. A bistable structure is characterized by a stress/strain curve with stable branches separated by unstable branches. We were interested in a particular bistable structure: one that, once activated, has a second stronger state with the ability to sustain higher loads. This allows for a better distribution of damage, and in addition the structure keeps its integrity for a longer time. Positive results on metallic configurations under tensile loading have prompted us to try and identify an equivalent composite configuration, able to show directional control of damage through this material-driven mechanism. Chain-like configurations with different reinforcements (T-300 carbon fibers, Kevlar® 49, E-glass and Dyneema®), polyurethane foam core and Kevlar stitching were conceived, designed, fabricated and investigated. Dyneema prototypes turned out to exhibit the desired bistable behavior as well as higher energy absorption than their baselines.

### 1. Introduction

Composites are typically characterized by high strength-to-weight and high stiffness-to-weight ratios, resistance to corrosion, and low magnetic signature. They can be tailored for a specific application. However, they are quite susceptible to delamination due to impact loading, for example, and a reduced load-carrying capacity may arise as a consequence of the impact event. Also, their brittleness is not attractive for applications where ductility, redundancy and high energy absorption are required, as in civil engineering.

Bistable structures for energy absorption were introduced by [Cherkaev and Slepyan 1995]. As discussed in the companion paper [Whitman and La Saponara 2007], the authors were asked to produce a proof-of-concept of these structures. Chain-like configurations were designed, manufactured and tested in tension. They behaved in a bistable manner, had higher energy absorption than the corresponding baselines, and a controlled type of failure. These results were obtained using metals such as 5052-H32 aluminum, C10100 copper and A36 steel.

The concepts of *main link*, *waiting link*, *percent wait* were also discussed, and will be repeated here for clarity sake. The *main link* is the part of the structure which is designed to break first. As it does, the structure does not lose its integrity because of the attached *waiting link*, which provides a redundant load path. The waiting link becomes active when the attached main link breaks. The interaction of main link and waiting link can be tuned so that the structure has the desired behavior. The *percent wait* is

---

*Keywords:* composite, energy absorption, bistable, Dyneema.

This paper is based upon work supported by the Army Research Office (grant ARO No. 41363-MA) and the National Science Foundation (grant CMMI-0621696, formerly CMS-0409514). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

the percent difference in waiting link length with respect to the attached main link length. It should be equal to the elongation to failure of the main link, however, the nonlinear behavior of the waiting link shows that this is not the case, and further tune-up of the percent wait is required for this material-driven mechanism to work.

Here we investigate structural configurations made with composite materials. The goal is, once again, to have a progressive failure where parts of the structure become active after a trigger damage event has occurred, and the damaged structure can carry a higher load with respect to its undamaged state and, more importantly, absorb more energy than its baseline. To the best of our knowledge, this study is novel.

## 2. Design and manufacturing

A variety of geometric configurations and percent waits was designed and manufactured with a waiting link that tried to mimic the curved waiting link present in the metallic configurations [Whitman and La Saponara 2007]. There were twenty-one iterations of this process (Table 1): nine had  $12 \times 12$  plain weave<sup>1</sup> T300 carbon from Toray Industries Inc., four used E-glass 7715 unidirectional cloth, three had a  $17 \times 17$  satin Kevlar® 49 type 285 weave, and five utilized a  $10 \times 10$  satin weave made with Dyneema®, a polyethylene fiber donated by DSM Corporation. Figure 1 shows the four materials tested, Figure 2 three of the geometries used ('small round core', 'medium round core' and 'large round core').

The weight and thickness of the fabrics are given in Table 2. The first three fabrics have almost the same thickness, and Dyneema is almost twice as thick — as mentioned, this material was donated. Note, however, that Dyneema's weight is very similar to T300 carbon's, so one may extrapolate a comparison of specific energy absorptions.

Each 'chain' has three elements. This number was chosen because the metallic 5052-H32 aluminum bistable chains with three elements had a consistent increase in energy absorption of about 30% with respect to the baseline, and three elements were adopted as starting point of the research on composite bistable structures [Whitman and La Saponara 2007].

In addition to the reinforcement fabrics mentioned above, the following materials were used:

- (a) Last-a-Foam 6704 core, with density  $64 \text{ kg/m}^3$  ( $4 \text{ lbs/ft}^3$ ). This is a closed cell polyurethane foam which is quite easy to cut (with a knife — not with hot wire, which will melt the foam), shape and machine (see Figure 3).
- (b) Kevlar 29 thread. As a main link break, the entire structure is subject to a shock. To reduce the chances of delamination, stitching was deemed necessary. The Kevlar 29 has shown to have the appropriate strength for the job, while being flexible enough to go through parts of the Consew sewing machine available.

The manufacturing process is similar for all configurations, and will be discussed here, with the aid of pictures representing one of the fabrics, Dyneema. A 152 mm wide  $\times$  406 mm long piece of fabric was cut, which is the minimum size that guarantees three elements. Most iterations required two plies for the waiting link and one for the main link (Table 1). The size of the cuts also fit in the 355 mm diameter quilting hoop required to hold the bottom ply (for the main link) tight, and to ensure that the plies forming

<sup>1</sup> $12 \times 12$  means that there are 12 bundles in the warp direction and 12 bundles in the fill direction.



**Figure 1.** The four materials used: from top to bottom, T300, E-Glass, Kevlar 49, Dyneema. The width of all specimens is 25.4 mm.

the waiting link had the planned 0/90 orientation. The foam core was cut and sanded to the desired shape. In most iterations, it was found that wrapping the foam in Teflon®-coated fiberglass would help. Finally, engineering paper and plain paper were used for the stitching process, respectively to use as ‘cross-hairs’, and to protect the bottom part of the fabric during the stitching process (Figures 4 and 5). Figures 6 and 7 show additional steps of the stitching process and the specimen ready for infiltration.

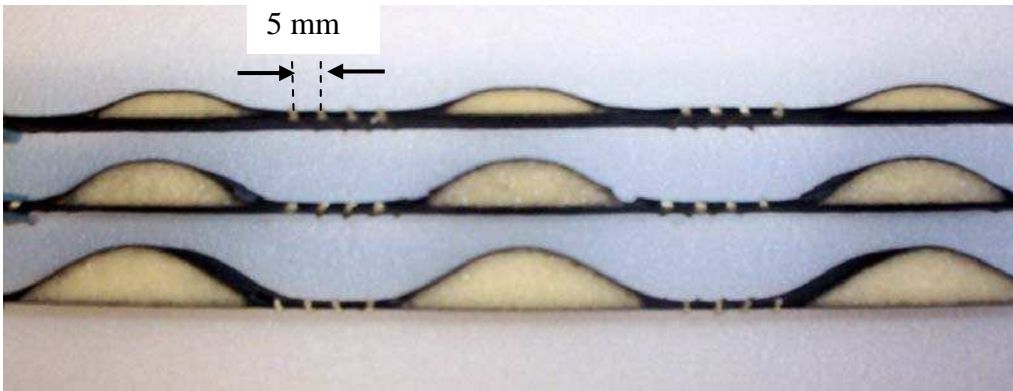
The specimen was infiltrated with Epon™ 862 epoxy and EpiKure™ 9553 hardener combined at a ratio of 100:17, through vacuum-assisted resin transfer molding, a low cost, out-of-autoclave process where the vacuum pressure differential forces the resin through the specimen (Figure 8). Curing is continued for two hours in an oven at 37° C.

Each element (composed of main link and waiting link) was 50 mm in length, and each chain-like specimen had three such elements. The material forming the node (section of the specimen with four stitched rows) was ignored for this study. Its effect will be considered for future work. Because of this, the baseline specimens had length equal to 150 mm. Tabs were applied to the specimen, which was cut into samples with width equal to 25 mm. These were then tested in tension in a 222 kN screw-driven machine at a rate of 1.5 mm/minute (Figure 9).

### 3. Testing and discussion of results

Tables 3 and 4 show test results for the baseline specimens and for all the iterations.

The specimens of the first thirteen iterations were tested with and without the core. The core was removed (scraped out from the side of the specimen) after the cure to study its effect on the specimen’s behavior. It was assumed that the core would be compressed by the waiting link’s straightening, and



**Figure 2.** Three of the geometries used in this study, applied to the carbon T300 specimens: from top to bottom, specimens with small round core, medium round core and large round core. The stitch rows are 5 mm apart in all specimens.

<i>Material</i>	<i>Core shape</i>	<i>Material</i>	<i>Core shape</i>	<i>Material</i>	<i>Core shape</i>
1 T300 carbon	trapezoid	8 T300 carbon	round gumdrop	15 Kevlar	medium round
2 T300 carbon	small round	9 E-glass	trapezoid	16 Kevlar	large round
3 T300 carbon	large round	10 E-glass	small round	17 Dyneema	small round
4 T300 carbon	large round	11 E-glass	large round	18 Dyneema	medium round
5 T300 carbon	small round	12 E-glass	medium round	19 Dyneema	large round
6 T300 carbon	small round	13 T300 carbon	medium round	20 Dyneema	very large round
7 T300 carbon	large round	14 Kevlar	small round	21 Dyneema	super-large round

**Table 1.** Iterations of materials and configurations for this work. Only iterations 1–5 did not have the core wrapped by Teflon-coated fiberglass. Only iterations 1–3 had two plies in the main links and four plies in the waiting links. All other iterations have one ply for the main links, two plies for the waiting links.

<i>Material</i>	<i>Weight, g/m<sup>2</sup> (oz/yd<sup>2</sup>)</i>	<i>Thickness, mm (in)</i>
12 × 12 T300 carbon plain weave	271 (8)	0.23 (0.009)
E-glass 7715 unidirectional cloth	246 (7.25)	0.23 (0.009)
17 × 17 satin Kevlar 49 weave	170 (5.0)	0.25 (0.010)
10 × 10 satin Dyneema weave	264 (7.8)	0.48 (0.019)

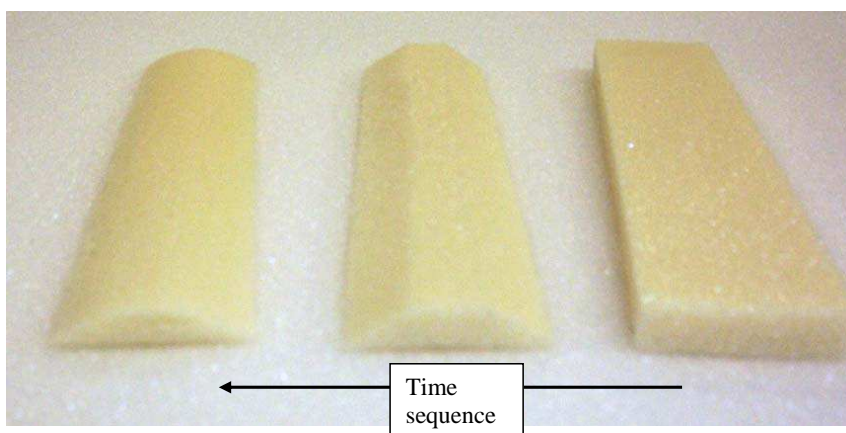
**Table 2.** Material information.

that it would push the main link out of plane, thus causing premature failure. Removing the core gave mixed results: in some cases, the specimen was damaged, in others the energy absorption grew with core removal; sometimes, the load/displacement curve was bistable, but in other cases, this did not happen. This inconsistency made us decide to stop the core removal for the last eight iterations.

Among the specimens, the percent wait varied from as little as 3.9% (iteration 2) to as high as 34% (iteration 21). The number of specimens per configuration varied between two (one with core, one without) to three (with core), depending on the success of the configuration. While such numbers of specimens may not give statistically significant results, we feel that this paper addresses a wide variety of different issues, and the paper gives qualitative results that may be indicative of the general performance of the structures.

Table 5 illustrates the energies of the round core configurations (small round, medium round and large round for the four materials, plus additional iterations for the Dyneema). These are given in terms of the baselines. Dyneema was the only material that exhibited consistent higher energy absorption with respect to its baseline, as will be discussed later in this paper. The E-glass and the Dyneema were the only materials where the damaged state was stronger than the undamaged state, as indicated by bold numbers in Table 5.

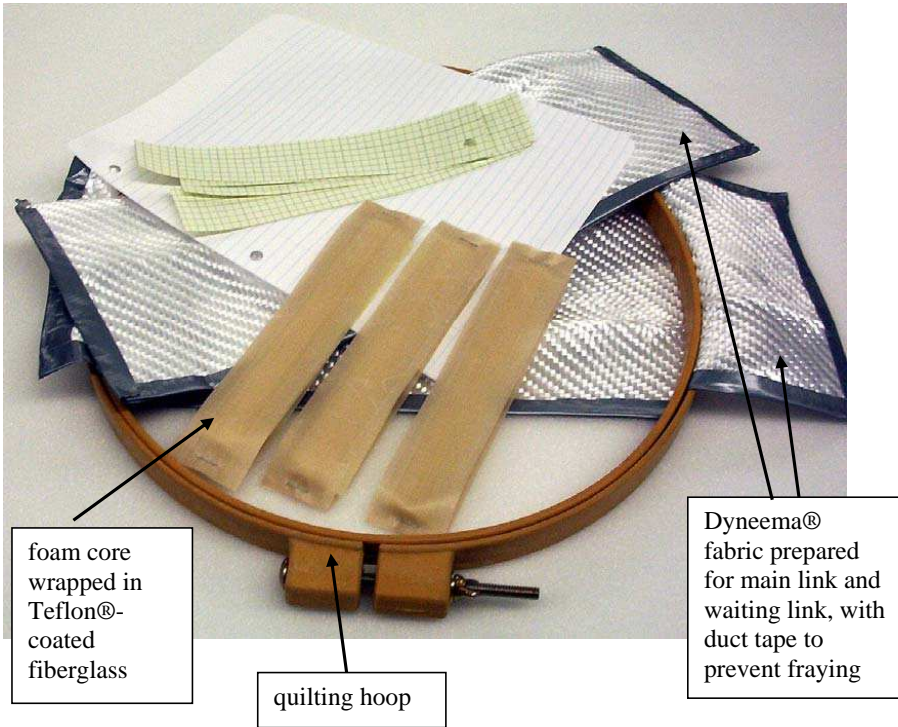
In the T300, Kevlar and E-glass cases, it is easy to extrapolate that the energy associated to the bistable specimens, calculated as the area under the curve, is significantly less than for the baselines (the numbers



**Figure 3.** Steps in the process of shaping the foam core for the specimens. From right to left: original core; intermediate stage, where the core was cut with a sharp box knife; final stage, where the core has been sanded to the shape on the left. The process was quite reliable and repeatable.

Reinforcement material	T300 carbon	E-glass	Kevlar	Dyneema
Energy of baseline (J)	$14.9 \pm 1.0$	$39.1 \pm 0.8$	$20.2 \pm 0.5$	$55.5 \pm 7.2$

**Table 3.** Energy of baseline specimens for each reinforcement used. Three specimens were tested for each material. Data shown is mean  $\pm$  one standard deviation.



**Figure 4.** Materials needed to stitch a specimen.



**Figure 5.** Sewing machine used to stitch, with Kevlar 29 thread, the plies of Dyneema inserted in the quilting hoop, with the aid of engineering paper. This process has also been used for the other materials discussed in the paper.





**Figure 6.** Left: picture of the top layer being folded back to allow the insertion of the wrapped core. Right: stitched Dyneema ready for core insertion.



**Figure 7.** Dyneema specimen ready for resin infiltration.

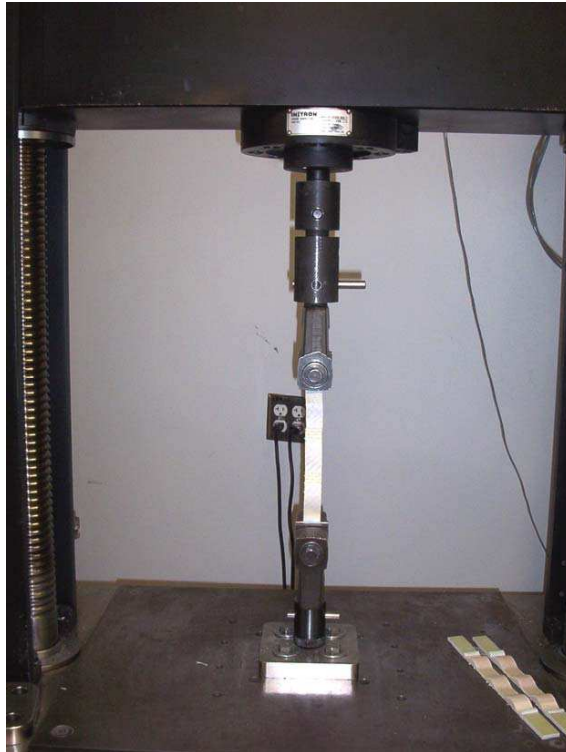
are shown in Table 5). The load drop after the main link failure was found to be proportional to the percent wait, so the percent wait was increased. This could lead to higher energy absorption (Figure 10).

Figure 11 shows the load/displacement curve of the carbon specimens with small round core, with and without core, compared to the three baseline specimens. One detects a bistable response of the specimens, but the damaged state is still weaker than the undamaged state; therefore this goal was not met.

However, several additional factors affect the structure's performance, such as the materials' strain and strength, matrix inclusions in the transition area between main link and waiting link, (which led to premature failure), manufacturing-related issues such as the waiting link's alignment, etc. Hence, while the general idea in Figure 10 seems to make sense, it cannot be said a priori that the energy absorption capacity of a bistable structure is proportional to the percent wait for all cases.



**Figure 8.** Specimen during vacuum-assisted resin transfer molding process (VARTM), after vacuum has been applied and before the specimen is cured in an oven. The red nylon mesh accelerates the infiltration process.



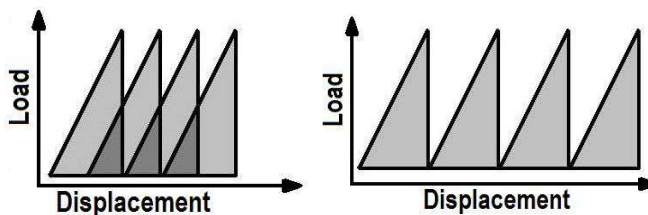
**Figure 9.** Test of a Dyneema specimen in a 222 kN screw-driven machine. Additional tabbed specimens are visible on the right.

Figure 12 shows side pictures of the fractured specimens in iterations 5–12, Figure 13 documents testing of one specimen and highlights a problem region of excess matrix.

**3.1. Discussion on T300 carbon, E-glass and Kevlar specimens.** The T300 carbon specimens proved unable to a) survive the strain needed to straighten the waiting link attached to the first failed main link, and b) carry the additional tensile load required to fail the next link in the chain. The average elongation to failure of the three-ply baseline specimens was equal to 1.87%, which is indicative of the strain capability

	<i>Material</i>	<i>Core shape</i>	<i>Percent wait (%)</i>	<i>Energy (J) with / without core</i>
1	T300 carbon	trapezoid	8.8	12.9 / 7.6
2	T300 carbon	small round	5.5	8.2 / 7.3
3	T300 carbon	large round	12	12.3 / 11.5
4	T300 carbon	large round	12	5.8 / 4.5
5	T300 carbon	small round	5.5	4.3 / 4.5
6	T300 carbon	small round	5.5	5.5 / 5.9
7	T300 carbon	large round	12	4.87 / 6.85
8	T300 carbon	round gumdrop	23.4	5.4 / 5.4
9	E-glass	trapezoid	8.8	10.3 / 11.5
10	E-glass	small round	5.5	16.6 / 15.4
11	E-glass	large round	12	11.1 / 4.06
12	E-glass	medium round	9	15.3 / 19.8
13	T300 carbon	medium round	9	5.54 / 6.85
14	Kevlar	small round	5.5	7.48 / –
15	Kevlar	medium round	9	7.76 / –
16	Kevlar	large round	12	8.28 / –
17	Dyneema	small round	5.5	83.9 ± 7.40 / –
18	Dyneema	medium round	9	84.2 ± 3.28 / –
19	Dyneema	large round	12	97.3 ± 12.7 / –
20	Dyneema	very large round	15	114 ± 5.05 / –
21	Dyneema	super-large round	34	133 ± 23.1 / –

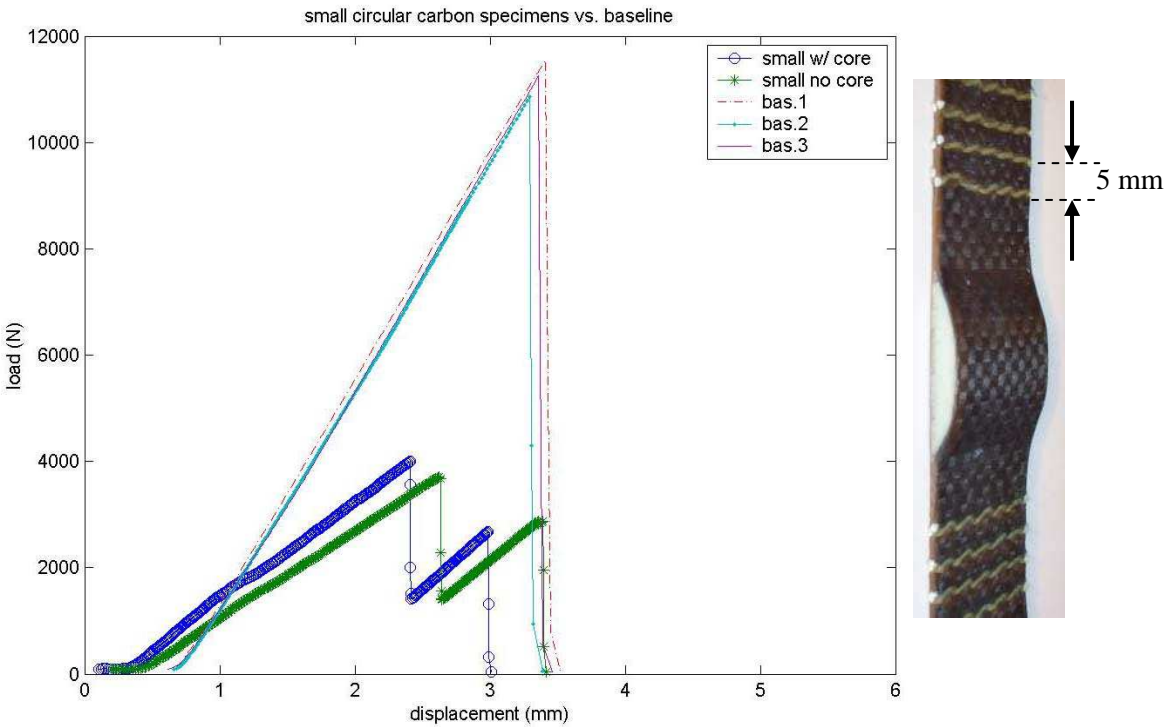
**Table 4.** Results of tensile tests for all configurations.



**Figure 10.** Response of small round core (left) with respect to medium and large core (right). The load drop after the main link breaks is proportional to the percent wait.

Segment Fiber	Energy (percent of baseline value)				
	Small round	Medium round	Large round	Very large round	Super-large round
T-300 carbon	37	37	33	N/A	N/A
<b>E-Glass</b>	<b>42</b>	<b>39</b>	<b>28</b>	N/A	N/A
Kevlar	37	38	41	N/A	N/A
<b>Dyneema</b>	<b>151</b>	<b>152</b>	<b>175</b>	<b>206</b>	<b>240</b>

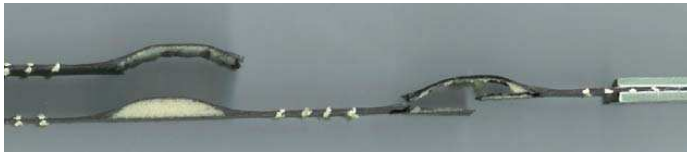
**Table 5.** Absorbed energies in terms of baselines, for the round core configurations of T-300 carbon, E-glass, Kevlar and Dyneema. E-glass and Dyneema were the only materials where the damaged state was stronger than the undamaged state (bold numbers). Dyneema was the only material with higher energy absorption than the baseline’s.



**Figure 11.** Load/displacement curves of carbon bistable specimens with small round core. One specimen had the core removed (‘no core’). The three baseline curves are shown, as well as a picture of the specimen.

of the chain-like specimens. The response was bistable in some of the specimens, however the damaged state was not stronger than the undamaged state (see Table 5 and Figures 11 and 14).

None of the T300 carbon specimens had the desired behavior. However, experimenting with geometries and use of Teflon-wrapped core was useful, and allowed to identify improved configurations and



Iteration 5, T300 carbon, small circular core, no wrapping of the core. The top specimen had the core removed prior to testing. The response of the two specimens was not bistable.



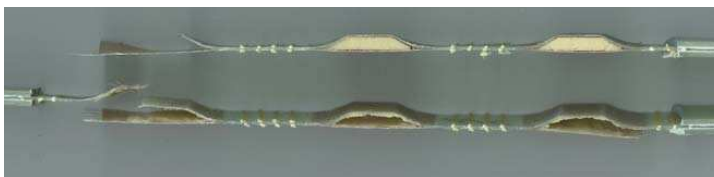
Iteration 6, T300 carbon, small circular core, with core wrapped.



Iteration 7, T300 carbon, large circular core



Iteration 8, T300 carbon, gumdrop core



Iteration 9, E-glass, trapezoid core. The core was removed before the test in the lower specimen.

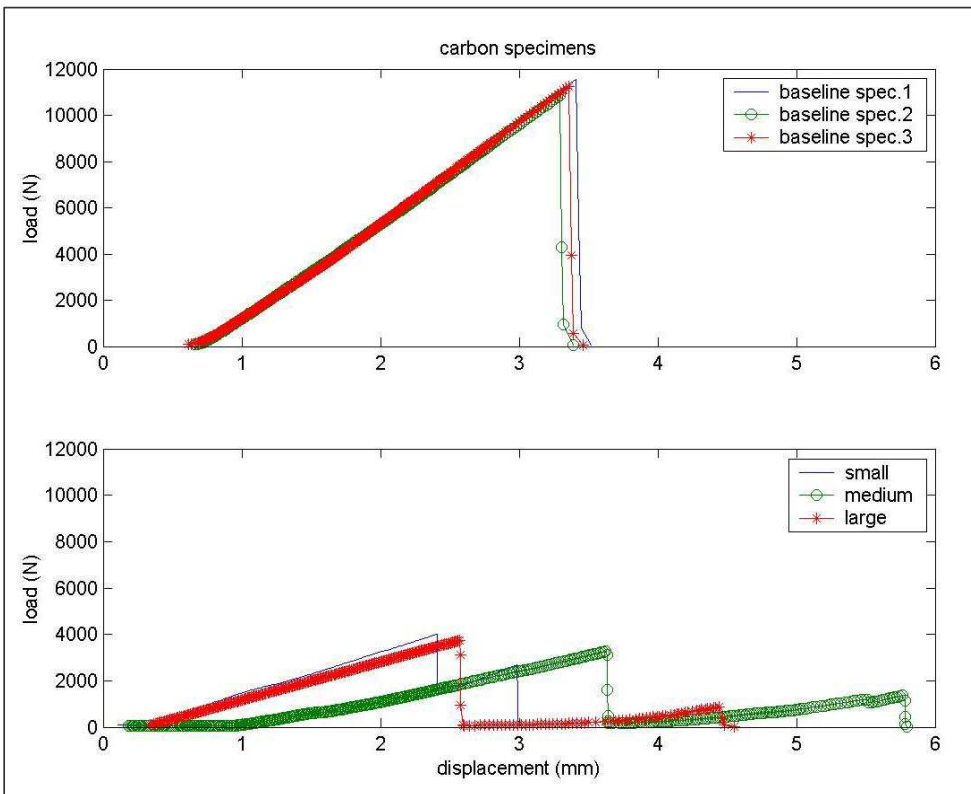


Iteration 12, E-glass, medium circular core

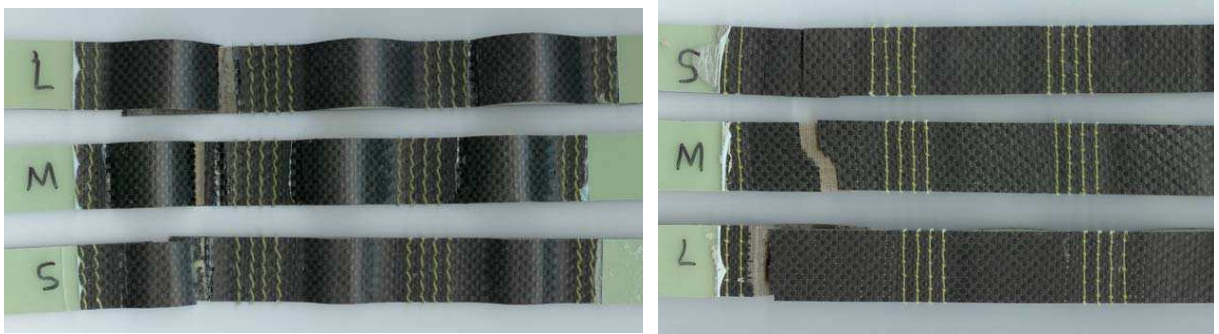
**Figure 12.** Side view of specimens shown after fracture. Among these specimens, the only iteration where the response not only was bistable but also has a stronger damaged state, is #12, the E-glass with medium round core. However, in all the cases in this figure, the absorbed energy was considerably less than the baselines’.



**Figure 13.** Picture of carbon specimen during testing. The matrix inclusion at the intersection of the main and waiting link popped out, causing failure of the specimen.



**Figure 14.** Load/displacement curves for baseline and bistable T300 carbon specimens.



**Figure 15.** Top and bottom views of fractured main links and waiting links in T300 carbon specimens with small, medium and large round cores. Specimen width is 25.4 mm.

observe and mitigate issues with the curvature of the waiting link: the intersection of main and waiting links tended to be plagued by matrix inclusions which ‘popped out’ during the testing, accelerating failure (note Figure 13). Figure 15 shows bottom and top views of the T300 specimens. Note how the main links failed both near the stitch location and in the center of the link, and the waiting links seem to fail at the concave transition of the waiting link layers.

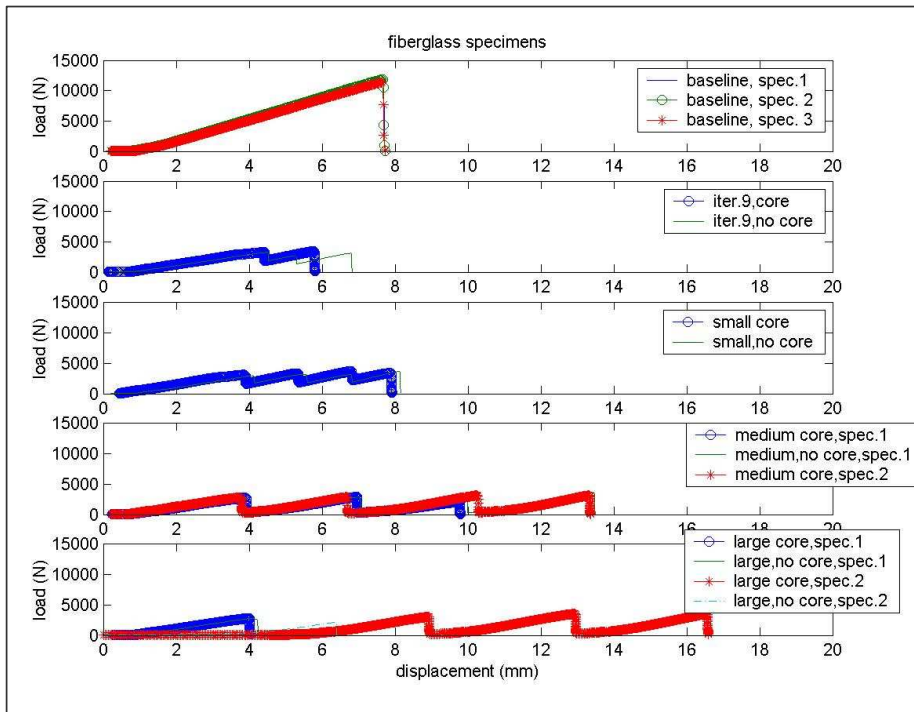
Three of the four iterations with E-glass specimens showed a bistable response with one of the desired behaviors (damaged state stronger than the undamaged state). The average elongation to failure of the baseline specimens was 3.62%. The much higher elongation to failure of this material with respect to carbon allowed a better performance in terms of bistable behavior, however the specimens did not have enough strength and durability to absorb more energy than their baselines (Figure 16, Table 5). Also, some E-glass specimens suffered from poor alignment of the waiting link material. Figure 17 shows fracture of main and waiting links in E-glass specimens. Similarly to the carbon specimens, fracture of the waiting link, occurs at the intersection of waiting link and main link, but the main link tended to fracture at the stitching.

The Kevlar specimens’ response was similar to the carbon’s. Figures 18 and 19 show the load/displacement curves and the fractured specimens. The average elongation to failure of the baseline was 2.54%.

The primary causes of failure for the T300 carbon, E-glass and Kevlar specimens include

- excessively acute radii in the waiting link,
- damage caused by core removal,
- susceptibility to matrix inclusions next to the core,
- excess matrix at the concave outer surface of the waiting link,
- high shock caused by energy release when the main link fails,
- high shock caused by a nonzero load following fracture,
- tearing due to poorly aligned layers (recall the manual manufacturing process).

**3.2. Discussion on Dyneema specimens.** Dyneema weave was received after the extensive testing of the previous three reinforcements, and it was the only material capable to absorb more energy than



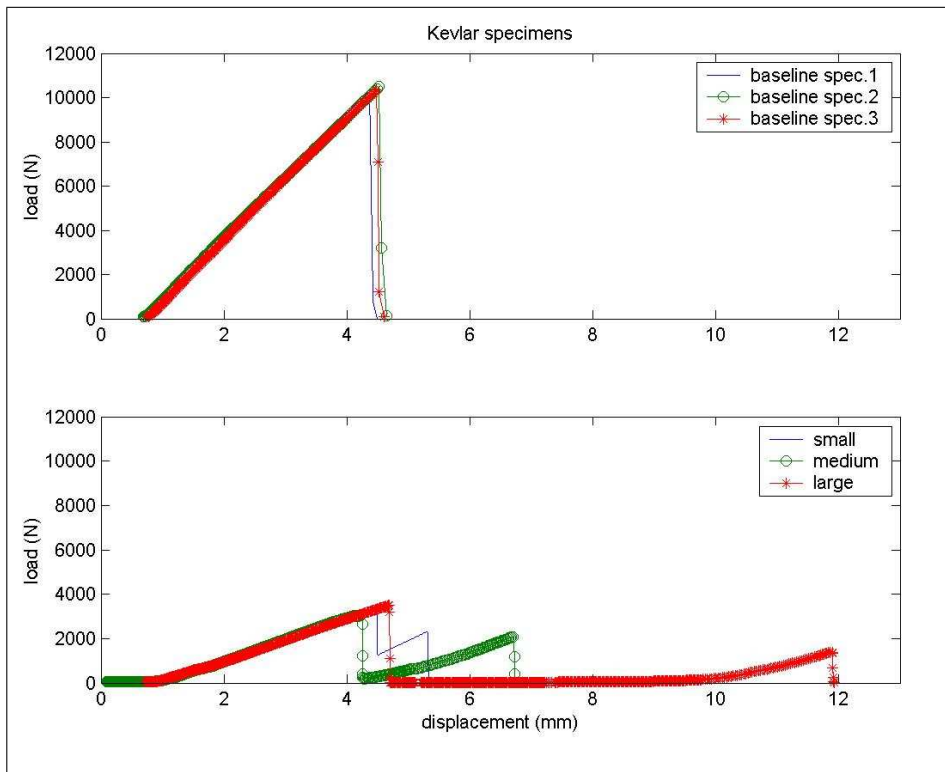
**Figure 16.** Load/displacement curves for baseline and bistable E-glass specimens.



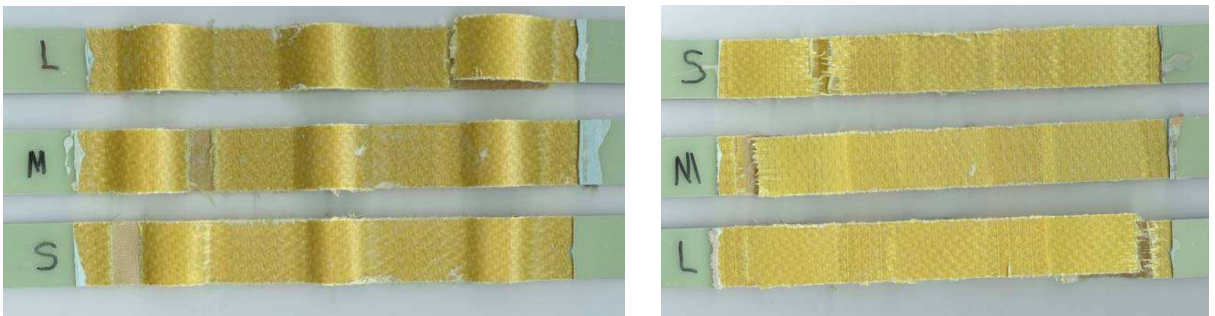
**Figure 17.** Top and bottom views of fractured main links and waiting links in E-glass specimens with small, medium and large round cores. Specimen width is 25.4 mm.

its baseline. The average strain of the baseline specimens was 4.7%. Note that the weave had about twice the thickness of the other weaves, but the weight is comparable to that of carbon and E-glass. In most Dyneema specimens, the whole structure was straight and in-plane until the first main link broke. Afterwards, the other main links moved out-of-plane and became somewhat curved. As the waiting link attached to the first fractured main link straightens, it crushes the core, while pushing the main link out of the way. When the last waiting link fails, the structure is still somewhat cohesive.





**Figure 18.** Load/displacement curves for baseline and bistable Kevlar specimens.



**Figure 19.** Top and bottom views of fractured main links and waiting links in Kevlar specimens with small, medium and large round cores. Specimen width is 25.4 mm.

The corresponding load/displacement curves show a nonlinear behavior (Figure 20). One of the reasons for the overall increase in absorbed energy is that both elongation to failure and strength are high enough that the waiting links are carrying considerable load by the time a main link fails. Figure 21 shows the lack of a complete fracture in main links and waiting links: different bundles seem to have had progressive failure. Further testing on the Dyneema involved specimens with larger percent waits

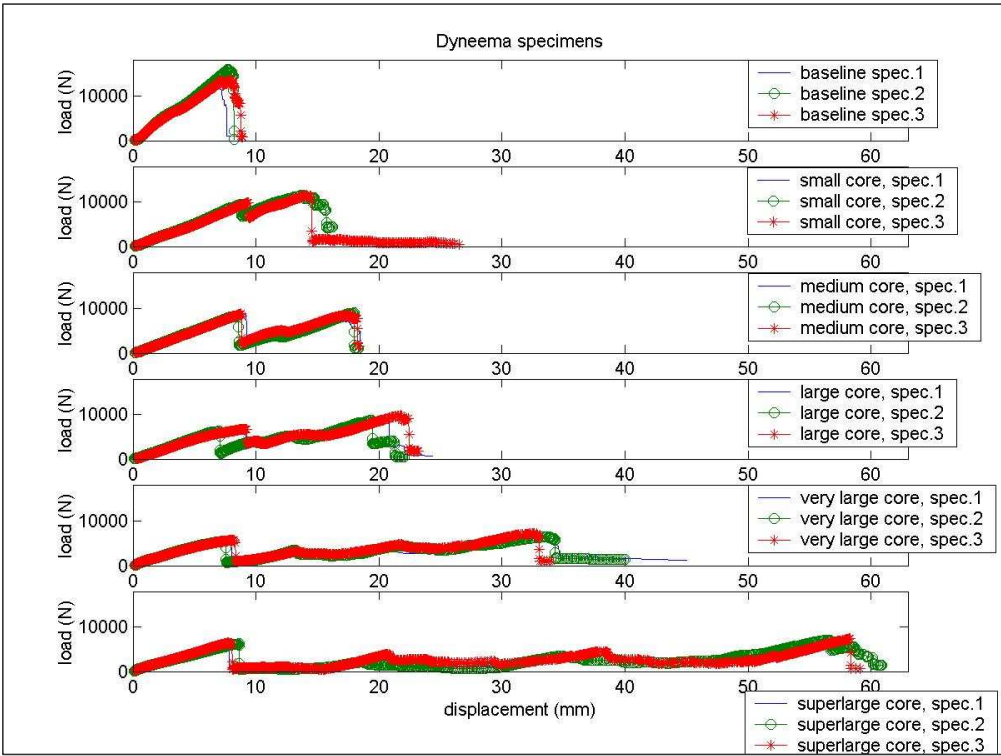


Figure 20. Load/displacement curves for baseline and bistable Dyneema specimens.

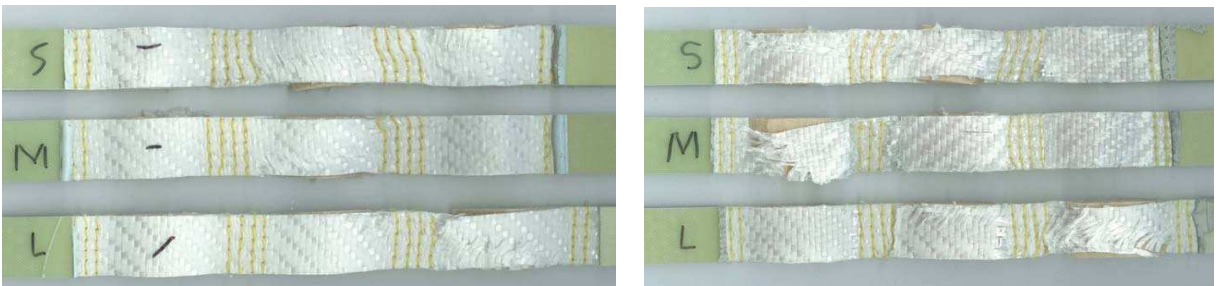


Figure 21. Top and bottom views of fractured main links and waiting links in Dyneema specimens with small, medium and large round cores. Specimen width is 25.4 mm.

(Table 5), and the absorbed energy consistently increased up to 240% of the baseline average energy. (For the nonlinear displacement/load curves of Dyneema, all data up to the point of the last significant load drop are used for the calculation of the energy. Because of that, the end ‘tails’ were not included in this calculation. The other materials’ load/displacement curves are linear and hence the calculation of the energy was straightforward.)

#### 4. Additional observations

Further experiments considered the effect of node composition (i.e. number of stitch rows), element length, and number of elements in a chain. These experiments were carried out on E-glass, the second best material, because the batch of Dyneema had been completely used. Current commercial demands make Dyneema very difficult to obtain for research purposes. However, we hope that it will be possible to complete this study in the future when supplies do come available.

#### 5. Summary and conclusions

This paper discussed the design, manufacturing and testing of composite bistable structures, with the goal of obtaining a) a bistable load/displacement curve which also exhibited a damaged state which higher load carrying capability than the undamaged state, b) higher energy absorption than the traditional laminate-type baselines. Twenty-one different configurations and four materials (T300 carbon, E-glass, Kevlar and Dyneema) were tested in quasistatic tension. Of the four materials, E-glass and Dyneema satisfied the first requirement, but only Dyneema was able to demonstrate a noticeable increase of energy absorption with respect to its baseline. Positive results could be obtained with other configurations of T300 carbon and Kevlar outside those designed in this paper.

This material-based mechanism is very affected by quality of manufacturing, local behavior (i.e. matrix inclusions), elongation to failure, strength, percent weight, node composition, element length. This first study on composite bistable structures under tension shows that they may offer a considerable increase of energy absorption and fail-safe fracture, which could affect the design of composites for all applications where crashworthiness and redundancy are critical aspects of the design.

#### Acknowledgments

We thank Dr. Andrej Cherkaev and Dr. Elena Cherkaev, Department of Mathematics, University of Utah, for their collaboration, and Dr. Martin Van Es, DSM Corporation, for the donation of Dyneema®.

The work of this paper was performed while Whitman was an M.S. student in the Department of Mechanical Engineering, University of Utah. The experimental work was conducted in the Strength of Materials Lab and in the Composite Manufacturing Lab of that department.

#### References

- [Cherkaev and Slepyan 1995] A. Cherkaev and L. Slepyan, "Waiting element structures and stability under extension", *Int. J. Damage Mec.* 4:1 (1995), 58–82.
- [Whitman and La Saponara 2007] Z. Whitman and V. La Saponara, "Bistable structures for energy absorption, I: metallic structures under tension", *J. Mech. Mater. Struct.* 2:2 (2007), 347–358.

Received 30 May 2006. Accepted 26 Sep 2006.

ZACHARY WHITMAN: zachary.whitman@swri.org  
Southwest Research Institute

VALERIA LA SAPONARA: vlasaponara@ucdavis.edu  
Department of Mechanical and Aeronautical Engineering, One Shields Ave, University of California, Davis, CA 95616-5294,  
United States

<http://mae.ucdavis.edu/vlasaponara/>

## PIEZOELECTRIC SHELL THEORIES WITH A *PRIORI* CONTINUOUS TRANSVERSE ELECTROMECHANICAL VARIABLES

ERASMO CARRERA AND SALVATORE BRISCHETTO

This paper addresses the static analysis of multilayer shells with embedded piezoelectric materials. The Reissner Mixed Variational Theorem is used to obtain transverse electromechanical variables (transverse shear and normal stresses, plus normal electrical displacement) which are *a priori* continuous at each layer-interface. The governing differential equations of doubly curved shells are derived by referring to the Unified Formulation in terms of a few fundamental nuclei. Formulation with discord interface continuity of transverse stresses and/or electrical displacements are discussed for comparison purpose. We address both equivalent single-layer models and layerwise models; up to fourth-order expansions in the thickness coordinate have been implemented. Numerical analysis has been restricted to closed-form solutions. Plates and simply supported cylindrical shells with orthotropic layers have been investigated. Both sensor and actuator configuration have been analyzed. The results obtained demonstrate the superiority of the proposed approach with respect to the other formulations considered, and its ability to furnish *a priori* interlaminar continuous transverse electrical displacement.

### 1. Introduction

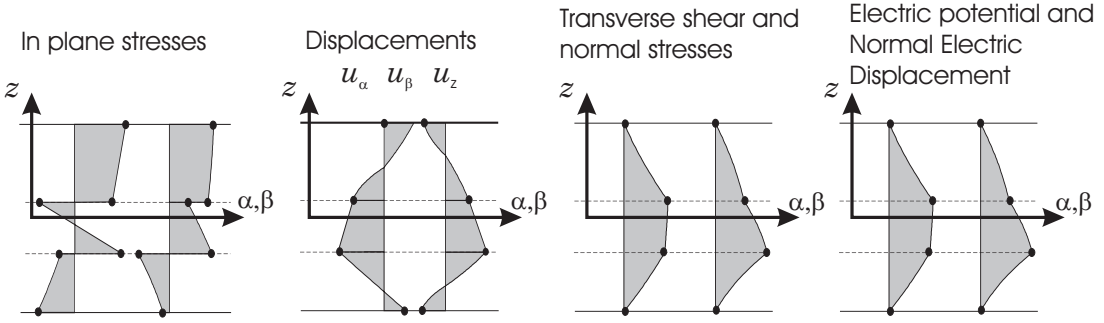
In recent years piezoelectric materials have been integrated with structural systems to build smart structures which are candidates for next generation of aerospace vehicles, as well as some advanced products in the automotive and ship industries. Piezoelectric materials are, in fact, capable of altering the response of the structures through sensing and actuation. By integrating surface bonded and embedded actuators in structural systems, the desired localized strains may be induced in the structures thanks to the application of an appropriate voltage to the actuators. Reviews on smart materials and structures appear in [Crawley 1994] and [Chopra 2002].

In most applications, the piezoelectric layers are embedded in multilayer structures made of anisotropic composite materials. The efficient use of piezoelectric materials in multilayer structures requires accurate evaluation of mechanical and electric variables in each layer. Classical shell models such as classical lamination theory (CLT) and first-order shear deformation theory (FSDT) can lead to large discrepancies with respect to the exact solution. Improvements can be introduced by using equivalent single-layer models (ESLMs), in which the number of unknown variables is independent of the number of constitutive layers, with higher-order kinematics. However, much better results can be obtained through the use of layerwise models (LWMs), in which the number of unknown variables depends on the number of layers. For recent indications of the superiority of LWMs over ESLMs, see [Robbins and Chopra 2006].

---

*Keywords:* piezoelectric shells, unified formulation, closed-form solutions, Reissner Mixed Variational Theorem, interlaminar continuity.

This work has been carried out in the framework of STREP EU project CASSEM under contract NMP-CT-2005-013517.



**Figure 1.**  $C^0_z$  requirements for the electromechanical case.

The advantages of the Reissner Mixed Variational Theorem (RMVT) with respect to other approaches that mostly make use of the Principle of Virtual Displacement (PVD) were shown in [Carrera 2001; 2003]. The Unified Formulation (UF) was used there to create an hierarchical shell formulation with variable kinematics (relative to displacements and transverse stresses) in each layer. Attention was restricted to pure mechanical problems. UF has been extended to closed-form and finite-element solutions of a piezoelectric plate in [Ballhause et al. 2004] and [Robaldo et al. 2006], respectively; PVD was used and only the displacements and the electrical potential were considered as unknown variables. The main advantage of RMVT is the possibility of fulfilling *a priori* the continuity conditions for the transverse electromechanical variables (electric displacement and stresses). In [Carrera 2001], these continuity conditions have been called  $C^0_z$ -requirements. Examples are given in Figure 1, which shows that some variables (mechanical displacements, transverse shear and normal stresses, electric voltage and transverse electrical displacements) must be  $C^0$  continuous in the thickness directions  $z$ , while the discontinuity of electromechanical properties at the layer interface requires a discontinuous first derivative of the same variables.

Attempts to introduce the  $C^0_z$ -requirements in piezoelectric continua have been made in [D’Ottavio and Kröplin 2006; Carrera and Boscolo 2006]. Closed form and FEs solutions were considered in these last papers, respectively. Attention was restricted to the fulfillment of  $C^0_z$ -requirements for transverse shear and normal stress components. Such an extension is herein stated as a “partial” RMVT application. The complete fulfillment of the  $C^0_z$ -requirements to both electrical and mechanical variables has been provided in the companion paper [Carrera and Fagiano 2007], devoted to FE analysis and plate geometries.. Such a contribution has been called a “full” extension of RMVT to piezoelectric continua: it allows one to describe *a priori* interlaminar continuous both transverse stresses and transverse electrical displacement component.

A few papers on piezoelectric shells exist in the literature, in particular for the FE method. Layerwise methods were considered in [Heyliger et al. 1996]. FE piezoelectric shells have been considered in [Lammering and Mesecke-Rischmann 2003]. Cho and Roh [2003] proposed geometrically exact shell elements, while Kögl and Bucalem [2005] gave the extension of MITC4 type element to piezoelectric shell structures. Review and assessment have been given in [Saravanos and Heyliger 1999].

Three dimensional piezoelectricity solutions have been addressed in [Wang et al. 2005; Shakeri et al. 2006; Chen et al. 1996; Dumir et al. 1997]. Wang et al. [2005] and Shakeri et al. [2006] dealt with

vibration problems. Chen et al. [1996] addressed cylindrical shell with very thin piezoelectric layers. Only a piezoelectric layer was instead considered in [Dumir et al. 1997]. No results are available in which both mechanical and piezoelectric layers (with thickness comparable to the mechanical layers) are analyzed.

It appears to be of interest to extend the full version of RMVT to piezoelectric shells which is proposed in this work. The present paper could be, in fact, considered as an extension of [Carrera 1999a; 1999b; D'Ottavio et al. 2006; Carrera and Brischetto 2007].

New benchmarks are proposed in this paper related to piezoelectric cylindrical shell that was originally proposed in [Varadan and Bhaskar 1991] for pure mechanical problems. Both actuator and sensor configurations are addressed. The role played by *a priori* continuous  $\mathcal{D}_z$  description has been outlined in most of the presented applications.

## 2. Overview of variational statements

The Principle of Virtual Displacement (PVD) for the pure mechanical static problems, can be written in the form

$$\int_V (\delta \boldsymbol{\epsilon}_{pG}^T \boldsymbol{\sigma}_{pC} + \delta \boldsymbol{\epsilon}_{nG}^T \boldsymbol{\sigma}_{nC}) dV = \delta L^e, \quad (1)$$

where  $\mathbf{u} = (u_\alpha, u_\beta, u_z)$  is the displacement vector;  $\boldsymbol{\sigma}_p = (\sigma_{\alpha\alpha}, \sigma_{\beta\beta}, \sigma_{\alpha\beta})$ ,  $\boldsymbol{\sigma}_n = (\sigma_{\alpha z}, \sigma_{\beta z}, \sigma_{zz})$ ,  $\boldsymbol{\epsilon}_p = (\epsilon_{\alpha\alpha}, \epsilon_{\beta\beta}, \epsilon_{\alpha\beta})$  and  $\boldsymbol{\epsilon}_n = (\epsilon_{\alpha z}, \epsilon_{\beta z}, \epsilon_{zz})$  are the in-plane and out-plane stresses and strains. The subscript  $C$  and  $G$  indicate the constitutive equations and the geometrical relations respectively (Sections 4–5).

The Reissner Mixed Variational Theorem assumes both the displacements  $\mathbf{u}$  and the normal stresses  $\boldsymbol{\sigma}_n$  according to the equation

$$\int_V (\delta \boldsymbol{\epsilon}_{pG}^T \boldsymbol{\sigma}_{pC} + \delta \boldsymbol{\epsilon}_{nG}^T \boldsymbol{\sigma}_{nM} + \delta \boldsymbol{\sigma}_{nM}^T (\boldsymbol{\epsilon}_{nG} - \boldsymbol{\epsilon}_{nC})) dV = \delta L^e. \quad (2)$$

Thus RMVT allows the *a priori* fulfillment of  $C_z^0$ -requirements of transverse shear and normal stresses. The subscript  $M$  indicates variables assumed from a given model (see Section 6).

The extension of RMVT to piezoelectric case requires the introduction of the internal electric work; see [D'Ottavio and Kröplin 2006; Carrera and Brischetto 2007; Carrera and Boscolo 2006; Garcia Lage et al. 2004]. The primary variables are displacements  $\mathbf{u}$ , normal stresses  $\boldsymbol{\sigma}_n$  and electric potential  $\Phi$ . This extension is referred to as partial P-RMVT:

$$\int_V (\delta \boldsymbol{\epsilon}_{pG}^T \boldsymbol{\sigma}_{pC} + \delta \boldsymbol{\epsilon}_{nG}^T \boldsymbol{\sigma}_{nM} + \delta \boldsymbol{\sigma}_{nM}^T (\boldsymbol{\epsilon}_{nG} - \boldsymbol{\epsilon}_{nC}) - \delta \mathcal{E}_G^T \mathcal{D}_C) dV = \delta L^e, \quad (3)$$

where  $\mathcal{E} = (\mathcal{E}_\alpha, \mathcal{E}_\beta, \mathcal{E}_z)$  is the electric field and  $\mathcal{D} = (\mathcal{D}_\alpha, \mathcal{D}_\beta, \mathcal{D}_z)$  is the electric displacement. Applications of P-RMVT to shell were discussed in [Carrera and Brischetto 2007].

The full extension of RMVT to piezomechanical problems (F-RMVT) must consider the transverse electrical displacement  $\mathcal{D}_z$  as an additional variable. F-RMVT states that

$$\int_V (\delta \boldsymbol{\epsilon}_{pG}^T \boldsymbol{\sigma}_{pC} + \delta \boldsymbol{\epsilon}_{nG}^T \boldsymbol{\sigma}_{nM} - \delta \mathcal{E}_{pG}^T \mathcal{D}_{pC} - \delta \mathcal{E}_{nG}^T \mathcal{D}_{nM} + \delta \boldsymbol{\sigma}_{nM}^T (\boldsymbol{\epsilon}_{nG} - \boldsymbol{\epsilon}_{nC}) - \delta \mathcal{D}_{nM}^T (\mathcal{E}_{nG} - \mathcal{E}_{nC})) dV = \delta L^e, \quad (4)$$

where  $\mathcal{D}_p = (\mathcal{D}_\alpha, \mathcal{D}_\beta)$  and  $\mathcal{D}_n = (\mathcal{D}_z)$  are in-plane and out-of-plane electric displacements, and  $\mathcal{E}_p = (\mathcal{E}_\alpha, \mathcal{E}_\beta)$  and  $\mathcal{E}_n = (\mathcal{E}_z)$  are in-plane and out-of-plane electric fields. F-RMVT allows the complete fulfillment of  $C_z^0$ -requirements.

Appropriate “constitutive” equations are required for (1)–(4), as discussed in Section 5.

### 3. Shell geometry

Geometry and notation for multilayer shells are shown in Figure 2; for details see [Rogacheva 1994]. The square of an infinitesimal linear segment in the layer, the associated infinitesimal area and volume elements are given by

$$\begin{aligned} ds_k^2 &= H_\alpha^k d\alpha_k^2 + H_\beta^k d\beta_k^2 + H_z^k dz_k^2, \\ d\Omega_k &= H_\alpha^k H_\beta^k d\alpha_k d\beta_k, \quad dV = H_\alpha^k H_\beta^k H_z^k d\alpha_k d\beta_k dz_k, \end{aligned} \tag{5}$$

where the metric coefficients are

$$H_\alpha^k = A^k (1 + z_k/R_\alpha^k), \quad H_\beta^k = B^k (1 + z_k/R_\beta^k), \quad H_z^k = 1, \tag{6}$$

$k$  is the layer index in the multilayer shell, and  $R_\alpha^k$  and  $R_\beta^k$  are the principal radii of curvature along the coordinates  $\alpha_k$  and  $\beta_k$ .  $A^k$  and  $B^k$  are the coefficients of the first fundamental form of  $\Omega_k$  ( $\Gamma_k$  is the  $\Omega_k$  boundary). In this paper, the attention has been restricted to shells with constant radii of curvature (cylindrical, spherical, toroidal geometries) for which  $A^k = B^k = 1$ .

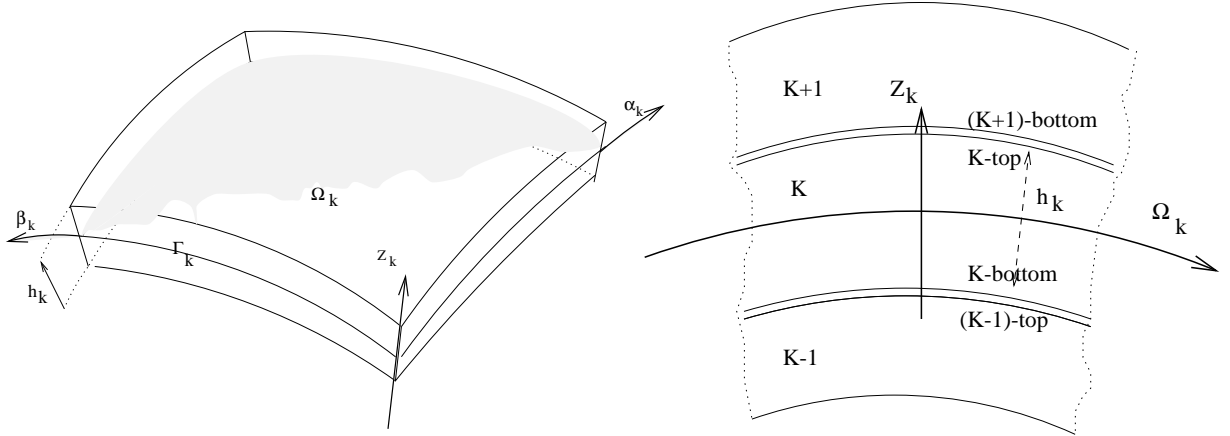


Figure 2. Geometry and notation for the layered shell.

### 4. Geometrical relations

The geometrical relations allow one to express the in-plane strain  $\epsilon_p$  and out-of-plane strain  $\epsilon_n$  in terms of displacement  $\mathbf{u}$ . At the same way the in-plane components  $\mathcal{E}_p$  and out-of-plane components  $\mathcal{E}_n$  of electric field can be expressed in terms of the electric potential  $\Phi$ . Then

$$\begin{aligned} \epsilon_p G &= [\epsilon_{\alpha\alpha}, \epsilon_{\beta\beta}, \epsilon_{\alpha\beta}]^T = (\mathbf{D}_p + \mathbf{A}_p) \mathbf{u}, & \epsilon_n G &= [\epsilon_{\alpha z}, \epsilon_{\beta z}, \epsilon_{zz}]^T = (\mathbf{D}_{n\Omega} + \mathbf{D}_{nz} - \mathbf{A}_n) \mathbf{u}, \\ \mathcal{E}_p G &= [\mathcal{E}_\alpha, \mathcal{E}_\beta]^T = -\mathbf{D}_{e\Omega} \Phi, & \mathcal{E}_n G &= [\mathcal{E}_z]^T = -\mathbf{D}_{en} \Phi. \end{aligned} \tag{7}$$



The explicit form of the introduced arrays follows:

$$\mathbf{D}_p = \begin{bmatrix} \partial_\alpha/H_\alpha & 0 & 0 \\ 0 & \partial_\beta/H_\beta & 0 \\ \partial_\beta/H_\beta & \partial_\alpha/H_\alpha & 0 \end{bmatrix}, \quad \mathbf{D}_{n\Omega} = \begin{bmatrix} 0 & 0 & \partial_\alpha/H_\alpha \\ 0 & 0 & \partial_\beta/H_\beta \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{D}_{nz} = \begin{bmatrix} \partial_z & 0 & 0 \\ 0 & \partial_z & 0 \\ 0 & 0 & \partial_z \end{bmatrix},$$

$$\mathbf{D}_{e\Omega} = \begin{bmatrix} \partial_\alpha/H_\alpha & 0 & 0 \\ 0 & \partial_\beta/H_\beta & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{D}_{en} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \partial_z \end{bmatrix},$$

$$\mathbf{A}_p = \begin{bmatrix} 0 & 0 & 1/(H_\alpha R_\alpha) \\ 0 & 0 & 1/(H_\beta R_\beta) \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{A}_n = \begin{bmatrix} 1/(H_\alpha R_\alpha) & 0 & 0 \\ 0 & 1/(H_\beta R_\beta) & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

For convenience, the electric potential, the in-plane and out-of-plane electric field have been treated as the  $[3 \times 1]$  vector with components  $\Phi = (\Phi, \Phi, \Phi)$ ,  $\mathcal{E}_{pG} = (\mathcal{E}_\alpha, \mathcal{E}_\beta, \mathcal{E}_\beta)$ ,  $\mathcal{E}_{nG} = (\mathcal{E}_z, \mathcal{E}_z, \mathcal{E}_z)$ . Such an artifice will allow us to preserve the  $[3 \times 3]$  dimension for the fundamental nuclei.

## 5. Constitutive equations

The constitutive equations of a piezoelectric continuum can be written in various forms; see [Ikeda 1996; Rogacheva 1994]. Different choices lead to the use of different electromechanical coefficients and field variables. The two most used forms are related to the Gibbs free energy  $G$  and the electric Gibbs energy  $G_2$ , respectively. These are often referred to as the  $d$ -form and the  $e$ -form, [Carrera and Brischetto 2007]. The formulation based on  $G_2$  is used in this work; we will present the constitutive equations for the PVD formulation and their split form in in-plane and out-of-plane components. We then use these equations to obtain the constitutive equations which are consistent with F-RMVT application.

**Classical PVD formulation.** The displacement  $\mathbf{u}$  and the potential  $\Phi$  are the unknown variables in the PVD setting. The electric Gibbs energy  $G_2$  is (see [Carrera and Brischetto 2007])

$$G_2(\boldsymbol{\epsilon}, \mathcal{E}) = \frac{1}{2} \boldsymbol{\epsilon}^T \mathbf{C}^\mathcal{E} \boldsymbol{\epsilon} - \frac{1}{2} \mathcal{E}^T \boldsymbol{\epsilon}^\epsilon \mathcal{E} - \mathcal{E}^T \boldsymbol{\epsilon} \boldsymbol{\epsilon}, \quad (8)$$

in which  $\mathbf{C}^\mathcal{E}$  is the elastic coefficients matrix related to constant electric field. The  $[6 \times 6]$  matrix  $\mathbf{C}$  of an orthotropic material in the rotated reference system assumes the form (see [Reddy 2004])

$$\mathbf{C}^\mathcal{E} = \begin{bmatrix} C_{11} & C_{12} & C_{16} & 0 & 0 & C_{13} \\ C_{12} & C_{22} & C_{26} & 0 & 0 & C_{23} \\ C_{16} & C_{26} & C_{66} & 0 & 0 & C_{63} \\ 0 & 0 & 0 & C_{55} & C_{45} & 0 \\ 0 & 0 & 0 & C_{45} & C_{44} & 0 \\ C_{31} & C_{32} & C_{36} & 0 & 0 & C_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{pp}^\mathcal{E} & \mathbf{C}_{pn}^\mathcal{E} \\ \mathbf{C}_{np}^\mathcal{E} & \mathbf{C}_{nn}^\mathcal{E} \end{bmatrix}, \quad (9)$$

$\mathbf{C}_{pp}^\mathcal{E}$ ,  $\mathbf{C}_{pn}^\mathcal{E}$ ,  $\mathbf{C}_{np}^\mathcal{E}$  and  $\mathbf{C}_{nn}^\mathcal{E}$  are  $[3 \times 3]$  submatrices related to in-plane  $p$  and out-of-plane (normal  $n$ ) strain/stress components.

The piezoelectric and dielectric coefficient matrices are

$$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & e_{15} & e_{14} & 0 \\ 0 & 0 & 0 & e_{25} & e_{24} & 0 \\ e_{31} & e_{32} & e_{36} & 0 & 0 & e_{33} \end{bmatrix}, \quad \boldsymbol{\varepsilon}^\epsilon = \begin{bmatrix} \varepsilon_{11}^\epsilon & \varepsilon_{12}^\epsilon & 0 \\ \varepsilon_{21}^\epsilon & \varepsilon_{22}^\epsilon & 0 \\ 0 & 0 & \varepsilon_{33}^\epsilon \end{bmatrix}, \quad (10)$$

where the dielectric coefficients refer to a constant strain state,  $\mathcal{E}$  is the electric field,  $\boldsymbol{\epsilon}$  is the  $[6 \times 1]$  strain vector with components  $\boldsymbol{\epsilon} = (\epsilon_{\alpha\alpha}, \epsilon_{\beta\beta}, \epsilon_{\alpha\beta}, \epsilon_{\alpha z}, \epsilon_{\beta z}, \epsilon_{zz})$ .

The stresses  $\boldsymbol{\sigma} = (\sigma_{\alpha\alpha}, \sigma_{\beta\beta}, \sigma_{\alpha\beta}, \sigma_{\alpha z}, \sigma_{\beta z}, \sigma_{zz})$  and electrical displacements  $\mathcal{D}$  are obtained upon direct differentiation of  $G_2$ :

$$\begin{aligned} \boldsymbol{\sigma} &= \frac{\partial G_2}{\partial \boldsymbol{\epsilon}} = \frac{\partial}{\partial \boldsymbol{\epsilon}} \left( \frac{1}{2} \boldsymbol{\epsilon}^T \mathbf{C}^\mathcal{E} \boldsymbol{\epsilon} - \frac{1}{2} \mathcal{E}^T \boldsymbol{\varepsilon}^{\epsilon\mathcal{E}} - \mathcal{E}^T \mathbf{e} \boldsymbol{\epsilon} \right) = \mathbf{C}^\mathcal{E} \boldsymbol{\epsilon} - \mathbf{e}^T \mathcal{E}, \\ \mathcal{D} &= -\frac{\partial G_2}{\partial \mathcal{E}} = -\frac{\partial}{\partial \mathcal{E}} \left( \frac{1}{2} \boldsymbol{\epsilon}^T \mathbf{C}^\mathcal{E} \boldsymbol{\epsilon} - \frac{1}{2} \mathcal{E}^T \boldsymbol{\varepsilon}^{\epsilon\mathcal{E}} - \mathcal{E}^T \mathbf{e} \boldsymbol{\epsilon} \right) = \mathbf{e} \boldsymbol{\epsilon} + \boldsymbol{\varepsilon}^{\epsilon\mathcal{E}}. \end{aligned} \quad (11)$$

The corresponding  $\mathbf{e}$ -form of the constitutive equations is split into in-plane and out-of-plane components:

$$\begin{aligned} \boldsymbol{\sigma}_p &= \frac{\partial G_2}{\partial \boldsymbol{\epsilon}_p} = \mathbf{C}_{pp}^\mathcal{E} \boldsymbol{\epsilon}_p + \mathbf{C}_{pn}^\mathcal{E} \boldsymbol{\epsilon}_n - \mathbf{e}_{pp}^T \mathcal{E}_p - \mathbf{e}_{np}^T \mathcal{E}_n, \\ \boldsymbol{\sigma}_n &= \frac{\partial G_2}{\partial \boldsymbol{\epsilon}_n} = \mathbf{C}_{np}^\mathcal{E} \boldsymbol{\epsilon}_p + \mathbf{C}_{nn}^\mathcal{E} \boldsymbol{\epsilon}_n - \mathbf{e}_{pn}^T \mathcal{E}_p - \mathbf{e}_{nn}^T \mathcal{E}_n, \\ \mathcal{D}_p &= -\frac{\partial G_2}{\partial \mathcal{E}_p} = \mathbf{e}_{pp} \boldsymbol{\epsilon}_p + \mathbf{e}_{pn} \boldsymbol{\epsilon}_n + \boldsymbol{\varepsilon}_{pp}^\epsilon \mathcal{E}_p + \boldsymbol{\varepsilon}_{pn}^\epsilon \mathcal{E}_n, \\ \mathcal{D}_n &= -\frac{\partial G_2}{\partial \mathcal{E}_n} = \mathbf{e}_{np} \boldsymbol{\epsilon}_p + \mathbf{e}_{nn} \boldsymbol{\epsilon}_n + \boldsymbol{\varepsilon}_{np}^\epsilon \mathcal{E}_p + \boldsymbol{\varepsilon}_{nn}^\epsilon \mathcal{E}_n, \end{aligned} \quad (12)$$

where

$$\begin{aligned} \mathbf{e}_{pp} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & \mathbf{e}_{pn} &= \begin{bmatrix} e_{15} & e_{14} & 0 \\ e_{25} & e_{24} & 0 \end{bmatrix}, & \mathbf{e}_{np} &= [e_{31} \quad e_{32} \quad e_{36}], & \mathbf{e}_{nn} &= [0 \quad 0 \quad e_{33}], \\ \boldsymbol{\varepsilon}_{pp} &= \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{12} & \varepsilon_{22} \end{bmatrix}, & \boldsymbol{\varepsilon}_{pn} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, & \boldsymbol{\varepsilon}_{np} &= [0 \quad 0], & \boldsymbol{\varepsilon}_{nn} &= [\varepsilon_{33}]. \end{aligned}$$

**F-RMVT constitutive equations.** In the F-RMVT setting one assumes the displacement  $\mathbf{u}$ , electric potential  $\Phi$ , transverse stresses  $\boldsymbol{\sigma}_n$  and normal electric displacement  $\mathcal{D}_n$ . The correspondent constitutive equations are

$$\begin{aligned} \boldsymbol{\sigma}_p &= \mathbf{C}_{\sigma\mathcal{D}} \mathcal{D}_n + \mathbf{C}_{\sigma\epsilon} \boldsymbol{\epsilon}_p + \mathbf{C}_{\sigma\sigma} \boldsymbol{\sigma}_n + \mathbf{C}_{\sigma\mathcal{E}} \mathcal{E}_p, \\ \boldsymbol{\epsilon}_n &= \mathbf{C}_{\epsilon\mathcal{D}} \mathcal{D}_n + \mathbf{C}_{\epsilon\epsilon} \boldsymbol{\epsilon}_p + \mathbf{C}_{\epsilon\sigma} \boldsymbol{\sigma}_n + \mathbf{C}_{\epsilon\mathcal{E}} \mathcal{E}_p, \\ \mathcal{D}_p &= \mathbf{C}_{\mathcal{D}\mathcal{D}} \mathcal{D}_n + \mathbf{C}_{\mathcal{D}\epsilon} \boldsymbol{\epsilon}_p + \mathbf{C}_{\mathcal{D}\sigma} \boldsymbol{\sigma}_n + \mathbf{C}_{\mathcal{D}\mathcal{E}} \mathcal{E}_p, \\ \mathcal{E}_n &= \mathbf{C}_{\mathcal{E}\mathcal{D}} \mathcal{D}_n + \mathbf{C}_{\mathcal{E}\epsilon} \boldsymbol{\epsilon}_p + \mathbf{C}_{\mathcal{E}\sigma} \boldsymbol{\sigma}_n + \mathbf{C}_{\mathcal{E}\mathcal{E}} \mathcal{E}_p, \end{aligned} \quad (13)$$

in which in-plane strains and electrical field components as well as transverse stresses and transverse electrical displacement are used to express  $\boldsymbol{\sigma}_p$ ,  $\boldsymbol{\epsilon}_n$ ,  $\mathcal{D}_p$  and  $\mathcal{E}_n$  as requested by F-RMVT in Equation (4).

The explicit forms of the matrices in Equation (13) are

$$\begin{aligned}
\mathbf{C}_{\sigma\mathcal{D}} &= (\mathbf{C}_{pn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T - \mathbf{e}_{np}^T)(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}, \\
\mathbf{C}_{\sigma\epsilon} &= \mathbf{C}_{pp} - \mathbf{C}_{pn}\mathbf{C}_{nn}^{-1}\mathbf{C}_{np} - (\mathbf{C}_{pn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T - \mathbf{e}_{np}^T)(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}(\mathbf{e}_{np} - \mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{C}_{np}), \\
\mathbf{C}_{\sigma\sigma} &= \mathbf{C}_{pn}\mathbf{C}_{nn}^{-1} - (\mathbf{C}_{pn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T - \mathbf{e}_{np}^T)(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}, \\
\mathbf{C}_{\sigma\mathcal{E}} &= \mathbf{C}_{pn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{pn}^T - \mathbf{e}_{pp}^T - (\mathbf{C}_{pn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T - \mathbf{e}_{np}^T)(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{pn}^T + \boldsymbol{\varepsilon}_{np}), \\
\mathbf{C}_{\epsilon\mathcal{D}} &= \mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}, \\
\mathbf{C}_{\epsilon\epsilon} &= -\mathbf{C}_{nn}^{-1}\mathbf{C}_{np} - \mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}(\mathbf{e}_{np} - \mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{C}_{np}), \\
\mathbf{C}_{\epsilon\sigma} &= \mathbf{C}_{nn}^{-1} - \mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}, \\
\mathbf{C}_{\epsilon\mathcal{E}} &= \mathbf{C}_{nn}^{-1}\mathbf{e}_{pn}^T - \mathbf{C}_{nn}^{-1}\mathbf{e}_{pn}^T(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{pn}^T + \boldsymbol{\varepsilon}_{np}), \\
\mathbf{C}_{\mathcal{D}\mathcal{D}} &= \mathbf{e}_{pn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1} + \boldsymbol{\varepsilon}_{pn}(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}, \\
\mathbf{C}_{\mathcal{D}\epsilon} &= \mathbf{e}_{pp} - \mathbf{e}_{pn}\mathbf{C}_{nn}^{-1}\mathbf{C}_{np} - (\mathbf{e}_{pn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{pn})(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}(\mathbf{e}_{np} - \mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{C}_{np}), \\
\mathbf{C}_{\mathcal{D}\sigma} &= \mathbf{e}_{pn}\mathbf{C}_{nn}^{-1} - \mathbf{e}_{pn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1} - \boldsymbol{\varepsilon}_{pn}(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}, \\
\mathbf{C}_{\mathcal{D}\mathcal{E}} &= \mathbf{e}_{pn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{pn}^T - (\mathbf{e}_{pn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{pn})(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{pn}^T + \boldsymbol{\varepsilon}_{np}) + \boldsymbol{\varepsilon}_{pp}, \\
\mathbf{C}_{\mathcal{E}\mathcal{D}} &= (\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}, \\
\mathbf{C}_{\mathcal{E}\epsilon} &= -(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}(\mathbf{e}_{np} - \mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{C}_{np}), \\
\mathbf{C}_{\mathcal{E}\sigma} &= -(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}, \\
\mathbf{C}_{\mathcal{E}\mathcal{E}} &= -(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{nn}^T + \boldsymbol{\varepsilon}_{nn})^{-1}(\mathbf{e}_{nn}\mathbf{C}_{nn}^{-1}\mathbf{e}_{pn}^T + \boldsymbol{\varepsilon}_{np}).
\end{aligned} \tag{14}$$

## 6. Unified formulation for shell theories

The Unified Formulation is a technique that allows one to handle in a unified manner a large variety of plate/shell modelings. According to UF, the governing equations are written in term of a few *fundamental nuclei*, which do not formally depend on the expansion  $N$  used in the  $z$ -direction, or in the variable description (LW or ESL).

The unknown variables  $\mathbf{u}$ ,  $\boldsymbol{\sigma}_n$ ,  $\boldsymbol{\Phi}$  and  $\mathcal{D}_n$  are expressed in term of the layer thickness coordinate as

$$\begin{aligned}
(\mathbf{u}^k(x, y, z), \boldsymbol{\Phi}^k(x, y, z), \boldsymbol{\sigma}_n^k(x, y, z), \mathcal{D}_n^k(x, y, z)) &= F_b(z)(\mathbf{u}_b^k(x, y), \boldsymbol{\Phi}_b^k(x, y), \boldsymbol{\sigma}_{nb}^k(x, y), \mathcal{D}_{nb}^k(x, y)) \\
&+ F_r(z)(\mathbf{u}_r^k(x, y), \boldsymbol{\Phi}_r^k(x, y), \boldsymbol{\sigma}_{nr}^k(x, y), \mathcal{D}_{nr}^k(x, y)) + F_t(z)(\mathbf{u}_t^k(x, y), \boldsymbol{\Phi}_t^k(x, y), \boldsymbol{\sigma}_{nt}^k(x, y), \mathcal{D}_{nt}^k(x, y)) \tag{15}
\end{aligned}$$

The subscript  $t$  and  $b$  denote the linear part of the thickness expansion ( $t$  and  $b$  will be used to denote top- and bottom-layer variable values in layerwise cases), while subscript  $r$  refers to higher-order terms:  $r = 1, \dots, N-1$ . In compact form,

$$(\mathbf{u}^k(x, y, z), \boldsymbol{\Phi}^k(x, y, z), \boldsymbol{\sigma}_n^k(x, y, z), \mathcal{D}_n^k(x, y, z)) = F_\tau(z)(\mathbf{u}^k(x, y), \boldsymbol{\Phi}^k(x, y), \boldsymbol{\sigma}_n^k(x, y), \mathcal{D}_n^k(x, y))_\tau. \tag{16}$$

Here  $(\mathbf{u}^k(x, y), \boldsymbol{\Phi}^k(x, y), \boldsymbol{\sigma}_n^k(x, y), \mathcal{D}_n^k(x, y))_\tau$  are two-dimensional unknowns, the  $F_\tau(z)$  are the base functions of the expansion, and the summation convention over repeated indices has been adopted. The base functions could be, in general, different for each variable. Different choices for  $F_\tau(z)$  will lead

to different plate/shell theories. The choices made in our study are briefly discussed below; detailed descriptions can be found in the works cited.

*Layerwise models.* The thickness functions are given by combinations of Legendre polynomials  $P_j$  as

$$F_t = \frac{P_0(\zeta_k) + P_1(\zeta_k)}{2}, \quad F_b = \frac{P_0(\zeta_k) - P_1(\zeta_k)}{2}, \quad F_r = P_r(\zeta_k) - P_{r-2}(\zeta_k), \quad r = 2, 3, \dots, N, \quad (17)$$

for  $\zeta = z_k/2h_k$ , where  $z_k$  is the local layer thickness coordinate and  $h_k$  is the layer thickness, so  $-1 \leq \zeta_k \leq 1$ . As mentioned,  $t$  and  $b$  denote top and bottom; that is, the chosen functions have the properties

$$\zeta_k = \begin{cases} 1 & : F_t = 1, F_b = 0, F_r = 0, \\ -1 & : F_t = 0, F_b = 1, F_r = 0, \end{cases} \quad (18)$$

Thanks to these properties the interlaminar continuity of the assumed variables can be easily linked in the assembly procedure from layer-level matrices to multilayer-level matrices.

*Equivalent single-layer model.* In this case the layerwise expansion is preserved for the transverse stresses, electric potential and electric displacements, while a Taylor-type expansion is used for the displacement components:

$$\mathbf{u}(x, y, z) = \mathbf{u}_\tau(x, y) z^\tau, \quad \tau = 0, N.$$

The base functions related to displacements can be chosen as

$$F_b(z) = 1, \quad F_r(z) = z^r, \quad r = 1, \dots, N - 1, \quad F_t(z) = z^N.$$

A refinement of the ESL formulation can be reached by adding to the displacement assumption a function that imposes a zigzag form onto the displacement distribution. Following Murakami [1986], who introduced this idea, we take the displacement with imposed zigzag as

$$\mathbf{u} = \mathbf{u}_0 + (-1)^k \zeta_k \mathbf{u}_Z + z^r \mathbf{u}_r, \quad r = 1, \dots, N - 1. \quad (19)$$

The zigzag term, denoted by the subscript  $Z$ , changes sign for each layer  $k$ . These terms can be represented in unified form, by referring the subscript  $t$  to the zigzag function:  $\mathbf{u}_t = \mathbf{u}_Z$ . In this case, the thickness functions are defined as

$$F_b = 1, \quad F_t = (-1)^k \zeta_k, \quad F_r = z^r \quad \text{for } r = 1, \dots, N - 1. \quad (20)$$

These ESLM with zigzag functions are considered with linear, parabolic and cubic expansions.

### 7. Governing equations

Upon substitution of the constitutive equation (13), the geometrical relations (7) and Equation (16) in Equation (4), and after integration by parts (see [Carrera 1998; 1999a; 1999b]), we obtain the governing differential equations for the piezoelectric layers consistent with the assumptions made.

The electromechanical equilibrium and compatibility equations on the domain  $\Omega^k$  are

$$\begin{aligned} \delta \mathbf{u}_\tau^k &: \mathbf{K}_{uu}^{k\tau s} \mathbf{u}_s^k + \mathbf{K}_{u\sigma}^{k\tau s} (\boldsymbol{\sigma}_n^k)_s + \mathbf{K}_{ue}^{k\tau s} \boldsymbol{\Phi}_s^k + \mathbf{K}_{u\mathcal{D}}^{k\tau s} (\mathcal{D}_n^k)_s = \mathbf{p}_\tau^k, \\ \delta (\boldsymbol{\sigma}_n^k)_\tau &: \mathbf{K}_{\sigma u}^{k\tau s} \mathbf{u}_s^k + \mathbf{K}_{\sigma\sigma}^{k\tau s} (\boldsymbol{\sigma}_n^k)_s + \mathbf{K}_{\sigma e}^{k\tau s} \boldsymbol{\Phi}_s^k + \mathbf{K}_{\sigma\mathcal{D}}^{k\tau s} (\mathcal{D}_n^k)_s = 0, \\ \delta \boldsymbol{\Phi}_\tau^k &: \mathbf{K}_{eu}^{k\tau s} \mathbf{u}_s^k + \mathbf{K}_{e\sigma}^{k\tau s} (\boldsymbol{\sigma}_n^k)_s + \mathbf{K}_{ee}^{k\tau s} \boldsymbol{\Phi}_s^k + \mathbf{K}_{e\mathcal{D}}^{k\tau s} (\mathcal{D}_n^k)_s = \mathbf{p}_{e\tau}^k, \\ \delta (\mathcal{D}_n^k)_\tau &: \mathbf{K}_{\mathcal{D}u}^{k\tau s} \mathbf{u}_s^k + \mathbf{K}_{\mathcal{D}\sigma}^{k\tau s} (\boldsymbol{\sigma}_n^k)_s + \mathbf{K}_{\mathcal{D}e}^{k\tau s} \boldsymbol{\Phi}_s^k + \mathbf{K}_{\mathcal{D}\mathcal{D}}^{k\tau s} (\mathcal{D}_n^k)_s = 0. \end{aligned} \tag{21}$$

Two types of boundary conditions on  $\Gamma^k$  are obtained: those of the Dirichlet type,

$$\mathbf{u}_s^k = \bar{\mathbf{u}}_s^k, \quad \boldsymbol{\Phi}_s^k = \bar{\boldsymbol{\Phi}}_s^k, \tag{22}$$

and those of Neumann type,

$$\begin{aligned} \mathbf{\Pi}_{uu}^{k\tau s} \mathbf{u}_s^k + \mathbf{\Pi}_{u\sigma}^{k\tau s} (\boldsymbol{\sigma}_n^k)_s + \mathbf{\Pi}_{ue}^{k\tau s} \boldsymbol{\Phi}_s^k + \mathbf{\Pi}_{u\mathcal{D}}^{k\tau s} (\mathcal{D}_n^k)_s &= \mathbf{\Pi}_{uu}^{k\tau s} \bar{\mathbf{u}}_s^k + \mathbf{\Pi}_{u\sigma}^{k\tau s} (\bar{\boldsymbol{\sigma}}_n^k)_s + \mathbf{\Pi}_{ue}^{k\tau s} \bar{\boldsymbol{\Phi}}_s^k + \mathbf{\Pi}_{u\mathcal{D}}^{k\tau s} (\bar{\mathcal{D}}_n^k)_s, \\ \mathbf{\Pi}_{eu}^{k\tau s} \mathbf{u}_s^k + \mathbf{\Pi}_{e\sigma}^{k\tau s} (\boldsymbol{\sigma}_n^k)_s + \mathbf{\Pi}_{ee}^{k\tau s} \boldsymbol{\Phi}_s^k + \mathbf{\Pi}_{e\mathcal{D}}^{k\tau s} (\mathcal{D}_n^k)_s &= \mathbf{\Pi}_{eu}^{k\tau s} \bar{\mathbf{u}}_s^k + \mathbf{\Pi}_{e\sigma}^{k\tau s} (\bar{\boldsymbol{\sigma}}_n^k)_s + \mathbf{\Pi}_{ee}^{k\tau s} \bar{\boldsymbol{\Phi}}_s^k + \mathbf{\Pi}_{e\mathcal{D}}^{k\tau s} (\bar{\mathcal{D}}_n^k)_s. \end{aligned} \tag{23}$$

The *fundamental nuclei* on the domain  $\Omega^k$  have the form

$$\begin{aligned} \mathbf{K}_{uu}^{k\tau s} &= \int_{A_k} ([-\mathbf{D}_p^\tau + \mathbf{A}_p^\tau]^T \mathbf{C}_{\sigma\epsilon} [\mathbf{D}_p^s + \mathbf{A}_p^s]) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{K}_{u\sigma}^{k\tau s} &= \int_{A_k} ([\mathbf{D}_{nz}^\tau - \mathbf{A}_n^\tau - \mathbf{D}_{n\Omega}^\tau]^T + [-\mathbf{D}_p^\tau + \mathbf{A}_p^\tau]^T \mathbf{C}_{\sigma\sigma}) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{K}_{ue}^{k\tau s} &= \int_{A_k} (-[-\mathbf{D}_p^\tau + \mathbf{A}_p^\tau]^T \mathbf{C}_{\sigma\mathcal{E}} \mathbf{D}_{e\Omega}^s) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{K}_{u\mathcal{D}}^{k\tau s} &= \int_{A_k} ([-\mathbf{D}_p^\tau + \mathbf{A}_p^\tau]^T \mathbf{C}_{\sigma\mathcal{D}}) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{K}_{\sigma u}^{k\tau s} &= \int_{A_k} ((\mathbf{D}_{nz}^s + \mathbf{D}_{n\Omega}^s - \mathbf{A}_n^s) - \mathbf{C}_{\epsilon\epsilon} (\mathbf{D}_p^s + \mathbf{A}_p^s)) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{K}_{\sigma\sigma}^{k\tau s} &= \int_{A_k} -\mathbf{C}_{\epsilon\sigma} \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{K}_{\sigma e}^{k\tau s} &= \int_{A_k} \mathbf{C}_{\epsilon\mathcal{E}} \mathbf{D}_{e\Omega}^s \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, & \mathbf{K}_{\sigma\mathcal{D}}^{k\tau s} &= \int_{A_k} -\mathbf{C}_{\epsilon\mathcal{D}} \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{K}_{eu}^{k\tau s} &= \int_{A_k} -\mathbf{D}_{e\Omega}^\tau{}^T \mathbf{C}_{\mathcal{D}\epsilon} (\mathbf{D}_p^s + \mathbf{A}_p^s) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, & \mathbf{K}_{e\sigma}^{k\tau s} &= \int_{A_k} -\mathbf{D}_{e\Omega}^\tau{}^T \mathbf{C}_{\mathcal{D}\sigma} \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{K}_{ee}^{k\tau s} &= \int_{A_k} \mathbf{D}_{e\Omega}^\tau{}^T \mathbf{C}_{\mathcal{D}\mathcal{E}} \mathbf{D}_{e\Omega}^s \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, & \mathbf{K}_{e\mathcal{D}}^{k\tau s} &= \int_{A_k} (\mathbf{D}_{en}^\tau{}^T - \mathbf{D}_{e\Omega}^\tau{}^T \mathbf{C}_{\mathcal{D}\mathcal{D}}) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{K}_{\mathcal{D}u}^{k\tau s} &= \int_{A_k} \mathbf{C}_{\mathcal{E}\epsilon} (\mathbf{D}_p^s + \mathbf{A}_p^s) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, & \mathbf{K}_{\mathcal{D}\sigma}^{k\tau s} &= \int_{A_k} \mathbf{C}_{\mathcal{E}\sigma} \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{K}_{\mathcal{D}e}^{k\tau s} &= \int_{A_k} (\mathbf{D}_{en}^s - \mathbf{D}_{e\Omega}^s \mathbf{C}_{\mathcal{E}\mathcal{E}}) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, & \mathbf{K}_{\mathcal{D}\mathcal{D}}^{k\tau s} &= \int_{A_k} \mathbf{C}_{\mathcal{E}\mathcal{D}} \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz. \end{aligned}$$

The *fundamental nuclei* on the boundary  $\Gamma^k$  are

$$\begin{aligned} \mathbf{\Pi}_{uu}^{k\tau s} &= \int_{A_k} \mathbf{I}_p^T \mathbf{C}_{\sigma\epsilon} (\mathbf{D}_p^s + \mathbf{A}_p^s) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, & \mathbf{\Pi}_{eu}^{k\tau s} &= \int_{A_k} \mathbf{I}_{e\Omega}^T \mathbf{C}_{\mathcal{D}\epsilon} (\mathbf{D}_p^s + \mathbf{A}_p^s) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{\Pi}_{u\sigma}^{k\tau s} &= \int_{A_k} (\mathbf{I}_{n\Omega}^T + \mathbf{I}_p^T \mathbf{C}_{\sigma\sigma}) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, & \mathbf{\Pi}_{e\sigma}^{k\tau s} &= \int_{A_k} \mathbf{I}_{e\Omega}^T \mathbf{C}_{\mathcal{D}\sigma} \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{\Pi}_{ue}^{k\tau s} &= \int_{A_k} -(\mathbf{I}_p^T \mathbf{C}_{\sigma\mathcal{E}} \mathbf{D}_{e\Omega}^s) \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, & \mathbf{\Pi}_{ee}^{k\tau s} &= \int_{A_k} -\mathbf{I}_{e\Omega}^T \mathbf{C}_{\mathcal{D}\mathcal{E}} \mathbf{D}_{e\Omega}^s \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \\ \mathbf{\Pi}_{u\mathcal{D}}^{k\tau s} &= \int_{A_k} \mathbf{I}_p^T \mathbf{C}_{\sigma\mathcal{D}} \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, & \mathbf{\Pi}_{e\mathcal{D}}^{k\tau s} &= \int_{A_k} \mathbf{I}_{e\Omega}^T \mathbf{C}_{\mathcal{D}\mathcal{D}} \mathbf{F}_\tau \mathbf{F}_s H_\alpha H_\beta dz, \end{aligned}$$

where some auxiliary arrays have been introduced to facilitate integration by parts:

$$\mathbf{I}_p = \begin{bmatrix} 1/H_\alpha & 0 & 0 \\ 0 & 1/H_\beta & 0 \\ 1/H_\beta & 1/H_\alpha & 0 \end{bmatrix}, \quad \mathbf{I}_{n\Omega} = \begin{bmatrix} 0 & 0 & 1/H_\alpha \\ 0 & 0 & 1/H_\beta \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{I}_{e\Omega} = \begin{bmatrix} 1/H_\alpha & 0 & 0 \\ 0 & 1/H_\beta & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (24)$$

**Closed-form solutions.** Navier-type closed-form solutions are possible for the governing equations derived in the previous section if the materials are *transversely isotropic*, that is, if they fulfill the conditions

$$C_{pp16} = C_{pp26} = C_{pn63} = C_{pn36} = C_{nn45} = 0, \quad e_{25} = e_{14} = e_{36} = 0, \quad \epsilon_{12}^\epsilon = \epsilon_{21}^\epsilon = 0. \quad (25)$$

The following harmonic assumptions can be made for the field variables:

$$\begin{aligned} (u_{\alpha_\tau}^k, \sigma_{\alpha z_\tau}^k) &= \sum_{m,n} (U_{\alpha_\tau}^k, S_{\alpha z_\tau}^k) \cos \frac{m\pi\alpha_k}{a_k} \sin \frac{n\pi\beta_k}{b_k}, & k &= 1, N_l, \\ (u_{\beta_\tau}^k, \sigma_{\beta z_\tau}^k) &= \sum_{m,n} (U_{\beta_\tau}^k, S_{\beta z_\tau}^k) \sin \frac{m\pi\alpha_k}{a_k} \cos \frac{n\pi\beta_k}{b_k}, & \tau &= t, b, r, \\ (u_{z_\tau}^k, \sigma_{zz_\tau}^k, p_{z_\tau}^k, \Phi_\tau^k, \mathcal{D}_{n_\tau}^k) &= \sum_{m,n} (U_{z_\tau}^k, S_{zz_\tau}^k, P_{z_\tau}^k, \hat{\Phi}_\tau^k, \hat{\mathcal{D}}_{n_\tau}^k) \sin \frac{m\pi\alpha_k}{a_k} \sin \frac{n\pi\beta_k}{b_k}, & r &= 2, N, \end{aligned} \quad (26)$$

in which  $a_k$  and  $b_k$  are the shell lengths in the  $\alpha_k$  and  $\beta_k$  directions, and  $m$  and  $n$  are the correspondent wave numbers. These assumptions correspond to simply supported boundary conditions.

Upon substitution of Equation (26), the governing equations on  $\Omega^k$  assume the form of a linear system of algebraic equations in the domain while the boundary conditions are exactly fulfilled.

The final form of that system of algebraic equations at the multilayer shell level could be formally written as follows:

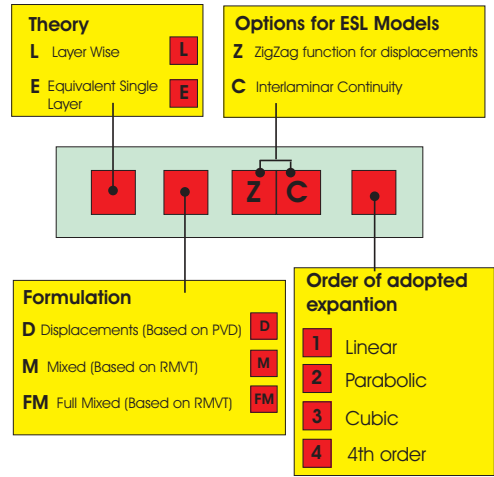
$$\begin{aligned} \hat{\mathbf{K}}_{uu} \hat{\mathbf{u}} + \hat{\mathbf{K}}_{u\sigma} \hat{\boldsymbol{\sigma}}_n + \hat{\mathbf{K}}_{ue} \hat{\boldsymbol{\Phi}} + \hat{\mathbf{K}}_{u\mathcal{D}} \hat{\mathcal{D}}_n &= \mathbf{p}, \\ \hat{\mathbf{K}}_{\sigma u} \hat{\mathbf{u}} + \hat{\mathbf{K}}_{\sigma\sigma} \hat{\boldsymbol{\sigma}}_n + \hat{\mathbf{K}}_{\sigma e} \hat{\boldsymbol{\Phi}} + \hat{\mathbf{K}}_{\sigma\mathcal{D}} \hat{\mathcal{D}}_n &= \mathbf{0}, \\ \hat{\mathbf{K}}_{eu} \hat{\mathbf{u}} + \hat{\mathbf{K}}_{e\sigma} \hat{\boldsymbol{\sigma}}_n + \hat{\mathbf{K}}_{ee} \hat{\boldsymbol{\Phi}} + \hat{\mathbf{K}}_{e\mathcal{D}} \hat{\mathcal{D}}_n &= \mathbf{p}_e, \\ \hat{\mathbf{K}}_{\mathcal{D}u} \hat{\mathbf{u}} + \hat{\mathbf{K}}_{\mathcal{D}\sigma} \hat{\boldsymbol{\sigma}}_n + \hat{\mathbf{K}}_{\mathcal{D}e} \hat{\boldsymbol{\Phi}} + \hat{\mathbf{K}}_{\mathcal{D}\mathcal{D}} \hat{\mathcal{D}}_n &= \mathbf{0}. \end{aligned} \quad (27)$$

An example of the explicit form of the fundamental nuclei for the  $k$ -layer is reported in the Appendix. The layer arrays are used to obtain multilayer matrices, using the assembly techniques described in [Carrera 2003].

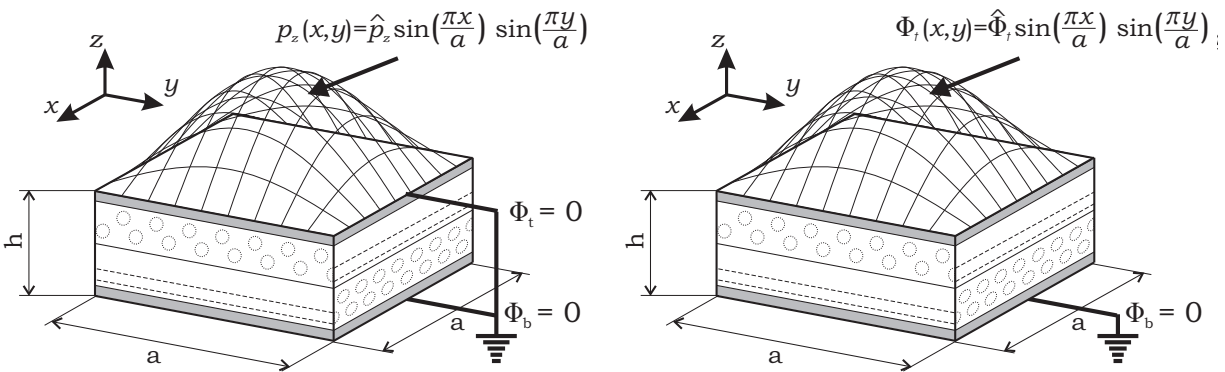
### 8. Numerical results

This section discusses numerical computations for piezoelectric shells, comparing F-RMVT results with the P-RMVT and PVD results described in [Carrera and Brischetto 2007].

**Acronyms of theories.** Depending on the used variational statement (PVD or RMVT), variables description and order of expansion  $N$ , a number of two-dimensional shell theories can be derived. To identify the various theories, appropriate acronyms are introduced, built as shown in the chart. The first field can be E or L, denoting the ESL or LW description, respectively; the second field can be D, M or FM according to PVD, RMVT application and RMVT application with normal electric displacement modeled; the last field can take on the numbers 1–4 according to the order of the adopted expansion in the thickness direction; a third Z and fourth C field (optional in the ESL case) denote the use of MZZF and/or IC fulfillment, respectively. FSDT results are obtained as particular case of ED1 ones, while CLT solutions have been computed by application of a penalty technique to a shear correction factor of FSDT. Stiffness coefficients related to FSDT and CLT are those of three-dimensional Hooke’s Law; that is, no plane stress problems have been imposed.



**Preliminary assessment.** The piezoelectric plates for which three-dimensional solutions were given in [Heyliger 1997] are considered. The two cases of a sensor (where a bisinusoidal distribution of mechanical pressure is applied) and an actuator (where a bisinusoidal distribution of electric voltage is applied) are considered in Figure 3, which also shows the Cartesian reference system used. The material properties are those in columns 1 and 2 of Table 1. A four-layer plate has been analyzed. The two external layers



**Figure 3.** Geometry and boundary conditions for the sensor (left) and actuator (right) configurations.

	PZT-4	Gr/EP	PVDF [Dumir et al. 1997]	Piezo [Chen et al. 1996]	Composite
$E_1$ [GPa]	81.3	132.38	2	2	172
$E_2$ [GPa]	81.3	10.756	2	2	6.9
$E_3$ [GPa]	64.5	10.756	2	2	6.9
$\nu_{12}$ [—]	0.329	0.24	1/3	0.29	0.25
$\nu_{13}$ [—]	0.432	0.24	1/3	0.29	0.25
$\nu_{23}$ [—]	0.432	0.49	1/3	0.29	0.25
$G_{23}$ [GPa]	25.6	3.606	0.75	0.7752	1.4
$G_{13}$ [GPa]	25.6	5.6537	0.75	0.7752	3.4
$G_{12}$ [GPa]	30.6	5.6537	0.75	0.7752	3.4
$e_{15}$ [C/m <sup>2</sup> ]	12.72	0	0	0	0
$e_{24}$ [C/m <sup>2</sup> ]	12.72	0	0	0	0
$e_{31}$ [C/m <sup>2</sup> ]	−5.20	0	−0.0015	0.046	0
$e_{32}$ [C/m <sup>2</sup> ]	−5.20	0	0.0285	0	0
$e_{33}$ [C/m <sup>2</sup> ]	15.08	0	−0.051	0	0
$\varepsilon_{11}/\varepsilon_0$ [—]	1475	3.5	—	—	—
$\varepsilon_{22}/\varepsilon_0$ [—]	1475	3.0	—	—	—
$\varepsilon_{33}/\varepsilon_0$ [—]	1300	3.0	—	—	—
$\varepsilon_{11}$ [pC/Vm]	$1.306 \times 10^4$	—	106.2	106.0	13060
$\varepsilon_{22}$ [pC/Vm]	$1.306 \times 10^4$	—	106.2	106.0	13060
$\varepsilon_{33}$ [pC/Vm]	$1.151 \times 10^4$	—	106.2	106.0	13060

**Table 1.** Electromechanical properties of the materials considered.

are made of piezoelectric material and have thickness  $h_e = 0.1h_{\text{tot}}$ . The two internal layers are of a unidirectional composite with fiber orientation  $0^\circ/90^\circ$  and thickness  $h_i = 0.4h_{\text{tot}}$ . Results for mechanical and electrical variables related to the sensor piezoelectric plates are given in Table 2, allowing further insight into the results already discussed in [Ballhause et al. 2004; D’Ottavio and Kröplin 2006; Carrera and Brischetto 2007]. The tables show the transverse distribution of the transverse normal stress and the normal electric displacement for a thick plate geometry. Higher-order LW results related to P-RMVT, F-RMVT and PVD applications are compared to exact three-dimensional solutions. Mixed theories fulfilling  $C_z^0$ -requirements (partial and full) give better normal stress results with respect to classical theories (LD4 and LD3 do not fulfill the continuity of  $\sigma_{zz}$  through the interfaces). The best results for  $\mathcal{D}_z$  are obtained via F-RMVT analysis under the assumption of *a priori* electrical and mechanical interlaminar continuity; indeed, LFM4 and LFM3 allow one to assume continuous  $\mathcal{D}_z$  through the interfaces.

The actuator case is considered in Table 3, which shows the electric potential obtained from high-order expansions compared with those from a 3D solution. PVD, P-RMVT and F-RMVT, are all suitable for



$z/h$	3D	LD4	LM4	LFM4	LD3	LM3	LFM3
0.500	1.0000	1.0000	1.0001	1.0001	1.0013	0.9994	0.9994
0.400	0.9515	0.9515	0.9516	0.9516	0.9502	0.9522	0.9522
0.400	0.9515	0.9518	0.9516	0.9516	0.9611	0.9522	0.9522
0.300	0.8520	0.8517	0.8517	0.8517	0.8479	0.8515	0.8515
0.200	0.7375	0.7376	0.7375	0.7375	0.7376	0.7376	0.7376
0.100	0.6169	0.6168	0.6169	0.6169	0.6207	0.6171	0.6171
0.000	0.4983	0.4986	0.4984	0.4984	0.4888	0.4977	0.4977
0.000	0.4983	0.4982	0.4984	0.4984	0.5067	0.4977	0.4977
-0.100	0.3805	0.3805	0.3804	0.3804	0.3771	0.3807	0.3807
-0.200	0.2614	0.2613	0.2614	0.2614	0.2613	0.2613	0.2613
-0.300	0.1482	0.1485	0.1484	0.1485	0.1518	0.1482	0.1482
-0.400	0.0487	0.0485	0.0487	0.0487	0.0403	0.0492	0.0492
-0.400	0.0487	0.0487	0.0487	0.0487	0.0499	0.0492	0.0492
-0.500	0.0000	0.0000	0.0000	0.0000	-0.0012	-0.0006	-0.0006
0.500	160.58	160.58	160.61	160.51	161.18	160.14	160.93
0.400	-0.3382	-0.3348	-0.4734	-0.3383	-0.9356	-4.1211	-0.3380
0.400	-0.3382	-0.3384	-0.3384	-0.3383	-0.3369	-0.3369	-0.3380
0.300	-0.1276	-0.1277	-0.1277	-0.1277	-0.1283	-0.1283	-0.1278
0.200	0.0813	0.0813	0.0813	0.0813	0.0813	0.0813	0.0813
0.100	0.2913	0.2914	0.2914	0.2914	0.2920	0.2920	0.2916
0.000	0.5052	0.5053	0.5053	0.5053	0.5038	0.5038	0.5050
0.000	0.5052	0.5053	0.5053	0.5053	0.5070	0.5070	0.5050
-0.100	0.7259	0.7262	0.7262	0.7262	0.7256	0.7256	0.7264
-0.200	0.9563	0.9565	0.9565	0.9565	0.9565	0.9565	0.9565
-0.300	1.1995	1.2000	1.2000	1.2000	1.2007	1.2007	1.1999
-0.400	1.4587	1.4590	1.4590	1.4590	1.4573	1.4573	1.4593
-0.400	1.4587	1.4559	1.3599	1.4590	1.9984	2.9753	1.4593
-0.500	-142.46	-142.46	-142.43	-142.39	-143.00	-142.69	-142.77

**Table 2.** Through the thickness distribution of (top) transverse normal stress  $\sigma_{zz}$  and (bottom) normal electric displacement  $\mathcal{D}_z$  in  $10^{13}$ , for the sensor case, with  $a/h = 4$ . Comparison of various approaches versus 3D solution.

this case; the electric potential is, in fact, a primary variable for all three models. The main advantage of F-RMVT with respect to P-RMVT is the fulfillment of  $C_z^0$ -requirements for normal electric displacement.

Further evaluation of the proposed model can be gleaned from the next few tables. A shell problem proposed in [Chen et al. 1996] is considered in Table 4, where 3D solutions are compared with respect to PVD, P-RMVT and F-RMVT models; mechanical displacements at middle surface are considered. This problem refers to a Ren shell [1987] with three mechanical cross-ply layers ( $90^\circ/0^\circ/90^\circ$ ) and two external piezoelectric layers (see fifth and fourth columns of Table 1). The thickness of each piezoelectric layer

$z/h$	3D	LD4	LM4	LFM4	LD3	LM3	LFM3
0.500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.400	0.9929	0.9929	0.9929	0.9929	0.9929	0.9929	0.9929
0.400	0.9929	0.9929	0.9929	0.9929	0.9929	0.9929	0.9929
0.300	0.8415	0.8418	0.8418	0.8418	0.8418	0.8418	0.8418
0.200	0.7014	0.7015	0.7015	0.7015	0.7014	0.7014	0.7014
0.100	0.5707	0.5709	0.5709	0.5709	0.5709	0.5709	0.5709
0.000	0.4476	0.4477	0.4477	0.4477	0.4477	0.4477	0.4476
0.000	0.4476	0.4477	0.4477	0.4477	0.4477	0.4477	0.4476
-0.100	0.3305	0.3307	0.3307	0.3307	0.3307	0.3307	0.3307
-0.200	0.2179	0.2179	0.2179	0.2179	0.2179	0.2179	0.2179
-0.300	0.1081	0.1082	0.1082	0.1082	0.1082	0.1082	0.1082
-0.400	-0.0010	-0.0010	-0.0010	-0.0010	-0.0010	-0.0010	-0.0010
-0.400	-0.0010	-0.0010	-0.0010	-0.0010	-0.0010	-0.0010	-0.0010
-0.500	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

**Table 3.** Through the thickness distribution of electric potential  $\Phi$  for the actuator case, with  $a/h = 4$ . Comparison of various approaches versus 3D solution.

$R_\beta/h_c$	2	4	10	50	100	2	4	10	50	100
	$\bar{u}_z (z = 0)$					$\bar{u}_\beta (z = 0)$				
Ref	1.440	0.459	0.144	0.0808	0.0785	5.294	1.549	0.480	0.269	0.262
LM4	1.443	0.458	0.144	0.0810	0.0787	5.308	1.547	0.479	0.270	0.262
LD4	1.443	0.458	0.144	0.0810	0.0787	5.305	1.547	0.479	0.270	0.262
LFM4	1.443	0.458	0.144	0.0810	0.0787	5.308	1.549	0.479	0.270	0.262
LM3	1.442	0.458	0.144	0.0810	0.0787	5.306	1.547	0.479	0.270	0.262
LD3	1.442	0.458	0.144	0.0810	0.0787	5.299	1.547	0.479	0.270	0.262
LFM3	1.442	0.458	0.144	0.0810	0.0787	5.306	1.547	0.479	0.270	0.262

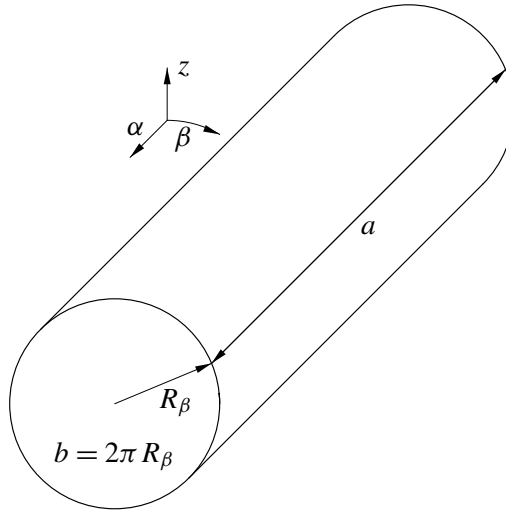
**Table 4.** Mechanical displacements for a five-layer piezomechanic Ren shell [Chen et al. 1996]; the numbers given are  $\bar{u}_z = 10 E_3 u_z h_c^3 / (P_z R_\beta^4)$  and  $\bar{u}_\beta = 100 E_3 u_\beta h_c^3 / P_z R_\beta^4$ . Comparison with exact 3D solution.

is one hundredth of the total ( $h_p = 1/100 h_c$ ). Loading conditions are:  $\Phi_t = \Phi_b = 0$  and  $P_z = 1$  at top. F-RMVT results are very similar to 3D solutions and to PVD and P-RMVT models: the piezoelectric layers are, in fact, very thin and mechanical displacements are the primary variables in the three proposed variational statements.

Table 5 shows a comparison of the results obtained with the present models with the three-dimensional solution from [Dumir et al. 1997] for a one-layer piezoelectric Ren shell, whose material properties are reported in column 3 of Table 1. The loading conditions are  $\Phi_t = \Phi_b = 0$  and  $P_z = 1$  at the top. F-RMVT

$R_\beta/h$	Ref	LM4	LM3	LFM4	LFM3	$R_\beta/h$	Ref	LM4	LM3	LFM4	LFM3
$\bar{w}(z = -h/2)$						$\bar{\sigma}_{\beta z}(z = 0)$					
<b>2</b>	-28.65	-28.63	-28.46	-28.63	-28.46	<b>2</b>	-0.6653	-0.6742	-0.6706	-0.6742	-0.6706
<b>4</b>	-20.55	-20.55	-20.54	-20.55	-20.54	<b>4</b>	-0.6238	-0.6261	-0.6252	-0.6261	-0.6252
<b>6</b>	← -18.72 →		<b>10</b>	← -17.60 →		<b>6</b>	-0.6055	-0.6065	-0.6061	-0.6065	-0.6061
<b>20</b>	-16.96	-16.95	-16.95	-16.95	-16.95	<b>10</b>	-0.5893	-0.5896	-0.5895	-0.5896	-0.5895
<b>100</b>	← -16.55 →		<b>500</b>	← -16.48 →		<b>20</b>	-0.5762	-0.5763	-0.5763	-0.5763	-0.5763
$\bar{w}(z = 0)$						$10^3\bar{\Phi}(z = 0)$					
<b>2</b>	-31.47	-31.45	-31.41	-31.45	-31.41	<b>2</b>	1.734	1.729	1.847	1.729	1.847
<b>4</b>	← -21.10 →		<b>6</b>	← -18.96 →		<b>4</b>	2.443	2.442	2.477	2.442	2.477
<b>10</b>	← -17.68 →		<b>20</b>	← -16.98 →		<b>6</b>	2.541	2.540	2.556	2.540	2.556
<b>100</b>	← -16.55 →		<b>500</b>	← -16.48 →		<b>10</b>	2.560	2.560	2.565	2.560	2.565
$\bar{w}(z = h/2)$						$10\bar{\mathcal{D}}_z(z = -h/2)$					
<b>2</b>	-31.31	-31.30	-31.14	-31.30	-31.14	<b>2</b>	5.908	-9.206	-10.27	6.381	7.326
<b>4</b>	-20.65	-20.65	-20.64	-20.65	-20.64	<b>4</b>	2.550	-16.26	-16.92	2.670	3.299
<b>6</b>			← -18.73 →			<b>6</b>	1.532	-31.74	-32.19	1.584	2.017
<b>10</b>	← -17.60 →		<b>20</b>	← -16.95 →		<b>10</b>	0.7830	-82.68	-82.94	0.8008	1.0622
<b>100</b>	← -16.55 →		<b>500</b>	← -16.48 →		<b>20</b>	0.2659	-320.7	-320.9	0.2702	0.4003
$\bar{v}(z = -h/2)$						$10\bar{\mathcal{D}}_z(z = h/2)$					
<b>2</b>	-23.61	-23.59	-23.38	-23.59	-23.38	<b>2</b>	0.8480	-5.641	-4.502	1.1484	-0.1072
<b>4</b>	-13.09	-13.09	-13.08	-13.09	-13.08	<b>4</b>	0.0213	-14.28	-13.62	0.1161	-0.5774
<b>6</b>	-10.23	-10.23	-10.22	-10.22	-10.22	<b>6</b>	-0.0996	-30.43	-29.98	-0.0557	-0.5155
<b>10</b>	← -8.177 →		<b>20</b>	← -6.778 →		<b>10</b>	-0.1575	-81.90	-81.64	-0.1413	-0.4121
<b>100</b>	← -5.738 →		<b>500</b>	← -5.539 →		<b>20</b>	-0.1860	-320.4	-320.3	-0.1820	-0.3143
$\bar{v}(z = h/2)$						$10\bar{\mathcal{D}}_\beta(z = 0)$					
<b>2</b>	2.046	2.045	2.029	2.045	2.029	<b>2</b>	-0.3070	-0.3060	-0.3270	-0.3060	-0.3270
<b>4</b>	-0.8806	-0.8806	-0.8799	-0.8806	-0.8799	<b>4</b>	-0.4324	-0.4323	-0.4384	-0.4323	-0.4384
<b>6</b>	-2.331	-2.331	-2.330	-2.331	-2.330	<b>6</b>	-0.4497	-0.4497	-0.4524	-0.4497	-0.4524
<b>10</b>			← -3.572 →			<b>10</b>	-0.4531	-0.4531	-0.4541	-0.4531	-0.4541
<b>20</b>	-4.528	-4.528	-4.527	-4.527	-4.527	<b>20</b>	-0.4496	-0.4496	-0.4498	-0.4496	-0.4498
<b>100</b>	← -5.297 →		<b>500</b>	← -5.451 →		<b>100</b>			← -0.4431 →		
$\bar{\sigma}_{zz}(z = 0)$						$10\bar{\mathcal{D}}_\beta(z = 0)$					
<b>2</b>	-0.2906	-0.2359	-0.2922	-0.2359	-0.2922	<b>2</b>	-0.4415	-0.4415	-0.4414	-0.4414	-0.4414
<b>4</b>	0.2170	0.2453	0.2181	0.2453	0.2181	<b>4</b>					
<b>6</b>	0.6356	0.6541	0.6365	0.6541	0.6365	<b>6</b>					
<b>10</b>	1.420	1.430	1.420	1.430	1.420	<b>10</b>					
<b>20</b>	3.319	3.325	3.320	3.325	3.320	<b>20</b>					
<b>100</b>	← 18.34 →		<b>500</b>	← 93.34 →		<b>100</b>					

**Table 5.** One-layer piezoelectric Ren shell: present analysis vs. exact solutions [Dumir et al. 1997]. Values of  $R\beta/h$  in bold;  $\leftarrow x \rightarrow$  indicates that all 5 row values equal  $x$ .



**Figure 4.** Proposed benchmark: cylindrical shell, geometry and notation from [Varadan and Bhaskar 1991].

and P-RMVT models are compared to exact solutions for different values of the thickness ratio. This table confirms the validity of F-RMVT model: the limitations for  $\mathcal{D}_z$  evaluation exhibited by P-RMVT model (see also [Carrera and Brischetto 2007]) have been overcome by F-RMVT.

**Proposed benchmarks.** In this section the shell problems for which 3D solution were given in [Varadan and Bhaskar 1991] for the pure mechanical problem are extended to the piezoelectric case. The problem was studied in [Carrera 1999a; 1999b] to assess the Unified Formulation in both PVD and RMVT for pure mechanical problems. The piezoelectric shells considered here are built by replacing and/or adding piezoelectric layers to the original shells. The material properties of these layers coincide with those already used for the piezoelectric plates.

The cylindrical shell considered by Bhaskar and Varadan has the geometric parameters  $a = 40$ ,  $b = 20\pi$ ,  $R_\alpha = \infty$ ,  $R_\beta = 10$ ,  $m = 1$ ,  $n = 8$ ; see Figure 4.

Two layouts are considered: one piezoelectric layer, and then four layers, the two external layers being made of piezoelectric material, and the two internal ones of carbon fiber cross-ply ( $0^\circ/90^\circ$ ). The mechanical and geometrical properties are those used in the plate case. Both actuator and sensor configurations are treated. For the actuator case the distribution of electric potential  $\Phi$  applied at top surface is

$$\Phi(\alpha, \beta) = \bar{\Phi} \sin \frac{\pi\alpha}{a} \sin \frac{8\pi\beta}{b}, \tag{28}$$

with  $\bar{\Phi}_t = 1$ ,  $\bar{\Phi}_b = 0$  and  $\bar{P}_z = 0$ .

Transverse mechanical pressure is applied at the bottom surface in the sensor case:

$$P_z(\alpha, \beta) = \bar{P}_z \sin \frac{\pi\alpha}{a} \sin \frac{8\pi\beta}{b}, \tag{29}$$

with  $\bar{\Phi}_{top} = \bar{\Phi}_b = 0$ , and  $\bar{P}_z = 1$ .

Results are given in the next few tables and figures, which compare different variational treatments for thick and thin shells, showing the electric potential, normal electric displacement, transverse displacement  $w \equiv u_z$ , and normal stress  $\sigma_{zz}$ . Table 6, left, referring to the one-layer actuator case, shows that higher-order expansions leads to the same results for both electric potential and mechanical displacement,

$R_\beta/h$	2	4	10	100	$R_\beta/h$	2	4	10	100
$\Phi(z=0)$					$\Phi(z=0)$				
LD4	0.3431	0.4611	0.5037	0.5254	LD4	0.0153	0.0355	0.0942	0.6513
LM1	0.5000	0.5000	0.5000	0.5000	LM1	0.0000	0.0000	0.0000	0.0000
LM4	0.3431	0.4611	0.5037	0.5254	LM4	0.0153	0.0355	0.0942	0.6513
LFM1	0.5000	0.5000	0.5000	0.5000	LFM1	0.0000	0.0000	0.0000	0.0000
LFM2	0.3415	0.4609	0.5037	0.5254	LFM2	0.0150	0.0350	0.0939	0.6513
LFM3	0.3436	0.4614	0.5037	0.5254	LFM3	0.0161	0.0359	0.0943	0.6514
LFM4	0.3431	0.4611	0.5037	0.5254	LFM4	0.0153	0.0355	0.0942	0.6513
$\mathcal{D}_z 10^{11}(z=h/2)$					$\mathcal{D}_z 10^9(z=h/2)$				
LD4	-605.73	-801.76	-1622.9	-10416	LD4	0.0224	0.1377	2.0958	1456.1
LM1	-350.00	-662.37	-1608.6	-11711	LM1	-0.1144	-0.2121	0.1311	1118.0
LM4	-605.73	-801.76	-1622.9	-10416	LM4	0.0224	0.1377	2.0958	1456.1
LFM1	-327.37	-642.20	-1599.7	-16656	LFM1	-0.1340	-0.3506	-1.7252	-197.53
LFM2	-529.21	-743.93	-1593.3	-16262	LFM2	0.0407	0.0378	-0.0937	-111.27
LFM3	-587.25	-787.72	-1616.7	-16266	LFM3	0.0045	0.0092	-0.1536	-111.76
LFM4	-584.80	-783.99	-1615.6	-16266	LFM4	0.0095	0.0028	-0.1646	-111.76
$W 10^{11}(z=0)$					$W 10^9(z=0)$				
LD4	-9.6220	-11.285	6.4540	11277	LD4	0.0566	0.3332	4.5483	3016.6
LM1	-21.851	-18.528	2.9362	9531.4	LM1	0.0623	0.2956	3.6212	2530.3
LM4	-9.6220	-11.285	6.4540	11277	LM4	0.0566	0.3332	4.5483	3016.6
LFM1	-21.850	-18.528	2.9353	9531.4	LFM1	0.0623	0.2955	3.6210	2530.3
LFM2	-9.4240	-11.030	6.5318	11276	LFM2	0.0530	0.3223	4.5169	3016.5
LFM3	-9.6061	-11.282	6.4540	11277	LFM3	0.0566	0.3331	4.5482	3016.6
LFM4	-9.6220	-11.285	6.4540	11277	LFM4	0.0566	0.3332	4.5483	3016.6
$\sigma_{zz}(z=h/2)$					$\sigma_{zz}(z=-h/2)$				
LD4	3.0350	2.5512	1.0498	-836.32	LD4	-1.5105	-3.0742	-33.331	-22413
LM1	-0.4037	-0.1264	0.6259	54.231	LM1	-1.0564	-3.6212	-20.878	-1473.3
LM4	0.0431	0.0114	0.0004	-0.0001	LM4	-1.1392	-1.0671	-1.0269	-1.0018
LFM1	-0.4037	-0.1264	0.6258	54.231	LFM1	-1.0551	-3.6198	-20.877	-1473.3
LFM2	0.4285	0.4332	0.1982	-0.3861	LFM2	-1.1380	-1.5644	-2.6532	-12.288
LFM3	-0.0260	-0.0347	0.0005	0.0098	LFM3	-1.3463	-1.3907	-1.4043	-1.2716
LFM4	0.0431	0.0114	0.0004	-0.0001	LFM4	-1.1392	-1.0671	-1.0269	-1.0018

**Table 6.** Proposed benchmark: one-layer piezoelectric shell in the Varadan–Bhaskar geometry (Figure 4). Comparison of approaches. Left, actuator case; right, sensor case.

regardless of variational statement. Normal stress fulfills the boundary conditions at the top ( $\sigma_{zz} = 0$  at the top for actuator case) in the case of P-RMVT and F-RMVT; however higher-order expansions are required. Normal electric displacement results related to F-RMVT applications are quite different with respect to PVD and P-RMVT ones (where  $\mathcal{D}_z$  is not an assumed variable). These differences are larger for the sensor case, as will be seen.

Table 6, right, refers to the sensor case. The same comments for the left half of the table are made for electric potential, mechanical displacement and normal stress evaluation. Mechanical loading is applied at the bottom of the shell; such a boundary condition is exactly fulfilled by P-RMVT and F-RMVT applications. F-RMVT is more effective with respect to the other models for normal electric displacement evaluations.

The comments made in connection with Table 6 are confirmed for Table 7, which refers to a four-layer configuration. The differences in the evaluation of  $\mathcal{D}_z$  are, in fact, larger for the sensor case. The presence of two mechanical layers reduces the electric coupling. Furthermore, for the sensor case the effect of the radii of curvature is larger than in the actuator case. The fulfillment of boundary conditions for transverse normal stress at the shell bottom demands the use of higher-order expansions and mixed theories; this is confirmed by an analysis of Table 7 (in contrast, higher-order PVD theories leads to an accurate evaluation of  $\sigma_{zz}$  in the case of a plate. This is not so in the shell case because of the curvature).

Figure 5 shows that discrepancies for  $\mathcal{D}_z$  among F-RMVT, P-RMVT and PVD increase in the sensor case. Figure 6 refers to a multilayer shell: the capability of F-RMVT analysis to fulfill the interlaminar continuity of  $\mathcal{D}_z$  is clearly shown. Transverse normal stress evaluation is given in Figure 7. The superiority of F-RMVT and P-RMVT with respect to PVD is confirmed.

### Appendix: Explicit forms of fundamental nuclei

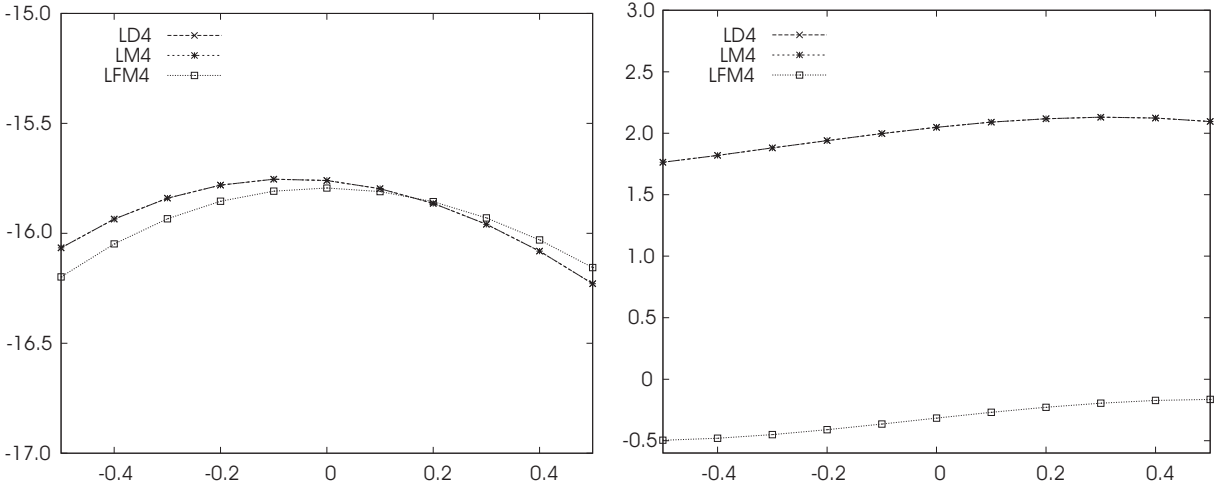
We detail here the algebraic forms related to closed-form solution for the nuclei  $\mathbf{K}_{u^{\mathcal{D}}}^{k\tau s}$ :

$$\begin{aligned}
 (K_{u^{\mathcal{D}}})_{11} &= \frac{\alpha J_{\beta}^{k\tau s} (C_{nn33}e_{31} - C_{pn13}e_{33})}{e_{33}^2 + C_{nn33}\varepsilon_{33}}, & (K_{u^{\mathcal{D}}})_{12} &= 0, & (K_{u^{\mathcal{D}}})_{13} &= 0, \\
 (K_{u^{\mathcal{D}}})_{21} &= \frac{\beta J_{\alpha}^{k\tau s} (C_{nn33}e_{32} - C_{pn23}e_{33})}{e_{33}^2 + C_{nn33}\varepsilon_{33}}, & (K_{u^{\mathcal{D}}})_{22} &= 0, & (K_{u^{\mathcal{D}}})_{23} &= 0, \\
 (K_{u^{\mathcal{D}}})_{31} &= \frac{-C_{nn33}e_{31} \frac{\lambda_D}{R_{\alpha}} J_{\beta}^{k\tau s} - C_{nn33}e_{32} \frac{\lambda_D}{R_{\beta}} J_{\alpha}^{k\tau s} + C_{pn13}e_{33} \frac{\lambda_D}{R_{\alpha}} J_{\beta}^{k\tau s} + C_{pn23}e_{33} \frac{\lambda_D}{R_{\beta}} J_{\alpha}^{k\tau s}}{e_{33}^2 + C_{nn33}\varepsilon_{33}}, & (K_{u^{\mathcal{D}}})_{32} &= 0, & (K_{u^{\mathcal{D}}})_{33} &= 0,
 \end{aligned}$$

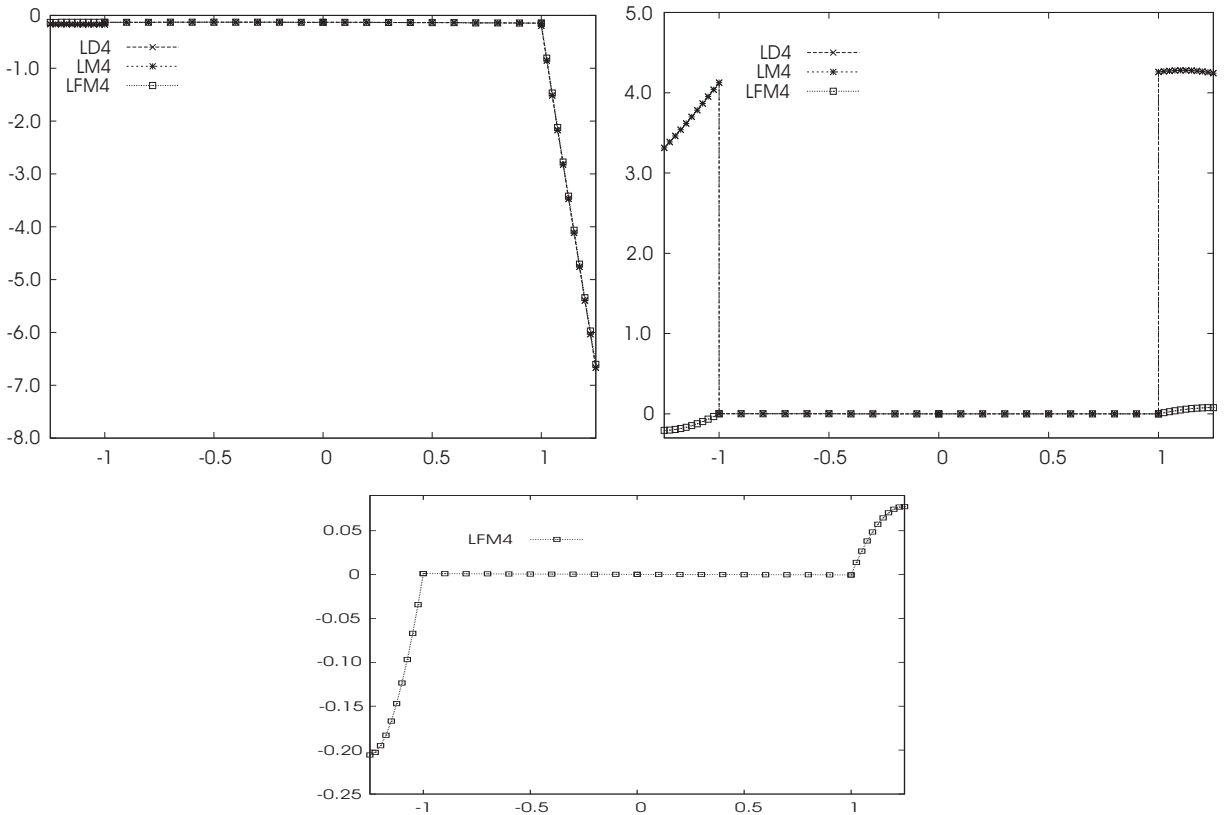
where we have introduced the integrals

$$J_{\alpha}^{k\tau s} = \int_{h_k} F_{\tau} F_s H_{\alpha} dz, \quad J_{\beta}^{k\tau s} = \int_{h_k} F_{\tau} F_s H_{\beta} dz. \tag{30}$$

The particular case of flat plates can be easily obtained by taking  $R_{\alpha}^k = R_{\beta}^k = \infty$ .



**Figure 5.** Proposed benchmark: plot of  $\mathcal{D}_z 10^9$  versus  $z$  for one-layer piezoelectric shell, with  $R_\beta/h = 10$ . Left, actuator case; right, sensor case.



**Figure 6.** Proposed benchmark: plot of  $\mathcal{D}_z 10^{10}$  versus  $z$  for multilayer piezoelectric shell, with  $R_\beta/h = 10$ . Top left, actuator case; top right, sensor case. The bottom figure plots the zoom of LFM4 results of sensor case.

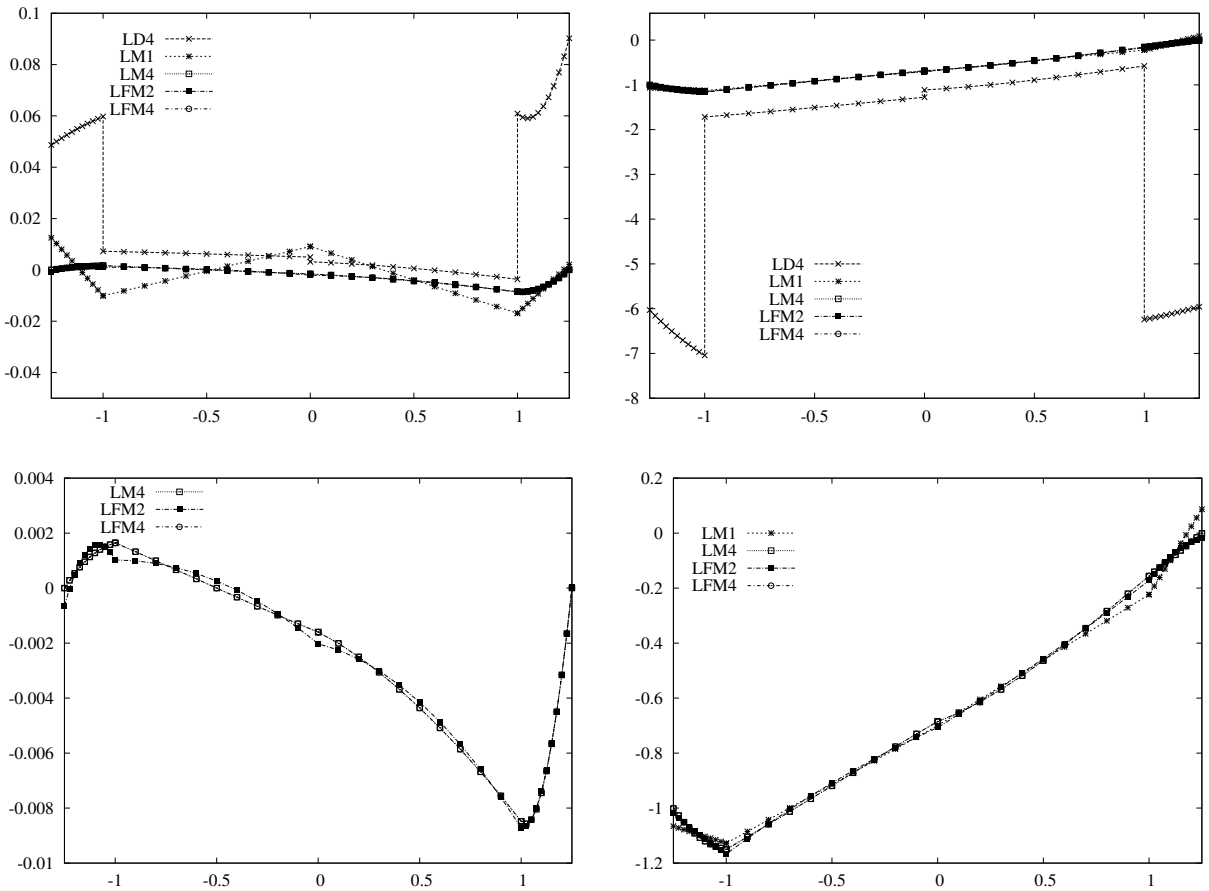
$R_\beta/h$	2	4	10	100	$R_\beta/h$	2	4	10	100
$\Phi (z = 0)$					$\Phi (z = 0)$				
LD4	0.4064	0.4829	0.5029	0.5009	LD4	0.0039	0.0157	0.0485	0.3414
LM1	0.4014	0.4824	0.5029	0.5009	LM1	0.0036	0.0154	0.0484	0.3415
LM4	0.4064	0.4829	0.5029	0.5009	LM4	0.0039	0.0157	0.0485	0.3414
LFM1	0.3805	0.4784	0.5041	0.5012	LFM1	0.0013	0.0133	0.0458	0.3287
LFM2	0.4059	0.4827	0.5029	0.5009	LFM2	0.0038	0.0156	0.0485	0.3414
LFM3	0.4065	0.4829	0.5029	0.5009	LFM3	0.0039	0.0157	0.0485	0.3414
LFM4	0.4064	0.4829	0.5029	0.5009	LFM4	0.0039	0.0156	0.0485	0.3414
$\mathcal{D}_z 10^9 (z = h/2)$					$\mathcal{D}_z 10^{11} (z = h/2)$				
LD4	-1.0754	-0.6666	-0.3322	-0.3494	LD4	9.8912	42.445	391.73	227910
LM1	-0.6686	-0.4157	-0.2172	-0.3373	LM1	5.1565	31.407	347.52	225160
LM4	-1.0754	-0.6666	-0.3322	-0.3494	LM4	9.8858	42.441	391.73	227910
LFM1	-1.0844	-0.6674	-0.3285	-0.3623	LFM1	1.1188	1.6464	1.9872	-1.6061
LFM2	-1.0639	-0.6600	-0.3268	-0.3622	LFM2	0.5401	0.6890	0.8068	-2.5408
LFM3	-1.0655	-0.6603	-0.3269	-0.3622	LFM3	0.6220	0.7813	0.9010	-2.4677
LFM4	-1.0654	-0.6603	-0.3269	-0.3622	LFM4	0.6092	0.7747	0.9001	-2.4676
$W 10^{11} (z = 0)$					$W 10^9 (z = 0)$				
LD4	-1.1542	-1.0208	-1.0048	2.4869	LD4	0.2633	0.9437	7.9334	4403.2
LM1	-1.2671	-1.0582	-1.0290	2.4730	LM1	0.2548	0.9386	7.9314	4404.3
LM4	-1.1542	-1.0208	-1.0048	2.4869	LM4	0.2633	0.9437	7.9334	4403.2
LFM1	-1.2320	-1.0342	-1.0112	2.4838	LFM1	0.2539	0.9377	7.9271	4403.0
LFM2	-1.1362	-1.0171	-1.0043	2.4869	LFM2	0.2623	0.9426	7.9307	4403.2
LFM3	-1.1560	-1.0209	-1.0048	2.4869	LFM3	0.2634	0.9438	7.9335	4403.2
LFM4	-1.1542	-1.0208	-1.0048	2.4869	LFM4	0.2633	0.9437	7.9334	4403.2
$\sigma_{zz} (z = h/2)$					$\sigma_{zz} (z = -h/2)$				
LD4	0.1416	0.0902	0.0757	-0.1835	LD4	-2.2444	-6.0302	-52.912	-32549
LM1	-0.0062	0.0021	0.0023	0.0006	LM1	-1.1772	-1.0658	-0.8449	11.453
LM4	0.0000	0.0000	0.0000	0.0000	LM4	-1.0013	-1.0010	-1.0006	-1.0000
LFM1	0.0055	0.0074	0.0043	0.0008	LFM1	-1.1791	-1.0667	-0.8440	11.512
LFM2	0.0008	0.0000	0.0001	0.0000	LFM2	-0.9992	-1.0172	-1.0857	-1.7388
LFM3	0.0001	0.0000	0.0000	0.0000	LFM3	-0.9983	-1.0057	-1.0087	-1.0071
LFM4	0.0000	0.0000	0.0000	0.0000	LFM4	-1.0013	-1.0010	-1.0006	-0.9999

**Table 7.** Proposed benchmark: multilayer piezoelectric shell in the Varadan–Bhaskar geometry (Figure 4). Comparison of approaches. Left, actuator case; right, sensor case.

**References**

[Ballhause et al. 2004] D. Ballhause, M. D’Ottavio, B. Kröplin, and E. Carrera, “A unified formulation to assess multilayered theories for piezoelectric plates”, *Comput. Struct.* **83**:15–16 (2004), 1217–1235.  
 [Carrera 1998] E. Carrera, “Evaluation of layer-wise mixed theories for laminated plates analysis”, *AIAA J.* **36** (1998), 830–839.





**Figure 7.** Proposed benchmark: plot of  $\sigma_{zz}$  versus  $z$  for multilayer piezoelectric shell, with  $R_\beta/h = 4$ . Left, actuator case; right, sensor case. The lower panes show details of the graphs. Mixed and full mixed theories are considered.

[Carrera 1999a] E. Carrera, “Multilayered shell theories accounting for layerwise mixed description, 1: governing equations”, *AIAA J.* **37**:9 (1999), 1107–1116.

[Carrera 1999b] E. Carrera, “Multilayered shell theories accounting for layerwise mixed description, 2: numerical evaluations”, *AIAA J.* **37**:9 (1999), 1117–1124.

[Carrera 2001] E. Carrera, “Developments, ideas and evaluations based upon the Reissner’s Mixed Theorem in the modelling of multilayered plates and shells”, *Appl. Mech. Rev.* **54**:4 (2001), 301–329.

[Carrera 2003] E. Carrera, “Theories and finite elements for multilayered plates and shells: a unified compact formulation with numerical assessment and benchmarking”, *Arch. Comput. Methods Eng.* **10**:3 (2003), 215–296.

[Carrera and Boscolo 2006] E. Carrera and M. Boscolo, “Classical and mixed finite elements for static and dynamics analysis of piezoelectric plates”, *Int. J. Numer. Methods Eng.* (2006). Published online 6 Nov 2006.

[Carrera and Brischetto 2007] E. Carrera and S. Brischetto, “Reissner mixed theorem applied to bending analysis of piezoelectric shells”, *J. Intell. Mater. Syst. Struct.* (2007). In press.

[Carrera and Fagiano 2007] E. Carrera and C. Fagiano, “Full and partial extension of mixed finite elements for static analysis of piezoelectric plates”, *J. Intell. Mater. Syst. Struct.* (2007). In press.

- [Chen et al. 1996] C.-Q. Chen, Y.-P. Shen, and X.-M. Wang, "Exact solution of orthotropic cylindrical shell with piezoelectric layers under cylindrical bending", *Int. J. Solids Struct.* **33**:30 (1996), 4481–4494.
- [Cho and Roh 2003] M. Cho and H. Y. Roh, "Development of geometrically exact new shell elements based on general curvilinear co-ordinates", *Int. J. Numer. Methods Eng.* **56**:1 (2003), 81–115.
- [Chopra 2002] I. Chopra, "Review of state of art of smart structures and integrated systems", *AIAA J.* **40**:11 (2002), 2145–2187.
- [Crawley 1994] E. F. Crawley, "Intelligent structures for aerospace: a technology overview and assessment", *AIAA J.* **32**:8 (1994), 1689–1699.
- [D'Ottavio and Kröplin 2006] M. D'Ottavio and B. Kröplin, "An extension of Reissner mixed variational theorem to piezoelectric laminates", *Mech. Adv. Mater. Struct.* **13**:2 (2006), 139–150.
- [D'Ottavio et al. 2006] M. D'Ottavio, D. Ballhause, B. Kröplin, and E. Carrera, "Closed-form solutions for the free vibration problem of multilayered piezoelectric shells", *Comput. Struct.* **84**:22–23 (2006), 1506–1518.
- [Dumir et al. 1997] P. Dumir, G. Dube, and S. Kapuria, "Exact piezoelectric solution of simply-supported orthotropic circular cylindrical panel in cylindrical bending", *Int. J. Solids Struct.* **34**:6 (1997), 685–702.
- [Garcia Lage et al. 2004] R. Garcia Lage, C. M. Mota Soares, C. A. Mota Soares, and J. N. Reddy, "Layerwise partial mixed finite element analysis of magneto-electro-elastic plates", *Comput. Struct.* **82**:17–19 (2004), 1293–1301.
- [Heyliger 1997] P. Heyliger, "Exact solutions for simply supported laminated piezoelectric plates", *J. Appl. Mech. (ASME)* **64** (1997), 299–306.
- [Heyliger et al. 1996] P. Heyliger, K. C. Pei, and D. Saravanos, "Layerwise mechanics and finite element model for laminated piezoelectric shells", *AIAA J.* **34**:11 (1996), 2353–2360.
- [Ikeda 1996] T. Ikeda, *Fundamentals of piezoelectricity*, Corrected paperback ed., Oxford Science Publications, Oxford University Press, New York, 1996.
- [Kögl and Bucalem 2005] M. Kögl and M. L. Bucalem, "Analysis of smart laminates using piezoelectric MITC plate and shell elements", *Comput. Struct.* **83**:15–16 (2005), 1153–1163.
- [Lammering and Mesecke-Rischmann 2003] R. Lammering and S. Mesecke-Rischmann, "Multifield variational formulations and related finite elements for piezoelectric shells", *Smart Mater. Struct.* **12**:6 (2003), 904–913.
- [Murakami 1986] H. Murakami, "Laminated composite plates theory with improved in-plane response", *J. Appl. Mech. (ASME)* **53** (1986), 661–666.
- [Reddy 2004] J. N. Reddy, *Mechanics of laminated composite plates and shells: theory and analysis*, CRC Press, Boca Raton, FL, 2004.
- [Ren 1987] J. G. Ren, "Exact solutions for laminated cylindrical shells in cylindrical bending", *Compos. Sci. Technol.* **29**:3 (1987), 169–187.
- [Robaldo et al. 2006] A. Robaldo, E. Carrera, and A. Benjeddou, "A unified formulation for finite element analysis of piezoelectric adaptive plates", *Comput. Struct.* **84**:22–23 (2006), 1415–1505.
- [Robbins and Chopra 2006] D. H. Robbins and I. Chopra, "The effect of laminated kinematic assumptions on the global response of actuated plates", *J. Intell. Mater. Syst. Struct.* **17**:4 (2006), 273–299.
- [Rogacheva 1994] N. L. Rogacheva, *The theory of piezoelectric shells and plates*, CRC Press, Boca Raton, FL, 1994.
- [Saravanos and Heyliger 1999] D. A. Saravanos and P. R. Heyliger, "Mechanics and computational models for laminated piezoelectric beams, plates, and shells", *Appl. Mech. Rev.* **52**:10 (1999), 305–320.
- [Shakeri et al. 2006] M. Shakeri, M. Eslami, and A. Daneshmehr, "Dynamic analysis of thick laminated shell panel with piezoelectric layer based on three dimensional elasticity solution", *Comput. Struct.* **84**:22–23 (2006), 1519–1526.
- [Varadan and Bhaskar 1991] T. K. Varadan and K. Bhaskar, "Bending of laminated orthotropic cylindrical shells: an elasticity approach", *Compos. Struct.* **17**:2 (1991), 141–156.
- [Wang et al. 2005] H. M. Wang, H. J. Ding, and Y. M. Chen, "Dynamic solution of a multilayered orthotropic piezoelectric hollow cylinder for axisymmetric plane strain problems", *Int. J. Solids Struct.* **42**:1 (2005), 85–102.

Received 24 Jul 2006. Accepted 9 Oct 2006.

ERASMO CARRERA: [erasmo.carrera@polito.it](mailto:erasmo.carrera@polito.it)

Dept. of Aeronautics and Aerospace Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy

SALVATORE BRISCHETTO: [salvatore.brischetto@polito.it](mailto:salvatore.brischetto@polito.it)

*Dept. of Aeronautics and Aerospace Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy*

Heshmat A. !Aglan and Mahmood Fateh

Fracture and fatigue crack growth analysis of rail steels Mahmood !Fateh with Heshmat A. Aglan  
Symmetry conditions and finite mechanisms Patrick W. !Fowler with Simon D. Guest

Single member actuation of kagome lattice structures Simon D. !Guest with Anthony C. H. Leung

Single member actuation of kagome lattice structures Simon D. !Guest with Anthony C. H. Leung  
Yoshihiro !Kanno and Izuru Takewaki

Behavior of paper on a high speed conveyor subjected to air jet impingement: a method for estimating  
bending stiffness Kalyan C. !Katuri with M. K. Ramasubramanian and Richard A. Venditti

Anthony C. H. !Leung and Simon D. Guest

An Eshelby-type approach for defect energetics in carbon nanotubes Jia !Lu with Liang Zhang

Massimiliano !Lucchesi, Cristina Padovani, Giuseppe Pasquinelli and Nicola Zani

Static analysis of masonry vaults, constitutive model and numerical analysis Cristina !Padovani with  
Massimiliano Lucchesi, Giuseppe Pasquinelli and Nicola Zani

Static analysis of masonry vaults, constitutive model and numerical analysis Giuseppe !Pasquinelli  
with Massimiliano Lucchesi, Cristina Padovani and Nicola Zani

M. K. !Ramasubramanian, Richard A. Venditti and Kalyan C. Katuri

Worst case plastic limit analysis of trusses under uncertain loads via mixed 0-1 programming Izuru  
!Takewaki with Yoshihiro Kanno

Behavior of paper on a high speed conveyor subjected to air jet impingement: a method for estimating  
bending stiffness Richard A. !Venditti with M. K. Ramasubramanian and Kalyan C. Katuri

Shmuel !Vigdergauz

Static analysis of masonry vaults, constitutive model and numerical analysis Nicola !Zani with Mas-  
similiano Lucchesi, Cristina Padovani and Giuseppe Pasquinelli

Liang !Zhang and Jia Lu

# SUBMISSION GUIDELINES

## ORIGINALITY

Authors may submit manuscripts in PDF format on-line. 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:

Marie-Louise Steele  
Division of Mechanics and Computation  
Durand Building, Room 262  
Stanford University  
Stanford CA 94305

## 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 are encouraged to use L<sup>A</sup>T<sub>E</sub>X and the standard article class, but submissions in other varieties of T<sub>E</sub>X, and, exceptionally in other formats, are acceptable. Electronic submissions are strongly encouraged in PDF format only; after the refereeing process we will ask you to submit all source material.

## REFERENCES

Bibliographical references should be listed alphabetically at the end of the paper and include the title of the article. All references in the bibliography should be cited in the text. The use of B<sup>I</sup>B<sup>T</sup><sub>E</sub>X is preferred but not required. Tags will be converted to the house format (see a current issue for examples), however, in the manuscript, the citation should be by first author's last name and year of publication, e.g. "as shown by Kramer, et al. (1994)". Links will be provided to all literature with known web locations and authors are encouraged to provide their own links on top of the ones provided by the editorial process.

## FIGURES

Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple, or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to:

[production@mathscipub.org](mailto:production@mathscipub.org)

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 2 No. 2 February 2007

---

Behavior of paper on a high speed conveyor subjected to air jet impingement: a method for estimating bending stiffness	201
M. K. RAMASUBRAMANIAN, RICHARD A. VENDITTI AND KALYAN C. KATURI	
Static analysis of masonry vaults, constitutive model and numerical analysis	221
MASSIMILIANO LUCCHESI, CRISTINA PADOVANI, GIUSEPPE PASQUINELLI AND NICOLA ZANI	
Worst case plastic limit analysis of trusses under uncertain loads via mixed 0-1 programming	245
YOSHIHIRO KANNO AND IZURU TAKEWAKI	
Shape optimization of a rigid inclusion in a shear-loaded elastic plane	275
SHMUEL VIGDERGAUZ	
Symmetry conditions and finite mechanisms	293
SIMON D. GUEST AND PATRICK W. FOWLER	
Single member actuation of kagome lattice structures	303
ANTHONY C. H. LEUNG AND SIMON D. GUEST	
An Eshelby-type approach for defect energetics in carbon nanotubes	319
LIANG ZHANG AND JIA LU	
Fracture and fatigue crack growth analysis of rail steels	335
HESHMAT A. AGLAN AND MAHMOOD FATEH	
Bistable structures for energy absorption I. Metallic structures under tension	347
ZACHARY WHITMAN AND VALERIA LA SAPONARA	
Bistable structures for energy absorption II. Composite structures under tension	359
ZACHARY WHITMAN AND VALERIA LA SAPONARA	
Piezoelectric shell theories with <i>a priori</i> continuous transverse electromechanical variables	377
ERASMO CARRERA AND SALVATORE BRISCHETTO	



1559-3959(200702)2:2;1-D