A computationally tractable framework for nonlinear dynamic multiscale modeling of membrane woven fabrics

Philip Avery | Daniel Z. Huang | Wanli He | Johanna Ehlers | Armen Derkevorkian | Charbel Farhat

Abstract
A general-purpose computational homogenization framework is proposed for the nonlinear dynamic analysis of membranes exhibiting complex microscale and/or mesoscale heterogeneity characterized by in-plane periodicity that cannot be effectively treated by a conventional method, such as woven fabrics. The framework is a generalization of the “finite element squared” (or FE²) method in which a localized portion of the periodic subscale structure is modeled using finite elements. The numerical solution of displacement driven problems involving this model can be adapted to the context of membranes by a variant of the Klinkel–Govindjee method originally proposed for using finite strain, three-dimensional material models in beam and shell elements. This approach relies on numerical enforcement of the plane stress constraint and is enabled by the principle of frame invariance. Computational tractability is achieved by introducing a regression-based surrogate model informed by a physics-inspired training regimen in which FE² is utilized to simulate a variety of numerical experiments including uniaxial, biaxial and shear straining of a material coupon. Several alternative surrogate models are evaluated including an artificial neural network. The framework is demonstrated and validated for a realistic Mars landing application involving supersonic inflation of a parachute canopy made of woven fabric.

KEYWORDS
artificial neural network, FE squared, membrane, multiscale, parachute, woven fabrics

1 INTRODUCTION
Nonlinear multiscale problems—defined here as nonlinear problems exhibiting vastly different scale features that are significant to the macroscopic behavior—are ubiquitous in science and engineering. They arise, for example, in the modeling of woven fabrics (see Figure 1) used in body armor and inflatable structures such as vehicle air bags, parachutes, and other atmospheric decelerators; and in the modeling of textiles within the context of forming processes for woven composites. Numerical methods that attempt to resolve all relevant scales typically lead to massive discretized problems. However, recent developments using a variety of alternative surrogate modeling techniques—including nonlinear, projection-based model order reduction (PMOR), kriging, and artificial neural networks (NNs)—to accelerate the solution of one or
more scales within the context of a computational homogenization framework present a coherent methodology by which a computationally tractable approximation can be attained without resorting to ad hoc approximations. Notably, thin shell and membrane discretizations have not been considered in this context prior to this work, although several frameworks for multiscale modeling of shells without emphasis on computational efficiency have been proposed. In particular, this article addresses the case of a hybrid discretization in which plane stress membrane elements are employed at the macroscopic scale, for the sake of convenience and numerical efficiency; but three-dimensional (3D) solid elements are preferred at the mesoscopic and/or microscopic scales for the sake of generality and in order to most precisely represent geometric features and deformation modes at these scales.

In this article, the presented multiscale framework is built on a macro–micro concept that generalizes to \( n \)-levels, although without loss of generality, only the two-level case is presented, anticipating that this case is sufficient for many problems of interest. The framework allows for the treatment of unilateral contact constraints at both macroscale and microscale. It features a microscale model discretized with solid elements, to allow accurate representation of microscopic geometric features such as yarns and voids. However, the proposed framework readily generalizes to alternative microscale discretizations such as shell elements. For the case of a macroscale model that is also discretized with solid elements, well-established localization/homogenization scale bridging strategies have been developed and typically provide a mapping between the 3D deformation gradient \( \mathbf{F} \) and first Piola–Kirchhoff stress \( \mathbf{P} \) tensors from which a constitutive relation is inferred. However, when the macroscale model is discretized with membrane elements and the microscale is discretized with solid elements, coupling between the two scales require careful attention. The treatment proposed in this article has two novel components:

- First, it is observed that due to the principle of material frame invariance, the conventional 3D \( \mathbf{F}–\mathbf{P} \) scale bridging approach can be reformulated using the polar decomposition of the deformation gradient to furnish a mapping between the right stretch tensor and its conjugate, the symmetric Biot stress tensor. Using this reformulation combined with some straightforward transformations, it is shown that an unconventional FE\(^2 \) material model can conveniently be used as a drop-in replacement for any conventional 3D material model formulated as a mapping between the commonly used Green–Lagrangian strain \( \mathbf{E} \) and its conjugate, the second Piola–Kirchhoff stress tensor \( \mathbf{S} \). The relevance of this development to the issue of multiscale membrane-solid coupling will be addressed in what follows.

- Material models used in membrane elements are typically of the plane stress variety. In some cases, a plane stress variant of a 3D material model for which the plane stress condition is enforced analytically can be derived. When an analytical solution is not available, numerical enforcement of the plane stress condition is commonly used, for example, in the case of J2 elastoplasticity. This involves solving numerically—using a root-finding method such as Newton’s method or the bisection method—a nonlinear equation to enforce the plane stress condition. Klinkel and Govindjee have shown how numerical enforcement of the plane stress condition can be used to construct an interface that, in principle, enables any 3D material model to be “converted” into a plane stress variant that can then be used in a shell or beam element. Because the plane stress condition is typically expressed by constraints on the out-of-plane components of the second Piola–Kirchhoff stress tensor, the method of Klinkel and Govindjee was presented in the most convenient way—that is, using the \( \mathbf{E}–\mathbf{S} \) conjugate pair. Here, it is simply noted that this method can be trivially adapted...
to membranes and furthermore can be conveniently used with any material model of the form $S = \dot{S}(E)$ including, but not limited to, constitutive relations inferred from a FE$^2$ computational homogenization formulated using polar decomposition as outlined above.

Motivated by the fact that the proposed plane stress constitutive law is essentially a mapping between two pairs of 3D vectors, a lightweight alternative is considered, in which a regression-based model is used as a surrogate for constitutive function evaluations that would otherwise require the solution of a finite element model of the microscale RVE. Three alternative surrogates, each capable of achieving computational tractability, are presented and evaluated: (1) the classical linear elastic model fitted to data using linear regression; (2) a quadratic model fitted to data using linear regression; and (3) an artificial NN model fitted to data using the PyTorch library. In each case, the data used to train and test the model is obtained by exercising the proposed high-fidelity multiscale membrane model on a series of numerical experiments intended to mimic the familiar physical experimental-based methodology typically used in the development of conventional material models. A novel, nonlinear, PMOR approach is also proposed for accelerating the training process and described in Appendix A.

The remainder of this article is organized as follows. Section 2 provides an overview of the proposed two-level multiscale framework with a locally attached microscale, focusing on the context of large-deformation structural mechanics with macroscopic discretization using membrane elements, microscale discretization using 3D solid elements, bridging between the scales and the solution of the discrete coupled multiscale problem including treatment of contact at both scales. In Section 3, three regression-based surrogate microscale models and their training methodologies are presented and compared. A numerical example is provided in Section 4 to evaluate the proposed framework, involving a realistic, coupled, multiscale fluid–structure simulation of the deployment of a disk-gap-band (DGB) parachute in the Martian atmosphere. Finally, conclusions are offered in Section 5.

2 | MULTISCALE FORMULATION FOR MEMBRANES BASED ON A LOCALLY ATTACHED MICROSTRUCTURE

In this section, a multiscale continuum mechanics formulation suitable for membranes and based on the concept of a locally attached microstructure is presented. As formulated, the stress–strain relationship for a heterogeneous membrane is not defined by a conventional plane stress constitutive law, but rather by: (a) the solution at each material point of one or more boundary value problems governing its microstructure; and (b) the numerical enforcement of the plane stress condition. Although the concept generalizes naturally to three or more scales, it is presented and applied here to problems that exhibit precisely two separate scales – for the sake of clarity. Specifically, the stress–strain relationship at the coarse scale is defined by the solution of boundary value problems at the fine scale, an appropriate scale transition method and a constraint on the out-of-plane components of the homogenized stress tensor. At the finest scale, where all heterogeneities can be adequately resolved and described by an available constitutive theory, it is defined by an analytical constitutive law. All considered length scales are assumed to be much larger than the molecular dimension so that the continuum assumption holds. Furthermore, scale separation is assumed to loosely couple the various scales through localization from coarse to fine scales and homogenization from fine to coarse scales. For further details, the reader is referred to References 11,14-16 for the concept of a locally attached microstructure. The approach adopted here can be interpreted as a generalization and/or application to the case of membranes of the localization/homogenization scale bridging strategy presented in Reference 11.

2.1 | Preliminaries

Consider a domain $\mathcal{B}_0 \subset \mathbb{R}^3$ defining a highly heterogeneous membrane structure of interest. Assume that its boundary $\partial \mathcal{B}_0$ is subject to prescribed displacements on $\partial \mathcal{B}_0^u \subset \partial \mathcal{B}_0$ and tractions on $\partial \mathcal{B}_0 \setminus \partial \mathcal{B}_0^u$. Let $\varphi_0^t : \mathcal{B}_0 \to \mathcal{B}_0^t$ denote the nonlinear transformation that maps a point in the reference configuration, $X_0 \in \mathcal{B}_0$, at time $t$, to a counterpart in the current configuration, $x_0(X_0; t) = \varphi_0^t(X_0) \in \mathcal{B}_0^t$. Here, the current configuration of the membrane $\mathcal{B}_0^t$ is defined as

$$\mathcal{B}_0^t = \{x_0 \in \mathbb{R}^3 \mid x_0 = \phi (\xi^{(1)}, \xi^{(2)}) + \xi^{(3)} n_0\},$$

where $\phi$ is a nonlinear function, and $n_0$ is the normal to the boundary of the reference configuration.
where the map $\Phi : \mathcal{A}_0 \to \mathbb{R}^3$ defines the reference position of the mid-surface of the membrane, and $\mathcal{E}_0 \subset \mathbb{R}^2$ are coordinates parameterizing the mid-surface, $n_0 \in \mathbb{R}^3$ is the unit normal to the mid-surface in the current configuration, $\xi^{(3)} \in [-h/2,h/2]$ is a coordinate parameterizing the direction normal to the surface and $h$ is the upper bound of the membrane thickness. Similarly, the reference configuration of the membrane $\mathcal{B}_0$ is defined as

$$\mathcal{B}_0 = \{ \mathbf{X}_0 \in \mathbb{R}^3 \mid \mathbf{X}_0 = \Phi (\xi^{(1)}, \xi^{(2)}) + \xi^{(3)} \mathbf{N}_0 \} .$$

where the map $\Phi : \mathcal{A}_0 \to \mathbb{R}^3$ defines the reference position of the mid-surface of the membrane and $\mathbf{N}_0 \in \mathbb{R}^3$ is the unit normal to the mid-surface in the reference configuration. The deformation of this domain is governed by a reduction of the finite deformation continuum equations to the mid-surface with a plane stress but otherwise unknown constitutive law due to the assumed highly heterogeneous fine scale structure. For this reason, generalizing the work described in References 11, 14-16, the deformation problem is solved here by locally attaching an appropriately defined microstructure to each mid-surface point, computing the stress–strain relationship at each such point through the solution of a microstructure boundary value problem, bridging the scales via a localization and homogenization strategy and numerically enforcing the plane stress constraint on the resulting homogenized stress tensor. An appropriately defined microstructure in this context is one that represents only a minuscule “representative surface element” of the membrane within which the entire thickness of the membrane is accounted for. Hence, the range of the in-plane coordinates $\xi^{(1)}, \xi^{(2)}$ in the microscale domain should be much smaller than in the macroscale domain, while the ranges of the normal coordinate $\xi^{(3)}$ should be identical in both domains. The separation of scales and assumed periodicity in only two of the three spatial dimensions are notable characteristics of the problem of interest and its proposed treatment that distinguish it from the ubiquitous alternative multiscale treatments devised for fully 3D scale bridging.

Here and throughout the remainder of this article, the subscripts 0 and 1 denote quantities associated with the coarse (zeroth) and fine (first) scales, respectively. For simplicity, an $l$th scale is also referred to as scale $l$ or level $l$, interchangeably. The deformation at both scales is governed by the finite deformation continuum equations, with the stress–strain relationship defined by the solution of a constrained boundary value problem formulated at a finer scale for level $l = 0$, or an assumed constitutive law at the fine scale designated by level $l = 1$. Let $\varphi_1 : \mathcal{B}_1 \to \mathcal{B}_1'$ denote the nonlinear transformation that maps a point in the fine scale reference configuration, $\mathbf{X}_1 \in \mathcal{B}_1$, to a counterpart in the fine scale current configuration, $\mathbf{x}_1 (\mathbf{X}_1) = \varphi_1 (\mathbf{X}_1) \in \mathcal{B}_1'$. As in the formulation of the macroscale problem, $\partial \mathcal{B}_1$ is defined as the boundary of $\mathcal{B}_1$ and $\partial \mathcal{B}_1^m$ as its part where a displacement is prescribed.

The boundary conditions at scale 0 are defined by the physical problem of interest, while those at scale 1 depend on the deformations at the coarse scale. The constitutive law at the fine scale is chosen based on the expected response of this scale, while there is no preassigned constitutive law at the coarse scale but rather a dependence on the response of the microstructure to evaluate the constitutive function. Arbitrarily complex fine scale constitutive relationships involving nonlinearities and path-dependency are allowed, although in what follows only constitutive functions described by

$$\mathbf{S}_1 = \hat{\mathbf{S}}_1 (\mathbf{E}_1)$$

are considered, where $\mathbf{S}_1$ and $\mathbf{E}_1$ denote the microscale second Piola–Kirchhoff stress tensor and Green–Lagrange strain tensor, respectively, and $\hat{\mathbf{S}}_1$ is the microscale constitutive function. At the coarse scale, the intent is to devise a multiscale, plane stress constitutive function of the form

$$\mathbf{S}_0^m = \hat{\mathbf{S}}_0^m (\mathbf{E}_0^m) ,$$

(1)

where $\mathbf{S}_0$ and $\mathbf{E}_0$ denote the macroscale second Piola–Kirchhoff stress tensor and Green–Lagrange strain tensor, respectively; and the superscript $m$ applied to a tensor quantity designates the restriction of the tensor to its in-plane membrane components. For example, the membrane part of $\mathbf{S}_0$ is given by

$$\mathbf{S}_0^m = \begin{bmatrix} S_0^{11} & S_0^{12} \\ S_0^{21} & S_0^{22} \end{bmatrix} .$$

The superscript $m$ applied to a constitutive function (e.g., $\hat{\mathbf{S}}_0^m$) indicates that the function is a particular plane stress type of constitutive relation that evaluates the in-plane membrane components of a stress tensor while constraining its
out-of-plane components to be identically zero. A general numerical procedure for constructing such a function will be described subsequently.

2.2 Scale bridging

Following the work presented in Reference 11, the boundary conditions on $\mathcal{B}_1$ are defined so that the pointwise deformation gradient tensor at level 0, $F_0$, is equal to the volumetric average of the deformation gradient tensor at level 1—that is,

$$F_0 = \frac{1}{|\mathcal{B}_1|} \int_{\mathcal{B}_1} F_1 \, dV. $$

This *localization* transmission condition can be conveniently enforced by prescribing a boundary deformation of the form

$$x_1|_{\partial \mathcal{B}_1^*} = X_1|_{\partial \mathcal{B}_1^*} F_0 + w_1 \tag{2}$$

subject to some conditions (see Reference 2), where $w_1$ represents the nonuniform part of the boundary deformation. Without loss of generality, the uniform essential boundary condition $w_1 = 0$ is assumed.

The pointwise first Piola–Kirchhoff stress tensor at level 0 is defined as the volumetric average of the stress tensor at level 1

$$P_0 = \frac{1}{|\mathcal{B}_1|} \int_{\mathcal{B}_1} P_1 \, dV. \tag{3}$$

This *homogenization* transmission-type condition can be conveniently determined from quantities defined solely on $\partial \mathcal{B}_1^*$ by applying a Gauss-type identity to (3)

$$P_0 = \frac{1}{|\mathcal{B}_1|} \int_{\partial \mathcal{B}_1^*} P_1 N_1 \otimes X_1 \, dA \tag{4a}$$

$$= \frac{1}{|\mathcal{B}_1|} X_1^T f_1|_{\partial \mathcal{B}_1^*}, \tag{4b}$$

where $f_1|_{\partial \mathcal{B}_1^*}$ is the vector of so-called *reaction forces* associated with the prescribed deformations (2) and the superscript $T$ designates the transpose operation.

In this context, the microscale volume measure $|\mathcal{B}_1|$ should be interpreted as the entire volume of a bounding box enclosing the microscale volume (see Figure 2), including both regions of solid material and voids. The height of the bounding box $h + \epsilon$ should be slightly larger than the minimum enclosing dimension $h$ in the $\zeta(3)$ direction (i.e., $\epsilon > 0$) so that the microscale volume $\mathcal{B}_1$ does not intersect the box’s upper and lower faces. The magnitude of $\epsilon$ is otherwise arbitrary, as the dependence of the homogenized stress tensor on this parameter will be subsequently canceled when evaluating the membrane stress resultant. Note that the boundary $\partial \mathcal{B}_1^*$ used to define the transmission-type conditions is entirely contained within the four side faces of the bounding box, that is, the faces whose normals coincide with the $\zeta(1)$ and $\zeta(2)$ axes.

Equations (2) and (4) constitute a relation of the form

$$P_0 = \hat{P}_0(F_0) \tag{5}$$

that is evaluated in three steps as follows:

1. First, the microscale problem defined with the boundary conditions given in (2) is solved.
2. Next, the solution of the microscale problem is postprocessed to obtain the reaction forces. This can be done either by a boundary integral as shown in (4a), or by a volume integral over the region of the domain adjacent to the boundary.$^{17}$
3. Finally, the reaction forces are combined and scaled according to (4b) to produce the homogenized first Piola–Kirchhoff stress tensor $P_0$.

Unfortunately, (5) is not directly compatible with the stated application of interest—namely, a plane stress relation of the form (1) expressed in terms of the in-plane components of the Green–Lagrange strain and second Piola–Kirchhoff stress tensors. To formally adapt the homogenization methodology to this setting, it is first assumed without loss of generality that the relation (5) satisfies the principle of material frame invariance, which can be stated as follows

$$\hat{P}_0 (QF_0) = Q\hat{P}_0 (F_0), \quad \forall Q \in SO(3),$$

where $SO(3)$ is the group of special orthogonal transformations defined as

$$SO(3) = \{ Q \in \mathbb{R}^3 : Q^T Q = QQ^T = I, \quad \det(Q) = 1 \}$$

and $I$ is the identity matrix of dimension three. Regarding the assumption of material frame indifference, it can be shown under some conditions that (5) is objective and hence (6) holds. For cases where (6) does not hold, the alternative formulation proposed next can be interpreted as imposing or restoring material frame invariance, which is generally considered to be appropriate for constitutive relations in solid mechanics.

From the polar decomposition of the deformation gradient

$$F_0 = R_0 U_0,$$

where $R_0 \in SO(3)$ is the rotation tensor and $U_0 \in \mathbb{R}^{3 \times 3}$ is the symmetric positive definite right stretch tensor, and the following expression of the unsymmetric Biot stress tensor

$$B_0 = R_0^T P_0$$

it follows from (6) with $Q = R_0^T$ that the homogenized constitutive law (5) can equivalently be stated as a relation between the right stretch tensor and the unsymmetric Biot stress using the same functional form, that is,

$$B_0 = \hat{P}_0 (U_0).$$

This result can be simply interpreted as a variant of the standard transmission-type conditions (2, 4) in which the right stretch tensor is used instead of the deformation gradient to compute the microscale prescribed boundary deformations and the homogenized stress tensor obtained by evaluating the constitutive function $\hat{P}_0$ is identified as the Biot measure rather than the first Piola–Kirchhoff. Specifically,

$$x_1|_{\partial \mathcal{B}_1} = X_1|_{\partial \mathcal{B}_1} U_0, \quad (8a)$$

$$B_0 = \frac{1}{|\mathcal{B}_1|} X_1^f|_{\partial \mathcal{B}_1} f_1|_{\partial \mathcal{B}_1}, \quad (8b)$$
A more convenient relation between the Green–Lagrange strain and the first Piola–Kirchhoff stress tensor can be obtained from (7) by applying well-known transformations as follows:

1. First, the right stretch tensor can be obtained from the Green–Lagrange strain using

\[ U_0 = (C_0)^{\frac{1}{2}} = \sum_{i=1}^{3} \lambda_i N_i \otimes N_i, \]  

where \( C_0 = 2E_0 + I \) is the right Cauchy–Green deformation tensor and \( \lambda_i^2 \) and \( N_i \) are the eigenvalues and eigenvectors, respectively, of \( C_0 \).

2. Second, the second Piola–Kirchhoff stress can be obtained from the Biot stress using

\[ 0.5 (S_0 U_0 + U_0 S_0) = T_0, \]

where \( T_0 \) is the symmetric part of the Biot stress tensor \( T_0 = 0.5 (B_0 + B_0^T) \). Note that (10) has the form of the Lyapunov equation whose solution is given by a linear system of equations, namely

\[ \text{vec} (S_0) = [I \otimes U_0 + U_0^T \otimes I]^{-1} \text{vec} (2T_0), \]

where \( \otimes \) denotes the Kronecker product and \( \text{vec} (\cdot) \) denotes vectorization. For example, the vectorization of \( S_0 \) is given by

\[ \text{vec} (S_0) = \begin{bmatrix} S^{(11)}_0 & S^{(21)}_0 & S^{(31)}_0 & S^{(12)}_0 & S^{(22)}_0 & S^{(32)}_0 & S^{(13)}_0 & S^{(23)}_0 & S^{(33)}_0 \end{bmatrix}^T. \]

Due to symmetry, the dimension of (11) can be further reduced to six.

Substituting (9) and (11) into (7) produces a constitutive function relating the macroscale Green–Lagrange strain and second Piola–Kirchhoff stress tensor of the form

\[ S_0 = \hat{S}_0 (E_0) \]

that is evaluated in five steps as follows:

1. First, the macroscale right stretch tensor \( U_0 \) is computed from the Green–Lagrange strain \( E_0 \).
2. Second, the microscale problem with prescribed boundary values given by (8a) is solved.
3. Third, the solution of the microscale problem is postprocessed to obtain the reaction forces.
4. Next, the reaction forces are combined and scaled according to (8b) to produce the homogenized unsymmetric Biot stress tensor \( B_0 \).
5. Finally, the Lyapunov equation (11) is solved to get the homogenized second Piola–Kirchhoff stress tensor \( S_0 \).

### 2.3 Three-dimensional finite strain material models for membrane elements

The 3D constitutive law (12) can be adapted to plane stress (and hence membrane elements) using a variant of the method proposed by Klinkel and Govindjee13 for using finite strain 3D material models in beam and shell elements, which in turn is closely related to earlier methods proposed by De Borst21 and Dvorkin et al.22 This method involves solving a local nonlinear equation using Newton’s method to enforce the plane stress condition. Specifically, the requirement that the out-of-plane components of the second Piola–Kirchhoff stress tensor are zero, that is,

\[ S^{(0)}_0 = \begin{bmatrix} S^{(33)}_0 & S^{(13)}_0 & S^{(31)}_0 \end{bmatrix}^T = 0 \]

is enforced by iteratively solving for the corresponding out-of-plane components \( E^{(0)}_0 \) of the Green–Lagrange strain tensor, which are treated as unknowns in the above equations. Each Newton iteration incurs a single evaluation of the 3D
constitutive function (12) and its constitutive tangent. Solving the plane stress equation (13) for $E_0^s$ given $E_0^m$, then evaluating the in-plane components of the second Piola–Kirchhoff stress tensor $P_0^m$ at the resulting configuration corresponds to evaluating a plane stress constitutive relation of the form (1), which can be used as a drop-in replacement for a conventional finite strain plane stress constitutive equation. This will be demonstrated in what follows using the general purpose finite element analyzer AERO-S.\textsuperscript{23,24}

Remark 1. The enforcement of the plane stress condition can be done using alternative conjugate pairings, if so desired, by substituting the transformations associated with the preferred stress and strain measures in (9–11).

To complete the description of this multiscale material model, it is noted that for a static analysis or a dynamic analysis using an implicit time-stepping scheme, the consistent constitutive tangent of the plane stress constitutive law, $\frac{\partial S_0^m}{\partial E_0^m}$, is typically required. This quantity is readily obtained using the constitutive tangent of the 3D constitutive law; its precise definition can be found in Reference 13.

### 2.4 Discrete governing equations

Here, the discretized form of the equations governing the multiscale problem of interest are presented, notably including contact at both scales. Specifically,

- At the macroscale, the solution of a dynamic contact problem is sought. The deforming bodies are discretized in space using membrane finite elements and in time using the explicit central difference time-integration scheme. The contact part of the problem is solved using an implicit approach.\textsuperscript{25}
- At the microscale, the solution of static contact problems with prescribed displacement boundary conditions semidiscretized using solid finite elements is sought.

With regards to notation, a distinction is made in this work between unconstrained degrees of freedom (dofs), that is, dofs that are not constrained by any essential boundary condition, and constrained dofs, that is, dofs that are constrained by essential boundary conditions. A matrix or vector defined over the set of unconstrained dofs is not designated by any specific symbol. However, a vector of constrained dofs is designated by the ring symbol as in $\vec{v}$ and a vector defined over the entire set of constrained and unconstrained dofs is designated by the overline symbol as in $\overline{v}$. In other words,

$$\vec{v} = \begin{bmatrix} v \\ \dot{v} \end{bmatrix}.$$  \hfill (14)

It is assumed, without loss of generality that the discrete form of the governing macroscale equations can be written as a differential-algebraic inequality as follows

$$M_0 \dot{u}_0 + f_0^{int} (\dot{u}_{0}^{(n+1)}) + G_0 \left( u_{0}^{(n+2)} \right) \lambda_0^{(n+1)} = f_0^{ext} (t^{(n+1)}),$$  \hfill (15a)

$$g_0 \left( u_{0}^{(n+2)} \right) \geq 0,$$  \hfill (15b)

$$\lambda_0^{(n+1)} \leq 0,$$  \hfill (15c)

$$\lambda_0^{(n+1)T} g_0 \left( u_{0}^{(n+2)} \right) = 0,$$  \hfill (15d)

where $M_0$ is the (diagonal) mass matrix, $f_0^{int}$ and $f_0^{ext}$ are the internal and external force vectors, $u_{0}^{(n+1)}$ and $u_{0}^{(n+1)}$ are the displacements and accelerations at time $t^{(n+1)}$, $g_0$ is the gap, a vector-valued constraint function representing the discretized nonpenetration condition, $G_0$ is the transpose of the constraint Jacobian matrix

$$G_0 = \left[ \frac{\partial g_0}{\partial u_0} \right]^T$$

and $\lambda_0^{(n+1)}$ is a vector of Lagrange multipliers at time $t^{(n+1)}$.  

Remark 2. The evaluation of the proposed multiscale, plane stress constitutive function—which encapsulates the microscale response and its coupling with the macroscale counterpart—is performed during the computation of the internal force vector \( \mathbf{f}_0^{\text{int}} \). Precisely, this computation is carried out in the same fashion as in the case of a conventional material law. In particular, the contribution of each finite element to this computation is determined using an appropriate quadrature rule and the evaluation of the constitutive function at each quadrature point.

Given some initial values \( \mathbf{u}_0^{(n)}, \mathbf{u}_0^{(n+1)}, \) and \( \mathbf{u}_0^{(n)} \) at time \( t^n \), the solution of the above inequality problem at time \( t^{(n+1)} \) is obtained using the following updating procedure:

1. Update the displacement state
   \[
   \mathbf{u}_0^{(n+1)} = \mathbf{u}_0^{(n)} + \Delta t_n \mathbf{u}_0^{(n)} + 0.5 \Delta t_n^2 \mathbf{u}_0^{(n)}.
   \]

2. Update the acceleration and velocity states using the predictor-corrector iterative method
   (a) predictor: \( k = 0 \)
   \[
   \mathbf{u}_0^{(n+1), 0} = \mathbf{M}_0^{-1} \left[ \mathbf{f}_0^{\text{int}} \left( t^{(n+1)} \right) - \mathbf{f}_0^{\text{int}} \left( \mathbf{u}_0^{(n+1)} \right) \right],
   \]
   \[
   \mathbf{u}_0^{(n+1), 0} = \mathbf{u}_0^{(n)} + 0.5 \Delta t_n \left[ \mathbf{u}_0^{(n+1), 0} + \mathbf{u}_0^{(n+1), 0} \right],
   \]
   \[
   \mathbf{u}_0^{(n+2), 0} = \mathbf{u}_0^{(n+1)} + \Delta t_n \mathbf{u}_0^{(n+1), 0} + 0.5 \Delta t_n^2 \mathbf{u}_0^{(n+1), 0}.
   \]

   (b) corrector iterations: \( k = 1, \ldots \)
   \[
   \mathbf{u}_0^{(n+1), k} = \mathbf{u}_0^{(n+1), k-1} + \Delta \mathbf{u}_0^{(n+1), k},
   \]
   \[
   \mathbf{u}_0^{(n+1), k} = \mathbf{u}_0^{(n+1), k-1} + 0.5 \Delta t_n \Delta \mathbf{u}_0^{(n+1), k},
   \]
   \[
   \mathbf{u}_0^{(n+2), k} = \mathbf{u}_0^{(n+2), k-1} + 0.5 \left[ \Delta t_n \Delta t_{n+1} + \Delta t_n^2 \right] \Delta \mathbf{u}_0^{(n+1), k}.
   \]

At each corrector iteration, the acceleration increment \( \Delta \mathbf{u}_0^{(n+1), k} \) is obtained by linearizing the gap function \( g_0 \) and solving the linearized subproblem

\[
\mathbf{M}_0 \Delta \mathbf{u}_0^{(n+1), k} + \mathbf{G}_0 \begin{pmatrix} \mathbf{u}_0^{(n+2), k-1} \end{pmatrix} \Delta \mathbf{u}_0^{(n+1), k} = -\mathbf{f}_0^{k-1}, \tag{16a}
\]
\[
\mathbf{G}_0 \begin{pmatrix} \mathbf{u}_0^{(n+2), k-1} \end{pmatrix} \Delta \mathbf{u}_0^{(n+1), k} \succeq -\mathbf{g}_0^{k-1}, \tag{16b}
\]
\[
\lambda_0^{(n+1), k} \leq 0, \tag{16c}
\]
\[
\mathbf{G}_0 \begin{pmatrix} \mathbf{u}_0^{(n+2), k-1} \end{pmatrix} \Delta \mathbf{u}_0^{(n+1), k} + \mathbf{g}_0^{k-1} = 0, \tag{16d}
\]

where
\[
\mathbf{f}_0^{k-1} = \mathbf{M}_0 \left[ \mathbf{u}_0^{(n+1), k-1} - \mathbf{u}_0^{(n+1), 0} \right],
\]
\[
\mathbf{g}_0^{k-1} = \frac{2}{\Delta t_n \Delta t_{n+1} + \Delta t_n^2} - \mathbf{g}_0 \left( \mathbf{u}_0^{(n+2), k-1} \right).
\]

The corrector subproblem (16) has the form of a quadratic program: it can be solved by the primal-dual active set method.\textsuperscript{26,27}

1. Initialize \( \Delta \mathbf{u}_0^{(n+1), k}, \lambda_0^{(n+1), k} \)
2. Iterate
   • Choose active set:
   \[
   \mathcal{A} = \left\{ i : \lambda_0^{(n+1), k} > 0 \wedge \mathbf{G}_0 \begin{pmatrix} \mathbf{u}_0^{(n+2), k-1} \end{pmatrix} \Delta \mathbf{u}_0^{(n+1), k} + \mathbf{g}_0^{k-1} < 0 \right\}.
   \]
• Set the inactive Lagrange multipliers to zero:

\[
\left[ \lambda_0^{(n+1), k} \right]_i = 0 \quad \forall \ i \in A.
\]

• Solve for \( \Delta \bar{u}_0^{(n+1), k} \) and the active Lagrange multipliers:

\[
M_0 \Delta \bar{u}_0^{(n+1), k} + G_0^{A} \left( \bar{u}_0^{(n+2), k-1} \right) \lambda_0^{(n+1), A, k} = -\bar{f}_0^k, \tag{17a}
\]

\[
G_0^{A} \left( \bar{u}_0^{(n+2), k-1} \right)^T \Delta \bar{u}_0^{(n+1), k} = -\bar{g}_0^{A, k-1}, \tag{17b}
\]

where the superscript \( A \) applied to a vector designates its restriction to the active set. Similarly, the superscript \( A \) applied to a matrix designates its column-wise restriction to the active set.

The active set method subproblem (17) is a linear saddle-point system. To solve for the active Lagrange multipliers, we first eliminate \( \Delta \bar{u}_0^{(n+1), k} \) and then solve the remaining Schur complement system

\[
\begin{bmatrix}
G_0^{AT} M_0^{-1} G_0^{A} \\
M_0 + \mu G_0^{A} G_0^{AT}
\end{bmatrix} \lambda_0^{A} = \bar{g}_0^{A} - G_0^{AT} M_0^{-1} \bar{f}_0. \tag{18}
\]

To simplify notation, the superscripts denoting the time-step index and predictor-corrector iteration have been omitted here but can be inferred from (17). After solving (18) for the Lagrange multipliers, the acceleration increment can be obtained from (17a).

If \( G_0^{A} \) is rank-deficient, then the active set iterations may not converge. In this case, a penalty parameter \( (\mu) \) can be used to regularize the system, leading to perturbed systems of the form\(^{28}\)

\[
\begin{bmatrix}
G_0^{AT} M_0^{-1} G_0^{A} + \frac{1}{\mu} I \\
M_0 + \mu G_0^{A} G_0^{AT}
\end{bmatrix} \lambda_0^{A} = \bar{g}_0^{A} - G_0^{AT} M_0^{-1} \bar{f}_0
\]

or equivalently,

\[
\begin{bmatrix}
M_0 + \mu G_0^{A} G_0^{AT}
\end{bmatrix} \Delta \bar{u}_0 = -\bar{f}_0 - \mu G_0^{A} \bar{g}_0^{A}.
\]

This completes the description of the discrete macroscale problem and its solution algorithm. Significantly, each time-step incurs only one evaluation of \( f^{\text{int}} \), which in the context of a multiscale simulation invariably dominates the computational cost of the entire time-step. In order to evaluate this discrete vector of internal forces, the homogenized stress tensor must be computed at each Gauss point of the macroscale finite element model, which in turn involves the iterative solution of the Klinkel–Govindjee plane stress equation with one solution of the discrete microscale governing equation required per iteration. In the presence of contact at the microscale—for example, nonpenetration and sliding of yarns in a woven fabric—the discrete form of the microscale governing equation has a similar form to that of the macroscale (15) but without the time-dependence and associated temporal discretization. The external force term is also identically zero and can be omitted; the problem is instead driven by prescribed values of the constrained dofs and can be described as follows

\[
\begin{align}
\text{\( f_1^{\text{int}} (\bar{u}_1) + G_1 (\bar{u}_1) \lambda_1 = 0, \tag{19a} \)} \\
g_1 (\bar{u}_1) & \geq 0, \tag{19b} \\
\lambda_1 & \leq 0, \tag{19c} \\
\lambda_1^T g_1 (\bar{u}_1) = 0. \tag{19d}
\end{align}
\]

All of the quantities \( F_1, \bar{u}_1, g_1, G_1, \) and \( \lambda_1 \) are microscale counterparts of the corresponding macroscale quantities previously defined. The above problem can be solved in a similar fashion to that of the macroscale problem by solving a series of linearized subproblems of the form

\[
K_1^{\text{m}} \Delta u_1^k + G_1 \left( u_1^{k-1} \right) \lambda_1^k = -f_1^{k-1}. \tag{20a}
\]
\[ G_1 \left( \bar{u}_1^{k-1} \right)^T \Delta u_1^k \geq -\tilde{g}_1^{k-1}, \]  
(20b)

\[ \lambda_1^k \leq 0, \]  
(20c)

\[ \lambda_1^{kT} \left[ G_1 \left( \bar{u}_1^{k-1} \right)^T \Delta u_1^k + \tilde{g}_1^{k-1} \right] = 0, \]  
(20d)

where \( K_{tgt}^1 \) is the microscale tangent stiffness matrix

\[ K_{tgt}^1 = \frac{\partial f_1^{int}}{\partial u_1}. \]

Problem (20) can again be solved by the dual-primal active set method proposed for the corresponding macroscale problem (16), although numerous alternatives exist.

The computational homogenization method described herein provides a very general framework for solving the problem of interest without resorting to any ad hoc approximation. However, without introducing any further approximation, the framework—although amenable to parallel implementation— is impractical for all but the most modest of applications due to its computational complexity. For example, the authors of this article estimate that to simulate the inflation of a parachute using a macroscale model comprising 182,554 nodes and 279,025 triangular membrane elements would require 49,604,444,444 constitutive function evaluations and a total run time of approximately 48 years on 1000 processing units. Hence, a regression-based surrogate modeling methodology is proposed to achieve computational tractability and described below. It is emphasized that this methodology relies exclusively on the general framework presented above to obtain “training data” that can be used to construct a low-dimensional surrogate model.

# 3 | REGRESSION-BASED SURROGATE MICROSCALE MODEL

Here, a methodology featuring a regression-based surrogate model is presented for dramatically accelerating the solution of nonlinear dynamic multiscale problems modeled using the multiscale formulation based on the concept of a locally attached microstructure overviewed above. The methodology features a novel training strategy based on the concept of a coupon test analogy.

Regression-based surrogate models can be loosely classified as follows:

1. Models whose forms are determined a priori and whose parameters are fitted to available data. Examples of such models are:
   - The St. Venant–Kirchhoff hyperelastic model, a two-parameter model characterized by a linear relationship between the second Piola–Kirchhoff stress and the Green–Lagrange strain.
   - Hyperviscoelastic models incorporating a hyperelastic model such as the St. Venant–Kirchhoff model, combined with a viscoelastic component based on a Prony series.

2. Models whose forms are not entirely predetermined but which are rather discovered, at least in part, by a regression/fitting process. An example of such a model is an artificial NN-based model. In this case, certain characteristics of the model may still be specified a priori, such as the number of hidden layers and the functional form of the activation function.

A training strategy, that is, a procedure for sampling a parameter space such as \( \left\{ E_0^{(11)}, E_0^{(12)}, E_0^{(22)} \right\} \) and collecting conjugate stress and strain data for the purpose of constructing a regression-based constitutive model is proposed here. The strategy employs a small coupon of the material of interest that is semidiscretized at the macroscale level by a single membrane element. It is emphasized that due to the overwhelming cost of an entire multiscale simulation based on a high-dimensional macroscale model, it is not practical to collect data specifically customized to a target application, as is sometimes done to train projection-based reduced-order models (PROMs). However, to some extent, the range of strains to which the coupon model is subjected during the training can be customized, for example, to target applications with small, medium, or large deformations. Due to the small size of the coupon macroscale model, it is feasible to collect data...
FIGURE 3  Microscale training displacement solution snapshots colored by von Mises stress contours for a few points sampled in the parameter space $\{E_{0}^{(11)}, E_{0}^{(12)}, E_{0}^{(22)}\}$: uniaxial tension $\{0.16, 0.0, 0.0\}$ (top left); biaxial tension $\{0.15, 0.15, 0.0\}$ (top right); uniaxial compression $\{-0.06, 0.0, 0.0\}$ (bottom left); and shear $\{0.0, 0.0, 0.125\}$ (bottom right).

comprehensively sampled on a regular grid within a 3D parameter space such as $\{E_{0}^{(11)}, E_{0}^{(12)}, E_{0}^{(22)}\}$ (recall that the prescribed microscale boundary displacements are obtained from a mapping from the in-plane components of the macroscale symmetric Green–Lagrange strain tensor). For training purposes, the macroscale strain can be indirectly specified by prescribing displacements on the boundary of the macroscale model of the coupon. More importantly, the generation and collection of multiscale data can be accelerated using the nonlinear PMOR approach described in Appendix A.

Figure 3 shows for several points sampled in the parameter space $\{E_{0}^{(11)}, E_{0}^{(12)}, E_{0}^{(22)}\}$, deformed configurations and corresponding von Mises stress contours obtained during a training performed for the application described in the following section.

4  APPLICATIONS

In this section, the computational homogenization framework proposed in this article for the nonlinear dynamic analysis of membranes is demonstrated for the coupled, multiscale, fluid–structure simulation of the supersonic inflation of a DGB parachute canopy made of a woven fabric during a Mars landing event. For this purpose, three regression-based surrogate models are considered to accelerate the microscale computations:

- A linear regression model defined by

$$\begin{bmatrix} S_{0}^{(11)} & S_{0}^{(22)} & S_{0}^{(12)} \end{bmatrix}^T = C_{l}\begin{bmatrix} E_{0}^{(11)} & E_{0}^{(12)} & 2E_{0}^{(12)} \end{bmatrix}^T$$  \hspace{1cm} (21)

and the symmetric matrix $C_{l}\in \mathbb{R}^3$.

- A quadratic regression model defined by

$$\begin{bmatrix} S_{0}^{(11)} & S_{0}^{(22)} & S_{0}^{(12)} \end{bmatrix}^T = C_{q}\begin{bmatrix} E_{0}^{(11)} & E_{0}^{(12)} & 2E_{0}^{(12)} & E_{0}^{(11)2} & E_{0}^{(12)2} & (2E_{0}^{(12)})^2 & 2E_{0}^{(12)}E_{0}^{(22)} & 2E_{0}^{(12)}E_{0}^{(22)} & E_{0}^{(11)}E_{0}^{(22)} \end{bmatrix}^T$$  \hspace{1cm} (22)

and $C_{q}\in \mathbb{R}^{3\times9}$. 

A linear model with a NN-based correction of the form
\[
\begin{bmatrix}
S^{(11)}_0 & S^{(22)}_0 & S^{(12)}_0
\end{bmatrix}^T = C_l \begin{bmatrix}
E^{(11)}_0 & E^{(22)}_0 & 2E^{(12)}_0
\end{bmatrix}^T + N \left( \begin{bmatrix}
E^{(11)}_0 & E^{(22)}_0 & 2E^{(12)}_0
\end{bmatrix}^T \right),
\]
(23)

where \( N \) denotes a fully connected NN designed for correcting the linear model by mapping the strain to a stress correction.

First, it will be shown that the NN-based model outlined above outperforms the two other surrogate models in terms of training and test errors. For this reason, only this model will be considered in Section 4.2 to describe the behavior of the canopy material in the simulation of the supersonic parachute inflation of the DGB parachute.

### 4.1 Artificial NNs

#### 4.1.1 Data generation

Different stress–strain tensor data pairs \( \left( S^{(i)}_0, E^{(i)}_0 \right), i = 1, \ldots, N \) are generated by performing a numerical coupon test \( N \) times, where \( N \) is the number of training data points. Each coupon test is graphically depicted in Figure 4, where the right triangle geometry representing a single finite element has two sides of length equal to 1 m. The displacements of the right angle node and all out-of-plane displacements are constrained to be zero; prescribed in-plane displacements are applied to the two other nodes to generate a specified target strain field.

Each microscale problem (see Section 2.2) is solved at the single Gauss quadrature point located at the center of the right triangle using the nonlinear PMOR approach described in Appendix A, which accelerates the generation of the homogenized strain and stress pairs.

The strain field at the macroscale (fabric coupon) level \( \begin{bmatrix}
E^{(11)}_0 & E^{(22)}_0 & 2E^{(12)}_0
\end{bmatrix}^T \) is sampled in a cube of extent \([-0.1, 0.25] \times [-0.1, 0.25] \times [-0.1, 0.25] \). Here, the range of strains is customized to match the application of interest, specifically, the supersonic inflation of a DGB parachute discussed in Section 4.2. The cube is uniformly sampled using 17 equidistant points in each strain component, which accounts for a total of 4913 training data points. Each training data point requires the solution of the discrete equations (15) governing the multiscale coupon problem. To facilitate the implementation of the sampling procedure, each data point is generated as a time-step of a single multiscale simulation in which the prescribed boundary conditions are varied in time along the trajectory shown in Figure 5. Each time-step can be interpreted as an independent static simulation; alternatively, each line segment of the trajectory can be interpreted as being associated with the numerical counterpart of a single physical coupon test in which two strain components are held fixed, while the third is varied. Crucially, the converged solution of the microstructure problem at the previous data point is used to initialize Newton’s method at the next data point. In total, the multiscale data generation procedure equipped with the nonlinear PMOR approach described in Appendix A consumes about 40 h wall-clock time on a single core.

To validate the surrogate models, another set of 4913 test data points is also generated by shifting the aforementioned trajectory.

![FIGURE 4](image.png) Schematic of the numerical coupon test with prescribed displacements (strain fields)
4.1.2 | Training

For the sake of computational efficiency, the number of hidden layers for the NN introduced in (23) is set to 1. Both tanh and ReLU activation functions are considered.

A good definition of the loss function can be

$$\sum_{i=1}^{N} ||S_0^{(i)} - M(E_0^{(i)})||^2 + \lambda ||\theta||^2,$$

(24)

where $M$ represents the surrogate model, $\theta$ denotes its hyperparameters, and $L_2$ regularization is added using the parameter $\lambda$, which is set here to $\lambda = 10^{-4}$. For a woven fabric material, however, the shear stress $S_0^{(12)}$ is typically several orders of magnitude smaller than the axial stresses $S_0^{(11)}$ and $S_0^{(22)}$ (see Figure 6). For this reason, the alternative weighted loss function

$$\sum_{i=1}^{N} \left( S_0^{(11),(i)} - M^{(11)}(E_0^{(i)}) \right)^2 + \left( S_0^{(22),(i)} - M^{(22)}(E_0^{(i)}) \right)^2 + w \left( S_0^{(12),(i)} - M^{(12)}(E_0^{(i)}) \right)^2 + \lambda ||\theta||^2,$$

(25)

where $w$ is a weighting constant is more appropriate. For a given training dataset, the value of $w$ can be automatically deduced from the application to this set of a simple scaling procedure.

In total, six regression-based surrogate models are considered:

- The linear model (21).
- The quadratic model (22).
- The NN-tanh model, which is model (23) where the NN is equipped with six tanh neurons and trained with the nonweighted loss function (24) (for comparison only).
- The NN-ReLU model, which is model (23) where the NN is equipped with six ReLU neurons and trained with the nonweighted loss function (24) (for comparison only).
- The NN-ReLU-W6 model, which is model (23) where the NN is equipped with six ReLU neurons and trained with the weighted loss function (25).
- The NN-ReLU-W20 model, which is model (23) where the NN is equipped with 20 ReLU neurons and trained with the weighted loss function (25).
FIGURE 6 Stress–strain curves predicted using the exact (reference) and surrogate microscale models (training dataset)
Both linear and quadratic models are trained without regularization ($\lambda = 0$). All NNs are trained using the limited-memory BFGS (L-BFGS-B) method with regularization ($\lambda = 10^{-4}$). In all cases, the line search routine provided in Reference 30 is used: it attempts to enforce the Wolfe conditions using a sequence of polynomial interpolations. Note that the BFGS algorithm is appropriate in this case because the datasets are relatively small; for larger datasets, the stochastic gradient descent method is suggested for training.

For both training and test datasets, the relative total errors and relative errors for each stress component are reported in Table 1. The reader can observe that the best if not all NN-based surrogate models lead to relative errors that are one order of magnitude smaller than those of the linear and quadratic regression models. The training/test data and all obtained predictions are also plotted in Figures 6 and 7, for each of the component-wise relations $S_{0}^{(11)} - E_{0}^{(11)}$, $S_{0}^{(22)} - E_{0}^{(22)}$, and $S_{0}^{(12)} - 2E_{0}^{(12)}$. The data shows that the woven fabric material is flexible with respect to shear and compression. In particular, the shear stresses are found to be two orders of magnitude smaller than the axial stresses under similar strains. Furthermore, the $S_{0}^{(11)} - E_{0}^{(11)}$ and $S_{0}^{(22)} - E_{0}^{(22)}$ curves are “flat” when the woven fabric is compressed, indicating that it does not take compression; their slopes suddenly change at zero and remain constant in the stretching regime. Due to these features (especially slope discontinuity in the stress–strain relations), the NN-based regression models deliver better approximations and therefore outperform the linear and quadratic regression models.

It is worth mentioning that the shear stress is relatively small but highly nonlinear. NNs trained with the nonweighted loss function focus mainly on the axial stresses and therefore fail to capture the nonlinearity in the shear stress (see Figures 6 and 7). On the other hand, the NNs trained with the weighted loss function deliver a reasonable accuracy for shear stress prediction. Moreover, increasing the number of neurons is shown to improve accuracy (see Table 1).

Regarding computational cost, the number of operations performed during a single evaluation of a surrogate model is $O(15)$ for the linear regression model, $O(51)$ for the quadratic one and $O(13 \times \#\text{neurons} + 15)$ for the NN-based regression models. For an explicit time-integration, computational cost is typically dominated by that associated with constitutive function evaluations. Hence, in the context an explicit time-integrator, adopting an NN-based surrogate microscale model instead of a linear one may increase the cost of a multiscale simulation by up to a factor of 6 in the case of 6 neurons, or 18 in the case of 20 neurons. However, in the presence of an implicit time-integration scheme, the computational cost associated with an equation solver is typically such that the additional cost incurred by an NN-based surrogate model over a linear one would be substantially less. In any case, the realistic, coupled, multiscale, fluid–structure simulation discussed in the following section is not computationally tractable without a surrogate microscale model. For this reason, and because of its superior accuracy, the surrogate model NN-RelU-W20 is chosen for performing this simulation—and model NN-ReLU is considered only for the purpose of performing a comparison.

### 4.2 Supersonic inflation of a DGB parachute for Mars landing

Finally, the proposed computational homogenization framework is equipped with the previously trained NN-ReLU models and applied here to simulate the supersonic inflation dynamics of a NASA DGB parachute system in the low-density, low-pressure, supersonic Martian atmosphere. While such a coupled, multiscale, fluid–structure simulation is crucial to the understanding of the effects of a woven fabric material on the performance of a parachute during the deceleration

| TABLE 1 | Relative total errors and relative errors for each stress component and each considered regression-based surrogate model, for both training and test datasets |
|----------|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|          | Linear | Quadratic | NN-tanh | NN-ReLU | NN-ReLU-W6 | NN-ReLU-W20 |
| Training set | 19.5% | 10.4% | 1.06% | 1.05% | 2.13% | 0.97% |
| $S_{xx}$ | 19.5% | 10.4% | 0.93% | 1.10% | 1.63% | 0.88% |
| $S_{yy}$ | 19.5% | 10.4% | 1.24% | 0.97% | 2.74% | 1.08% |
| $S_{xy}$ | 42.4% | 38.6% | 39.1% | 38.3% | 19.2% | 8.20% |
| Test set | 14.0% | 9.87% | 0.91% | 1.03% | 2.40% | 0.95% |
| $S_{xx}$ | 14.0% | 9.87% | 0.91% | 1.17% | 1.83% | 1.00% |
| $S_{yy}$ | 14.0% | 9.87% | 0.91% | 0.87% | 2.83% | 0.90% |
| $S_{xy}$ | 43.0% | 34.81% | 38.7% | 39.2% | 21.6% | 9.37% |
FIGURE 7  Stress–strain curves predicted using the exact (reference) and surrogate microscale models (test dataset)
process, its main purpose here is twofold: (1) demonstrate the computational tractability of the proposed computational framework for a realistic application; and (2) to validate (partially) it using flight data from the landing on Mars of NASA’s rover Curiosity.

Specifically, the DGB parachute system considered here is that which successfully deployed in 2012 for the Mars landing of Curiosity (see Figure 8-left). This aerodynamic decelerator system consists of three main components:\(^3\!\!2\)

- The canopy, which is made of a woven nylon fabric material (see Figure 1).
- The suspension lines, which are made of Technora T221 braided cords.
- The reentry vehicle.

Its geometric and material properties are listed in Table 2.

The simulation discussed herein starts from the line stretch stage where the suspension line subsystem is deployed, the canopy is folded (see Figure 8-right) and the entire system is prestressed by the folding pattern.\(^3\!\!4\) The incoming supersonic flow is at the state defined by the free-stream conditions \(M_\infty = 1.8\), \(\rho_\infty = 0.0067\ \text{kg m}^{-3}\) and \(p_\infty = 260\ \text{Pa}\).

Since the Martian atmosphere is mainly composed of carbon dioxide, the viscosity of this gas is modeled using Sutherland’s viscosity law with the constant \(\mu_0 = 1.57 \times 10^{-6}\ \text{kg m}^{-1}\text{s}^{-1}\) and the reference temperature \(T_0 = 240\ \text{K}\). The Reynolds number based on the canopy diameter is \(4.06 \times 10^6\). Hence, at the beginning of the simulation, the flow is assumed to have transitioned to the turbulent regime—which is modeled here using Vreman’s eddy viscosity subgrid-scale model for turbulent shear flow\(^3\!\!7\) equipped with the model constant \(C_s = 0.07\).

Given the expected large motions and deformations of the parachute system during its inflation, the flow computations are performed using the large eddy simulation capability of the AERO-F flow solver\(^2\!\!3,2\!\!4\) and its embedded

---

**Figure 8** Simulation of the supersonic parachute inflation dynamics of a NASA disk-gap-band system: System configuration (left); and embedding computational fluid domain and embedded initial folded configuration (right)

**Table 2** Simulation of the supersonic parachute inflation dynamics of a NASA disk-gap-band system: Geometric and material properties of the system\(^3\!\!2,3\!\!3,3\!\!6\)

| Component       | Parameter | Description                  | Value         |
|-----------------|-----------|-------------------------------|---------------|
| Canopy          | \(D\)     | diameter                      | 15.447 m      |
|                 | \(t\)     | thickness                     | \(7.6 \times 10^{-5}\) m |
|                 | \(E\)     | microscale yarn Young’s modulus | 3497 MPa\(^a\) |
|                 | \(\nu\)   | microscale yarn Poisson’s ratio | 0.2           |
|                 | \(\rho^C\) | density                       | 1154.25 kg m\(^{-3}\) |
|                 | \(\alpha\) | porosity                      | 0.08          |
| Suspension lines| \(L\)     | length                        | 36.56 m       |
|                 | \(D\)     | diameter                      | \(3.175 \times 10^{-3}\) m |
|                 | \(E\)     | Young’s modulus               | 29.5 GPa      |
|                 | \(\rho^{SL}\) | density                       | 1154.25 kg m\(^{-3}\) |

\(^a\) The microscale yarn Young’s modulus is roughly estimated from the Young’s modulus of the macroscale nylon.
boundary method for fluid–structure interaction known as the finite volume method with exact two-material Riemann problems.\textsuperscript{33,38-42} AERO-F incorporates a parallel adaptive mesh refinement (AMR) capability based on newest vertex bisection,\textsuperscript{43,44} which enables it to capture various interactions between the fluid system, the nonlinear parachute system including its suspension lines and the forebody.

The canopy of the DGB parachute consists of band and disk gores that are represented here by 279,025 geometrically nonlinear membrane elements. The suspension line subsystem contains 80 lines, each of which is discretized by 500 geometrically nonlinear beam elements. The reentry vehicle is modeled as a fixed rigid body; it is embedded, together with the entire aerodynamic decelerator system, in the embedding computational fluid domain (see Figure 8). This domain is a box of size 200 m × 160 m × 160 m. It is initially discretized by a mesh with 2,778,867 nodes and 16,308,672 tetrahedra. During the fluid–structure interaction simulation, AMR is applied to track and resolve the boundary layers and flow

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure9}
\caption{Simulation of the supersonic parachute inflation dynamics of a disk-gap-band system: Time-evolutions of the structural system and Mach number field during deployment.}
\end{figure}
features. For this purpose, the specified characteristic mesh sizes near the reentry vehicle and canopy are 2.5 and 5 cm, respectively; those in the wake and near the shock are set to 10 cm.

Since the canopy is made of a woven nylon fabric with an 8% void fraction, its permeability is modeled using a homogenized porous wall model.45,46 Due to the massive self-contact of the parachute canopy during its dynamic inflation, the explicit central difference time-integration scheme is used to advance in time the semidiscrete state of the structural system. A small amount of Rayleigh damping is applied to stabilize this system.

First, a quasi-steady state of the flow past the folded parachute configuration shown in Figure 8-right is computed assuming that this configuration is rigid and fixed. Using the computed CFD solution and the aforementioned prestressed state of the structural model of the parachute system as initial fluid and structural conditions, respectively, the coupled, multiscale, fluid–structure simulation of the inflation dynamics of the DGB parachute is performed in the time-interval \([0, 0.8]\) s. The length of this time-interval is such that it covers the inflation process as well as a few breathing cycles of the DGB parachute system. While the explicit central difference time-integrator is applied to advancing in time the semidiscrete structural system for the reason mentioned above, the implicit, 3-point backward difference formula scheme is applied to time-integrate the semidiscrete fluid state. The fluid and structural discretizations are coupled using the stability-preserving, second-order, time-accurate, implicit–explicit fluid–structure staggered solution procedure presented in Reference 47 and the fluid–structure coupling time-step is set to \(\Delta t_{FS} = 10^{-5}\) s.

Figure 9 graphically depicts the time-evolutions of the dynamic inflation of the DGB parachute and the flow Mach number around it. The parachute is fully inflated at approximately \(t = 0.24\) s; after this time, it starts the breathing cycles expected from a violent, high-speed, dynamic, inflation process.

Figure 10 reports the time-histories of the total drag force predicted by the coupled, fluid–structure simulations described above. For validation and reference, respectively, this figure includes the measured time-history of the total drag generated by the parachute system during the Mars landing of NASA’s rover Curiosity32 and its counterpart predicted using the same aforementioned simulation but equipped with the classical St. Venant–Kirchhoff model.34 The reader can

![Figure 10](image_url)

**FIGURE 10** Simulation of supersonic the parachute inflation dynamics of a NASA disk-gap-band system: Time-history of the total drag recorded during the Mars landing of Curiosity32 (blue); counterpart time-histories obtained using the coupled, multiscale, fluid–structure simulation equipped with the NN-ReLU (orange) and NN-ReLU-W20 (green) surrogate microscale models; and counterpart time-history obtained using the same simulation but equipped with the St. Venant–Kirchhoff (red) constitutive model.

![Figure 11](image_url)

**FIGURE 11** Simulation of the supersonic parachute inflation dynamics of a NASA disk-gap-band system: Time-histories of the maximum von Mises stresses predicted by the coupled, fluid–structure interaction simulations equipped with the NN-ReLU (orange) surrogate microscale model, the NN-ReLU-W20 (green) counterpart model and the St. Venant–Kirchhoff (red) constitutive model.
Simulation of the supersonic parachute inflation dynamics of a NASA disk-gap-band system: Time-histories of the von Mises stress field predicted by the coupled, fluid–structure interaction simulations equipped with the NN-ReLU (left) surrogate microscale model, the NN-ReLU-W20 counterpart model (middle) and the St. Venant–Kirchhoff (right) constitutive model.
observe that the NN-ReLU and NN-ReLU-W20 surrogate microscale models deliver stable results that are in reasonably good agreement with the flight data. The effect on drag performance of the constitutive relation used to model the woven nylon fabric material is found to be weak.

Figure 11 reports the time-histories of the maximum von Mises stresses—an indicator of material failure—predicted by the aforementioned coupled, fluid–structure interaction simulations. Similar stress results are delivered by the NN-ReLU and NN-ReLU-W20 surrogate microscale models, which indicates that for this application, the shear effect of the woven nylon fabric is not significant. However, the results delivered by the St. Venant–Kirchhoff model show that the flexibility with respect to shearing and compression of the multiscale woven fabric model highlighted in Section 4.1.2 leads to lower von Mises stresses in the parachute breathing cycle, after full inflation. This disparity between the results obtained using the classical St. Venant–Kirchhoff and multiscale models is also highlighted in Figure 12, which depicts the time-evolutions of the entire von Mises stress fields. Although further (experimental) investigation is required to conclude which model is more reliable, this comparison illustrates the potential of a multiscale constitutive model for improving the prediction of material failure.

The combined offline/online computational costs of all coupled, fluid–structure interaction simulations discussed above are reported in Table 3, component-by-component. It is worth mentioning that the estimated simulation time for the direct FE² simulation—based on the multiplication of the number of FE² model evaluations on each CPU and the cost of a single evaluation—is also reported in Table 3: it suggests that the NN-based surrogate microscale models lead to speed up factors of the order of $7 \times 10^4$ (all training costs included), which demonstrates the potential of NN-based surrogate microscale models for constitutive modeling.

5 | CONCLUSIONS

The general framework for computationally tractable, nonlinear, multiscale modeling of membrane, woven fabrics presented in this article is enabled by the coherent exploitation of several key, established pillar methodologies:

- The computational homogenization approach known as finite element squared (FE²) based on the concept of a locally attached microstructure.
- The numerical enforcement of the membrane’s plane stress condition.
- The “discovery” of a surrogate, microscale model such as an artificial NN-based regression model using data generated by many multiscale numerical simulations of the behavior of a small woven fabric coupon.
- The acceleration of the training of the above NN using nonlinear PMOR and hyperreduction.

The proposed computational framework encompasses a cascade of multiscale models, ranging from the highest fidelity (without any surrogate model) to the lowest (linear regression surrogate model). The proposed discovery at the finest scale of a surrogate constitutive model by means of numerical coupon testing is analogous to the experimental
testing procedure used to identify the parameters (e.g., Young’s modulus or Poisson’s ratio) of conventional material models. A highlight of the overall approach is that while experimental data is typically limited to uniaxial tension (occasionally biaxial and/or shear data may also be available), numerical data suffers from no such limitation. Using the concept of numerical coupon testing, an entire parameter space of physically admissible combinations of normal and shear strains can be explored in order to characterize complex and unconventional materials, and support the discovery of a constitutive law. Furthermore, using PMOR equipped with hyperreduction (see Appendix A), the aforementioned exploration of a large parameter space can be performed in a multiscale setting in practical wall-clock time. PMOR continues to be an active and fertile area of research that can be leveraged to extend and improve this framework. In particular, the recent emergence of in-situ training methodologies\(^4\) presents an attractive option to streamline and enhance PMOR utilization by eliminating the conventional and potentially cumbersome offline-online decomposition of computational effort and vulnerabilities associated with extrapolation.

All of the above conclusions are supported in this article by the successful demonstration of the proposed computational framework for the simulation of a supersonic parachute inflation dynamics problem in Martian atmospheric entry conditions, for which flight data is available. For this application, a NN-based surrogate microscale model is constructed and trained in a large parameter space in 40 h wall-clock time using a single computational core. This surrogate model is shown to enable the proposed overall nonlinear multiscale framework to achieve computational tractability. Specifically, the coupled, multiscale, fluid–structure interaction simulation of the supersonic, dynamic inflation process of the parachute and a few of its breathing cycles is completed in about 116 h wall-clock time (less than 5 days) on 576 cores of a Linux cluster (82% of this wall-clock time is consumed by the computation of the turbulent flow). Particularly, the NN-based surrogate microscale model is shown to reduce by almost five orders of magnitude the wall-clock time that would otherwise be required for performing the multiscale structural dynamics computations within the same fluid–structure interaction simulation using only the proposed tailoring of the FE\(^2\) framework to membrane woven fabrics. Equally importantly, the time-history of the total drag force predicted using the proposed computational framework is found to match well its flight-recorded counterpart.

**ACKNOWLEDGMENTS**

Philip Avery, Daniel Huang, Johanna Ehlers, and Charbel Farhat acknowledge partial support by the Jet Propulsion Laboratory (JPL) under Contract JPL-RSA No. 1590208 and partial support by the National Aeronautics and Space Administration (NASA) under Early Stage Innovations (ESI) Grant NASA-NNX17AD02G. Parts of this work were completed at the JPL, California Institute of Technology, under a contract with NASA. Optical microscope photographs were taken by Cheyenne Hua at the JPL. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of JPL or NASA.

**DATA AVAILABILITY STATEMENT**

The data that support the findings of this study are openly available in the repository at https://github.com/Zhengyu-Huang/Fabric-Data.git.

**ORCID**

Philip Avery https://orcid.org/0000-0002-1310-6790

Charbel Farhat https://orcid.org/0000-0003-2563-8820

**REFERENCES**

1. Gereke T, Döbrich O, Hübner M, Cherif C. Experimental and computational composite textile reinforcement forming: a review. Compos A Appl Sci Manuf. 2013;46:1-10.
2. Zahr MJ, Avery P, Farhat C. A multilevel projection-based model order reduction framework for nonlinear dynamic multiscale problems in structural and solid mechanics. Int J Numer Methods Eng. 2017;112(8):855-881.
3. Yvonnet J, He QC. The reduced model multiscale method (R3M) for the non-linear homogenization of hyperelastic media at finite strains. J Comput Phys. 2007;223(1):341-368.
4. He W, Avery P, Farhat C. In-situ adaptive reduction of nonlinear multiscale structural dynamics models. Int J Numer Methods Eng. 2020;121(22):4971-4988.
5. Knap J, Barton N, Hornung R, Arsenlis A, Becker R, Jefferson D. Adaptive sampling in hierarchical simulation. Int J Numer Methods Eng. 2008;76(4):572-600.
6. Le B, Yvonnet J, He QC. Computational homogenization of nonlinear elastic materials using neural networks. Int J Numer Methods Eng. 2015;104(12):1061-1084.
7. Lu X, Giovanis DG, Yvonnet J, Papadopoulos V, Detrez F, Bai J. A data-driven computational homogenization method based on neural networks for the nonlinear anisotropic electrical response of graphene/polymer nanocomposites. *Comput Mater Sci.* 2019;64(2):307-321.

8. Coenen E, Kouznetsova V, Geers M. Computational homogenization for heterogeneous thin sheets. *Int J Numer Methods Eng.* 2010;83(9-10):1180-1205.

9. Ha-Minh C, Kanit T, Boussu F, Imad A. Numerical multi-scale modeling for textile woven fabric against ballistic impact. *Comput Mater Sci.* 2011;50(7):2172-2184.

10. Larsson R, Landervik M. A stress-resultant shell theory based on multiscale homogenization. *Comput Methods Appl Mech Eng.* 2013;263:1-11.

11. Miehe C, Schröder J, Schotte J. Computational homogenization analysis in finite plasticity simulation of texture development in polycrystalline materials. *Comput Methods Appl Mech Eng.* 1999;171(3-4):387-418.

12. Simo J, Govindjee S. Exact closed-form solution of the return mapping algorithm in plane stress elasto-viscoplasticity. *Eng Comput.* 1988;5(3):254-258.

13. Klinkel S, Govindjee S. Using finite strain 3D-material models in beam and shell elements. *Eng Comput.* 2002;19(3):254-271.

14. Smit R, Brekelmans W, Meijer H. Prediction of the mechanical behavior of nonlinear heterogeneous systems by multi-level finite element modeling. *Comput Methods Appl Mech Eng.* 1998;155(1-2):181-192.

15. Feyel F, Chaboche JL. FE2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre SiC/Ti composite materials. *Comput Methods Appl Mech Eng.* 2000;183(3-4):309-330.

16. Kouznetsova V, Brekelmans W, Baajens F. An approach to micro-macro modeling of heterogeneous materials. *Comput Mech.* 2001;27(3):37-48.

17. Melbo H, Kvamsdal T. Goal oriented error estimators for Stokes equations based on variationally consistent postprocessing. *Comput Methods Appl Mech Eng.* 2003;192(5-6):613-633.

18. Ciarlet PG. *Mathematical Elasticity: Volume I: Three-Dimensional Elasticity.* Amsterdam, Netherlands: North-Holland; 1988.

19. Yvonnet J, Monteiro E, He QC. Computational homogenization method and reduced database model for hyperelastic heterogeneous structures. *Int J Multisc Comput Eng.* 2011;3(1):201-225.

20. Ogden RW. Nonlinear elasticity, anisotropy, material stability and residual stresses in soft tissue. *Biomechanics of Soft Tissue in Cardiovascular Systems.* New York, NY: Springer; 2003:65-108.

21. De Borst R. The zero-normal-stress condition in plane-stress and shell elastoplasticity. *Commun Appl Numer Methods.* 1991;7(1):29-33.

22. Dvorkin EN, Pantuso D, Repetto EA. A formulation of the MITC4 shell element for finite strain Elasto-plastic analysis. *Comput Methods Appl Mech Eng.* 1995;125(1-4):17-40.

23. Geuzaine P, Brown G, Harris C, Farhat C. Aeroelastic dynamic analysis of a full F-16 configuration for various flight conditions. *AIAA J.* 2003;41(3):363-371.

24. Farhat C, Geuzaine P, Brown G. Application of a three-field nonlinear fluid–structure formulation to the prediction of the aeroelastic parameters of an F-16 fighter. *Comput Fluids.* 2003;32(1):3-29.

25. Salveson MW, Taylor RL. *Solution of Dynamic Contact Problems by Implicit/Explicit Methods.* Rept. UCRL-CR-125780. Livermore, CA: Lawrence Livermore National Laboratory; 1996.

26. Hintermüller M, Ito K, Kunisch K. The primal-dual active set strategy as a semismooth Newton method. *SIAM J Optim.* 2002;13(3):865-888.

27. Hübner S, Wohlmuth BI. A primal–dual active set strategy for non-linear multibody contact problems. *Comput Methods Appl Mech Eng.* 1999;194(27-29):3147-3166.

28. Hintermüller M, Kunisch K. Path-following methods for a class of constrained minimization problems in function space. *SIAM J Optim.* 2006;17(1):159-187.

29. Byrd RH, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. *SIAM J Sci Comput.* 1995;16(5):1190-1208.

30. Moré JJ, Thuente DJ. Line search algorithms with guaranteed sufficient decrease. *ACM Trans Math Softw.* 1994;20(3):286-307.

31. Huang DZ, Xu K, Farhat C, Darve E. Learning constitutive relations from indirect observations using deep neural networks. *J Comput Phys.* 2020;416.

32. Cruz JR, Way DW, Shidner JD, Davis JS, Adams DS, Kipp DM. Reconstruction of the Mars science laboratory parachute performance. *J Spacecr Rocket.* 2014;51(4):1185-1196.

33. Huang DZ, Avery P, Farhat C. An embedded boundary approach for resolving the contribution of cable subsystems to fully coupled fluid-structure interaction. *Int J Numer Methods Eng.* 2020.

34. Huang DZ, Avery P, Farhat C, Rabinovich J, Derkevorkian A, Peterson LD. Modeling, simulation and validation of supersonic parachute inflation dynamics during Mars landing. *AIAA Scitech 2020 Forum.* Reston, VA: American Institute of Aeronautics and Astronautics; 2020:0313.

35. Lin JK, Shook LS, Ware JS, Welch JV. *Flexible Material Systems Testing.* NASA Report CR-2010-216854; Frederica, Delaware: ILC Dover, LP; 2010.

36. Hill JL. *Mechanical Property Determination for Flexible Material Systems* [PhD thesis]. North Ave NW, Atlanta, GA: Georgia Institute of Technology; 2016:30332.

37. Vreman A. An eddy-viscosity subgrid-scale model for turbulent shear flow: algebraic theory and applications. *Phys Fluids.* 2004;16(10):3670-3681.
APPENDIX A. MULTISCALE PROJECTION-BASED MODEL ORDER REDUCTION APPROACH FOR SPACE EXPLORATION

For the sake of completeness, a nonlinear, multiscale, PMOR/hyperreduction framework is presented here for dramatically accelerating the training of a regression-based artificial NN in a large parameter space, in view of using it as a surrogate microscale model. The described approach constitutes a generalization of the framework first presented in Reference 2 to:

- Include a treatment of contact based on the PMOR method originally proposed in Reference 48, which features a nonnegative matrix factorization (NMF) scheme for the construction of a positive reduced-order basis (ROB) for the contact forces.
- Accommodate a novel training strategy based on the concept of a coupon test analogy introduced in Section 3.

For the sake of simplicity and clarity, but without any loss of generality, the proposed PMOR approach is described here only for the microscale level of a two-scale (macro–micro) model. Specifically, proper orthogonal decomposition (POD) is used to construct a PROM at the microscale level and a computational approach based on the energy conserving sampling and weighting (ECSW) method is used to hyperreduce the constructed PROM. Training is performed offline (i.e., a priori) using a small, multiscale coupon model.

How to cite this article: Avery P, Huang DZ, He W, Ehlers J, Derkevorkian A, Farhat C. A computationally tractable framework for nonlinear dynamic multiscale modeling of membrane woven fabrics. *Int J Numer Methods Eng.* 2021;122:2598–2625. https://doi.org/10.1002/nme.6634
A.1 Reduction of the primal unknowns

At the microscale (scale 1), the number of primal dofs $n_1$ of the computational model is reduced by searching for the primal solution $u_1$ of the typical microscale problem in a carefully constructed low-dimensional subspace, that is,

$$u_1 \approx V_1 y_1,$$

(A1)

where $V_1 \in \mathbb{R}^{n_1 \times r_1}$ is a ROB representing a low-dimensional subspace, $y_1 \in \mathbb{R}^{r_1}$ is the vector of generalized coordinates of $u_1$ in this basis and $r_1 \ll n_1$. The ROB is chosen to be orthonormal with respect to the identity matrix, that is,

$$V_1^T V_1 = I.$$

As mentioned above, the ROB is constructed using POD and the method of snapshots. To this end, $m_1$ solution snapshots of (19), $\{u_1^{(1)}, \ldots, u_1^{(m_1)}\}$, are computed at scale 1 for different prescribed boundary displacements and collected in the primal snapshot matrix

$$Y_u = \begin{bmatrix} u_1^{(1)} & \cdots & u_1^{(m_1)} \end{bmatrix}.$$

Then, this matrix is compressed using the singular value decomposition method and $V_1$ is constructed using the first $r_1$ singular vectors of $Y_u$, where $r_1$ is determined from the application of a retention criterion to the energy of the singular values.

From (2), it follows that the constrained dofs of the microscale displacement vector $\hat{u}_1$ lies in a low-dimensional subspace associated with a vector of generalized coordinates identified as the column-wise vectorization of the right stretch strain tensor $U_0 - I$, that is,

$$\hat{u}_1 = \Pi_1 \begin{bmatrix} \hat{X}_1 & 0 & 0 \\ 0 & \hat{X}_1 & 0 \\ 0 & 0 & \hat{X} \end{bmatrix} \text{vec}(U_0 - I) = \hat{V}_1 \hat{y}_1,$$

where $\Pi_1$ is a permutation matrix and $\hat{X}_1$ is a matrix whose three columns represent the $x$, $y$, and $z$ nodal coordinates, respectively, of the constrained nodes located on the boundary of the microscale model. The definition of $\hat{y}_1$ follows from to the notational convention (14). Hence, a basis encompassing both unconstrained and constrained dofs can be represented, up to a permutation, as

$$\hat{V}_1 = \begin{bmatrix} V_1 & 0 \\ 0 & \hat{V}_1 \end{bmatrix}.$$

The dimensionality of the discrete governing equations (19) is reduced at scale 1 by performing a Galerkin projection, that is, substituting (A1) in these equations and projecting the first of them onto the column space of $V_1$. This leads to the PROM

$$V_1^T f_1^\text{int} (\hat{V}_1 \hat{y}_1) + V_1^T G_1 (\hat{V}_1 \hat{y}_1) \lambda_1 = 0,$$

(A2a)

$$g_1 (\hat{V}_1 \hat{y}_1) \geq 0,$$

(A2b)

$$\lambda_1 \leq 0,$$

(A2c)

$$\lambda_1^T g_1 (\hat{V}_1 \hat{y}_1) = 0.$$
substantial bottleneck in the solution of (A2). For this reason, a number of hyperreduction methods have been proposed to overcome this bottleneck introduced by nonlinear terms. For solid mechanics and structural dynamics problems, the ECSW method is preferred due to its desirable structure-preserving and numerical stability properties. However, any other efficient hyperreduction method can be equally used, in principle, to overcome the aforementioned computational bottleneck.

As introduced in Reference 49, the ECSW method amounts to a “mesh reduction and quadrature” algorithm which samples a set of mesh elements \( \mathcal{V}_1' \subset \mathcal{V}_1 \) and attributes to each sampled element \( e \) a positive weight \( \alpha_e^1 > 0 \) such that

\[
\tilde{V}_1^T \tilde{f}_i^1 \left( \tilde{V}_1 \tilde{y}_1 \right) = \sum_{e \in \mathcal{V}_1'} \left( V_1^e \right)^T \tilde{f}^i_{\text{int}} \left( \tilde{V}_1 \tilde{y}_1 \right) \\
\approx \sum_{e \in \mathcal{V}_1'} \alpha_e^1 \left( V_1^e \right)^T \tilde{f}^i_{\text{int}} \left( \tilde{V}_1 \tilde{y}_1 \right) = \tilde{f}^i_{\text{int}} \left( \tilde{y}_1 \right).
\]

In the above expressions, the superscript \( e \) designates the restriction of a global vector or matrix to element \( e \) and the reduced mesh \( \mathcal{V}_1' \) (\( |\mathcal{V}_1'| \ll |\mathcal{V}_1| \)) can be computed using Lawson and Hanson’s nonnegative least squares algorithm, or an alternative L1 minimization algorithm, in a training step that seeks to minimize the size of \( \mathcal{V}_1' \) while maintaining an acceptable approximation error for the ensemble of the training data.

In addition to achieving a computational complexity that scales with the size \( n_1 \) of the PROM only in the computation of the components of the internal force vector corresponding to unconstrained dofs, ECSW and its reduced mesh ensure that scale transmission is performed efficiently, that is, without any operation whose computational complexity scales of the componentsofthe internal force vector corresponding to unconstrained dofs, ECSW and its reduced mesh ensure that scale transmission is performed efficiently, that is, without any operation whose computational complexity scales with \( |\mathcal{V}_1| \). This is evident in the transmission to finer scales, where \( \tilde{u}_1^e \) is required for each \( e \in \mathcal{V}_1' \); as for transmission to coarser scales, the homogenized macroscopic unsymmetric Biot stress tensor is approximated as

\[
\text{vec}(B_0) \approx \frac{1}{|\mathcal{V}_1'|} \tilde{f}_i \left( \tilde{y}_1 \right),
\]

where \( \mathcal{V}_1' \subset \mathcal{V}_1 \) denotes the subset of surface elements contained in the reduced mesh \( \mathcal{V}_1' \) and \( \tilde{f}_i \) is in general the restriction of the total vector of reduced forces—both internal and contact—to the constrained generalized coordinates, that is,

\[
\tilde{f}_i = f^i_{\text{int}} \left( \tilde{y}_1 \right) + V_1^T \tilde{G}_1 \left( \tilde{V}_1 \tilde{y}_1 \right) \lambda_1.
\]

In some cases though, the mesh of the microscale model can be constructed in such a way that the contact forces will contribute nothing to this quantity. This requires maintaining a separation of at least one element between the contact surface and the boundary of the mesh.

A.2 Reduction of the dual unknowns

At the microscale (here, scale 1), the number of dual dofs \( m_1 \) of the computational model can also be reduced by searching for the dual solution \( \lambda_1 \) of the problem of interest in another carefully constructed low-dimensional subspace, that is,

\[
\lambda_1 \approx W_1 z_1, \tag{A3}
\]

where \( W_1 \in \mathbb{R}^{n_1 \times r_1} \) is a dual ROB representing a low-dimensional subspace, \( z_1 \in \mathbb{R}^{r_1^1} \) is the vector of generalized coordinates of \( \lambda_1 \) in this basis and \( r_1^1 \ll n_1^1 \). The dual ROB is chosen such that it has no negative entry in any of its vectors.

In this work, the dual ROB is constructed using NMF. Specifically, \( m_1 \) solution snapshots of (19), \( \{ \lambda^{(1)}, \ldots, \lambda^{(m_1)} \} \), are computed at scale 1 for different prescribed boundary displacements and collected in the dual snapshot matrix

\[
Y_1^d = \begin{bmatrix} \lambda_1^{(1)} & \cdots & \lambda_1^{(m_1)} \end{bmatrix}
\]

and the dual ROB \( W_1 \) is constructed from the left factor of the NMF of \( Y_1^d \).

The dimensionality of the reduced governing equations (A2) is further reduced at scale 1 by substituting (A3) and projecting the gap function \( g_1 \) onto the column space of \( W_1 \). Hyperreduction of the internal force is also performed. This leads to the PROM.
Typically, the evaluations of the gap function and its Jacobian do not require a reconstruction of the full state but only its restriction to the contact surface. Furthermore, Galerkin projection of the contact force term can be optimized by accounting for the sparsity of the Jacobian. Specifically, only the row-wise restriction of $G_1$ to the contact surface is nonzero. Nevertheless, these evaluations may still incur a substantial computational cost. In principal, hyperreduction can be applied to further accelerate the evaluation of the reduced gap function and its Jacobian. This is an active topic of research but is not employed in the present work. However, it is noted that in the case of linear constraints, the proposed reduction of the dual variables leads to terms involving reduced-order matrices that are precomputable and as such does not generate any bottleneck in the online solution of the reduced-order discrete microscale equations. Consequently, just like in the case of any other linear terms, the efficient processing of such terms does not require any hyperreduction.