MD-inferred neural network monoclinic finite-strain hyperelasticity models for $\beta$-HMX: Sobolev training and validation against physical constraints

Nikolaos N. Vlassis, Puhan Zhao, Ran Ma, Tommy Sewell, WaiChing Sun

Abstract

We present a machine learning framework to train and validate neural networks to predict the anisotropic elastic response of the monoclinic organic molecular crystal $\beta$-HMX in the geometrical nonlinear regime. A filtered molecular dynamic (MD) simulations database is used to train the neural networks with a Sobolev norm that uses the stress measure and a reference configuration to deduce the elastic stored energy functional. To improve the accuracy of the elasticity tangent predictions originating from the learned stored energy, a transfer learning technique is used to introduce additional tangential constraints from the data while necessary conditions (e.g. strong ellipticity, crystallographic symmetry) for the correctness of the model are either introduced as additional physical constraints or incorporated in the validation tests. Assessment of the neural networks is based on (1) the accuracy with which they reproduce the bottom-line constitutive responses predicted by MD, (2) detailed examination of their stability and uniqueness, and (3) admissibility of the predicted responses with respect to continuum mechanics theory in the finite-deformation regime. We compare the neural networks’ training efficiency under different Sobolev constraints and assess the models’ accuracy and robustness against MD benchmarks for $\beta$-HMX.

1 Introduction

Plastic-bonded explosives (PBXs) are highly filled polymer composites in which crystallites of one or more energetic constituents are held together by a continuous polymeric binder phase. The filler (i.e., explosive) mass fraction is typically 90%-95% and typically exhibits a wide range of crystallite sizes, spanning several orders of magnitude up to a maximum of a few hundred microns. Detonation initiation in PBXs is often achieved by transmitting a mechanical shock wave into the explosive charge. Shock passage leads to an abrupt increase in stress, strain, and temperature in the material. In thermodynamic terms, the magnitude of the increase of these properties is given by the Hugoniot jump relations, which yield the locus of thermodynamic states immediately behind the shock discontinuity as a function of the input shock strength (with a parametric dependence on the initial thermodynamic state of the material). However, except in the case of very strong shocks, the stresses and temperatures achieved due to bulk hydrodynamic heating in ‘perfect’ crystal are insufficient to lead to prompt ignition of chemistry. Rather, it is thought that additional energy localization mechanisms—such as pore collapse, shear banding, and interfacial debonding and subsequent frictional heating—in the microstructure of the PBX are required to achieve the necessary local thermodynamic states required for rapid, sustained chemistry. These regions of locally high temperature, stress, and...
strain rate are known as *hot spots* (Bowden et al., 1985). If a given hot spot is sufficiently intense, chemistry will commence. Although the initial chemical events, so-called *primary reactions*, are typically endothermic, subsequent *secondary reactions* will follow leading to large, localized heat release and formation of small-molecule products. This results in thermal and stress pulses that propagate into the surrounding material. If the spatial density of such hot spots in a sample is sufficiently high, interactions among them will lead to accelerating chemistry culminating in detonation initiation.

The elastic properties of the constituents in a PBX play an important role in determining the states on the Hugoniot locus. The most obvious connection is their appearance in the reactant equation of state (EOS). For a useful summary, see Hooks et al. (2015). The isotropic EOS can be built around the isothermal compression curve, typically by fitting $V = V(P)$ to the 3rd-order Birch-Murnaghan (B-M) equation of state or some other convenient functional form at room temperature or zero kelvin. For the B-M EOS, the fitting variables are the bulk modulus $K$ and the initial pressure derivative $K'$. More sophisticated models account for crystal elastic anisotropy by incorporating the full elastic tensor. The advantage is a higher fidelity description of the elastic response, but doing so for a material under shock conditions requires knowledge of the pressure and temperature dependence of the elastic coefficients, which in most cases is only available from simulations (Pereverzev and Sewell, 2020). Furthermore, the possibility of coupling between the volumetric and deviatoric responses may make it difficult to frame a proper inverse problem for experiments (Borja, 2013; Bryant and Sun, 2018; Ma et al., 2021).

The substance octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (HMX) is the energetic constituent in many PBXs. HMX exhibits several crystal polymorphs [Cady]. The thermodynamically stable form on the 300 K isotherm, for pressures between 0 and approximately 30 GPa is known as $\beta$-HMX, for which the crystal structure is monoclinic with a unit cell containing two molecules Cady et al. (1963). Numerous theoretical studies of HMX physical properties and thermo-mechanical response to shocks have been reported; we do not discuss them here, but Das et al. (2021) provides a recent entry point into that literature. All MD simulations discussed below were performed for $\beta$-HMX in the $P2_1/n$ space group setting.

Previous work, such as Pereverzev and Sewell (2020)), has obtained pressure- and temperature-dependent elastic coefficients by applying small strain increment to a sample at thermal equilibrium at the desired thermodynamic state and determining the corresponding stress and elasticity tangential tensor at that state. Another feasible alternative that we consider here is to assume that the finite strain elasticity of $\beta$–HMX is that of a *Green elastic material* or hyper-elastic material (Marsden and Hughes, 1994; Ogden, 1997). In this approach we postulate that (1) the state of the stress in the current configuration can be solely determined by the state of the deformation of the current configuration relative to one choice of a reference configuration such as the crystal lattice vectors at (300 K, 1 atm) and (2) there exists an elastic stored energy functional of which the derivative with respect to the strain measure is the energy-conjugated stress measure. Comparing to the former approach, which tabulates the elasticity tensor at prescribed states for a given pressure and temperature, the hyperelasticity approach has several distinct advantages. First, the prediction of the elastic strain energy, stress measure, and elastic tangential stress are all bundled together into one scalar-valued tensor function, instead of separate calculations for stress and elastic tangent that might not be consistent with each other. Second, unlike the more widely used tabular approach, the hyperelasticity model does not require pressure as an input to predict elastic constitutive responses and hence enables consistency easily. Finally, by assuming the existence of such an elastic stored energy, the stability, and uniqueness of the constitutive responses as well as other attributes such as convexity, material frame indifference, and symmetry can be more easily analyzed mathematically (Ogden, 1997; Borja, 2013).

Nevertheless, with a few exceptions, such as Holzapfel and Ogden (2009); Holzapfel et al. (2004); Latorre and Mőntáns (2015), the majority of hyperelasticity models are limited to isotropic materials or materials of simple symmetry such as transverse isotropic and orthotropic. Hyperelastic models for materials of lower symmetry such as monoclinic or triclinic are less common (Clayton, 2010). This can be attributed to the fact that the strain and the stress for anisotropic materials are not necessarily co-axial, and handcrafting a mathematical expression for the energy functional that leads to accurate
predictions of stress and tangent, therefore, becomes a challenging task.

To overcome this technical barrier, we introduce a transfer learning approach that generates a neural network model for the hyperelastic response of β-HMX from molecular dynamics (MD) simulations. Our new contributions, to the best knowledge of the authors, are listed below:

1. Traditional supervised learning approaches often employ objective/loss functions that match the stress-strain responses (Ghaboussi et al., 1991; Lefik and Schrefler, 2003; Heider et al., 2020; Frankel et al., 2019), the elastic stored energy (Le et al., 2015; Teichert et al., 2019), or matching the energy, stress, or elastic tangent fields (Vlassis et al., 2020; Vlassis and Sun, 2021) with the raw data considered as the ground truth. This direct approach, however, is not suitable for MD data where the change of one state to another will lead to fluctuation that makes direct Sobolev training not productive (Czarnecki et al., 2017). To overcome this problem, we introduce a pre-training step in which the data are pre-processed through a filter and the underlying non-fluctuating patterns are extracted to train the neural network models.

2. We introduce a transfer learning approach where the additional desirable attributes (e.g. frame invariance) and necessary conditions for the correctness of the constitutive laws (e.g. material symmetry) can be enforced with a simple re-training.

3. We also introduce a post-training validation procedure where the focus is not only on predicting stress-strain responses but on the desirable properties of the elastic tangential operator. To compare to the previous literature that employs measures in the geometrical linear regime to measure anisotropy, we introduce a reverse mapping from Cuitino and Ortiz (1992) that generates the infinitesimal small-strain tangent from the finite strain counterpart. With these metrics available, we can examine the convexity and strong ellipticity of the learned function and also evaluate whether predicted constitutive responses exhibit the same evolution of anisotropy as the MD benchmark while ensuring that the filtering process does not lead to non-physical responses at the continuum level. The accuracy of the model is assessed by comparing MD-simulated and learned stresses as functions of strain, and by comparing the pressure-dependent tangent stiffness from the learned model against explicit predictions of the elastic tensor reported recently (Pereverzev and Sewell, 2020) for β-HMX states on the 300 K hydrostatic isothermal compression curve. The latter comparison, in particular, provides an incisive test of the accuracy of the learned functional, as this information was not used explicitly as part of the training set.

The rest of the paper is organized as follows. We first provide a brief account of the database generation procedure, including pertinent details of the MD simulations, the procedure to generate stress-strain data from the MD predictions, and the procedure to filter out the high-frequency responses (Section 2). We briefly review the setup of our hyperelastic model (Section 3) and then outline the major ingredients for the supervised learning of the hyperelastic energy functional, including the Sobolev training, the Hessian sampling techniques for controlling the higher-order derivatives and the way to incorporate the physical constraints in the training procedure (Section 4). This section is followed by the validation procedure that tests the attributes of the learned hyperelasticity models with physical constraints not included in the training problems (Section 5). The results of the numerical experiments are reported in Section 6 followed by concluding remarks in Section 7.

As for notations and symbols, bold-faced and blackboard bold-faced letters denote tensors (including vectors which are rank-one tensors); the symbol ‘·’ denotes a single contraction of adjacent indices of two tensors (e.g., \( a \cdot b = a_i b_i \) or \( c \cdot d = c_{ij} d_{jk} \)); the symbol ‘‘’ denotes a double contraction of adjacent indices of tensor of rank two or higher (e.g., \( C : \epsilon = C_{ijkl} \epsilon_{kl} \)); the symbol ‘\( \otimes \)’ denotes a juxtaposition of two vectors (e.g., \( a \otimes b = a_i b_j \)) or two symmetric second-order tensors (e.g., \( (a \otimes \beta)_{ijkl} = a_{ij} \beta_{kl} \)). We also define identity tensors: \( I = \delta_{ij} \), \( \bar{I} = \delta_{ik} \delta_{jl} \), and \( \bar{I} = \delta_{il} \delta_{jk} \), where \( \delta_{ij} \) is the Kronecker delta. We denote the Eulerian coordinate as \( \{x_1, x_2, x_3\} \) and the corresponding three orthonormal basis vectors as \( e_1, e_2 \) and \( e_3 \) accordingly. As for sign conventions, unless specified, the directions of the tensile stress and dilative pressure are considered as positive.
2 Database generation via molecular dynamics simulations

In this section, we discuss the specifics of the MD simulation setup used to generate the database used for the hyperelastic energy functional discovery. We provide a theoretical background for the simulations as well as details on the system setup. We demonstrate the output results for the simulations and describe the post-processing procedure to render them suitable for our machine learning algorithms.

Training data for the neural networks are obtained by computing the Cauchy stress tensor for isothermal samples as functions of imposed tensorial strains. The strains used correspond variously to uniaxial compression or tension, pure shear, and combination strains. The imposed strains are restricted to states below the threshold for mechanical failure of $\beta$-HMX as predicted by the MD. By learning the underlying free-energy functional, we can extract the hyperelastic response from second-order and higher-order strain derivatives.

2.1 Force field

The MD simulations were performed using LAMMPS Plimpton (1995) in conjunction with a modified version of the all-atom, fully flexible, non-reactive force field originally developed for HMX by Smith and Bharadwaj (S-B). Smith and Bharadwaj (1999); Bedrov et al. (2000); Kroonblawd et al. (2016); Mathew and Sewell (2018); Chitsazi et al. (2020) Intramolecular interactions in the S-B force field are modeled using harmonic functions for covalent bonds, three-center angles, and improper dihedral (“wag”) angles; and truncated cosine expansions for proper dihedrals. Intermolecular non-bonded interactions between atoms separated by three or more covalent bonds (i.e., 1-4 and more distant intramolecular atom pairs) are modeled using Buckingham-plus-charge (exponential-6-1) pair terms. Here and in Refs. Zhao et al. (2020); Kroonblawd and Fried (2020); Das et al. (2021), a steep repulsive pair potential was incorporated between non-bonded atom pairs to prevent ‘overtopping’ of the exponential-6-1 potential at short non-bonded separations $R$, which can occur under shock-wave loading due to the global maximum in the potential at distances of approximately 1 Å with a divergence to negative infinity as $R \to 0$. This is accomplished by superposing on the Buckingham potential a Lennard-Jones 12-6 potential in a way such that the $R^{-12}$ repulsive core strictly prevents overtopping while having practically no effect on the potential, and therefore on the interatomic forces, for non-bonded distances more than approximately 1 Å. Evaluation of dispersion and Coulomb pair terms was computed using the particle-particle particle-mesh (PPPM) k-space method (Hockney and Eastwood, 1988) with a cutoff value of 11 Å and with the PPPM precision set to $10^{-6}$.

2.2 MD Simulation cell setup

Three-dimensionally periodic (3-D) primary simulation cells were generated starting from the unit-cell lattice parameters for $\beta$-HMX (P2$_1$/n space group setting) predicted by the force field (at 300 K and 1 atm), by simple replication of the unit cell in 3-D space. This results in a monoclinic-shaped primary simulation cell. The mapping of the crystal frame to the Cartesian lab frame is $a || \hat{x}$, $b || \hat{y}$, and $c$ in the $+z$ space. Starting primary cell sizes for the uniaxial compression and uniaxial tension cases were approximately 30 nm parallel to the strain direction and approximately 10 nm transverse to it; those for pure shear deformation were approximately 10 nm $\times$ 10 nm $\times$ 10 nm; and those for biaxial compression were approximately 30 nm $\times$ 30 nm $\times$ 30 nm. Figure 1 depicts a unit cell of $\beta$-HMX and snapshots of representative simulation cells prior to the beginning of deformation. Table 1 contains details of the system sizes used.

2.3 Simulation details

MD trajectories were propagated using the velocity Verlet integrator in LAMMPS (Verlet, 1967; Swope et al., 1982). Primary cells constructed as described in the preceding paragraph were thermally equi-
Figure 1: Unit cell of \( \beta \)-HMX (panel (a)) and snapshots of representative simulation cells for (b) uniaxial compression and tension, (c) shear deformation, and (d) biaxial compression. Cyan for carbon, navy for nitrogen, red for oxygen, and white for hydrogen.

Table 1: System sizes for uniaxial compression and tension, pure shear deformation, and biaxial compression production simulations.

| Simulation                        | \( L_x \) (nm) | \( L_y \) (nm) | \( L_z \) (nm) | Number of Molecules |
|-----------------------------------|----------------|----------------|----------------|-------------------|
| Compression/Tension along \( \hat{x} \) | 30.3           | 10.5           | 10.6           | 12,880            |
| Compression/Tension along \( \hat{y} \) | 10.5           | 30.3           | 10.6           | 12,992            |
| Compression/Tension along \( \hat{z} \) | 10.5           | 10.5           | 30.4           | 12,800            |
| Shear deformation                 | 10.5           | 10.5           | 10.6           | 4,480             |
| Biaxial compression               | 30.3           | 30.3           | 30.4           | 106,720           |

Librated in the isochoric-isothermal (NVT) ensemble at 300 K by initially selecting atomic velocities from the 300 K Maxwell distribution followed by 20 ps of trajectory integration. Temperature control was achieved using the Nosé-Hoover thermostat (Nosé, 1984; Hoover, 1985) as implemented in LAMMPS with the damping parameter set to 50.0 fs. A 0.2 fs time step was used for the thermal equilibration.

Fifteen isothermal MD production simulations, comprising three apiece for uniaxial compression, uniaxial tension, and biaxial compression, and six for pure shear (i.e., positive and negative shear directions for three distinct shear cases) were performed at \( T = 300 \) K using NVT integration in conjunction with the LAMMPS fix deform command. The integration time step was 0.20 fs and the thermostat damping parameter was set to 20.0 fs. The system potential energy, temperature, pressure, Cauchy stress-tensor components, and primary cell lattice vectors were recorded at 10 fs intervals for subsequent analysis.

For the uniaxial compression and tension simulations, strain was applied parallel to the long direction of the primary cell while holding both the transverse cell lengths and the tilt factors constant. The strain rate was set to the constant value \( \pm 0.1/100 \) ps, applied uniformly at each time step. The uniaxial simulations were performed for 300 ps, resulting in a total strain of 0.3 for those cases.
For the shear simulations, the system was deformed along one of the three tilt factors (i.e., xy, xz, and yz) while the cell edge lengths were maintained at constant values. A constant strain rate of ± 0.1/100 ps was applied for 300 ps, resulting in total positive or negative shear strains of 0.3.

For the biaxial compression simulations, the primary cell was compressed along two axes simultaneously in the lab frame (i.e., x and y, y and z, or x and z) while holding the third cell length and the tilt factors constant. The strain rate was set to ± 0.05/100 ps along both directions. Trajectory integration was performed for 300 ps resulting in a strain of 0.15 along each of the two affected directions.

2.3.1 MD results

Figure 2 contains the system potential energy, pressure, Cauchy stress-tensor components, and lattice vectors vs. time for the case of uniaxial compression along ˆy. The effects of deformation are clearly evident in the potential energy and stress-tensor components (panels (a) and (c)), where it can be seen that the sample yields at t ≈ 190 ps. Data for times up to approximately 10 ps before failure were used for further analysis using machine learning.

The Cauchy stress is obtained from the standard LAMMPS command and the expression can be found there (cf. lam).

2.4 Filtering MD simulation data

The raw data from the MD simulations are not expected to be smooth, due to thermal fluctuations. These fluctuations may depend on the thermostat employed and the size of the system. This temperature fluctuation, however, is not supposed to be captured by the hyperelasticity energy functional, which is only designed to capture the macroscopic constitutive responses.

To deal with the MD data, we can either introduce a regularization process during the machine
learning training or we can simply filter out the Gaussian noise that might otherwise affect the convexity and therefore the stability of the hyperelasticity model.

While one can filter the Cauchy stress tensor on a component-by-component basis, such a strategy may lead to a filtered Cauchy stress that depends on the coordinate system. Thus, this strategy should be avoided. While there are potentially more sophisticated techniques for filtering tensorial and multi-dimensional data (e.g. Muti and Bourennane (2005)), here we introduce a spectral decomposition on the Cauchy stress such that

$$\sigma = \sum_{a=1}^{3} \sigma_a n^a \otimes n^a. \quad (1)$$

Following this step, a 1D moving average filter is applied to each of the eigenvalues of the Cauchy stress and to the Euler angles that represent the orthogonal basis vector—$n^a$. To remove the noise, we used a 1D uniform filter on the data series that works similar to a rolling-average window. The temporal length of the filter window is equal to that of 300 MD observations. This length of the filter window is selected after a manual trial-and-error such that we may suppress the noise of the tensorial time series without greatly distorting the global recorded constitutive response. Note that a highly fluctuated stress data may increase the difficulty of Sobolev training the hyperelasticity energy functional but also affect the stability of the constitutive responses at the continuum scale. Hence, this preliminary step is necessary.

![Figure 3: Filtering of MD simulation data with a uniform filter for a compression test along the x axis. The filtering is performed twice for two MD simulations with different thermostat coupling parameters and thus different oscillations.](image)

To examine whether the filter introduces significant bias to the filter data, we apply our filtering procedure to two MD simulations with the same strain path but initiated from different initial conditions. The filtered and unfiltered constitutive responses are compared for both cases, as shown in Fig. 3. The two MD simulations demonstrate different fluctuation patterns but the filtered responses are very close. The uniform filter used to process the data appears to capture almost identical behaviors for both simulations.

3 Finite strain hyperelastic neural network functional for $\beta$-HMX

In this work, we will approximate a finite strain hyperelastic energy functional for $\beta$-HMX using a feed-forward neural network architecture trained with a modified Sobolev training loss function that incorporates additional physical constraints via a transfer learning technique. The following assumptions and setup have been made:

1. There exists one stress-free configuration for the $\beta$-HMX for which the stored elastic energy is zero. This configuration constitutes the reference configuration for the deformation mapping.
2. We assume that all the data used in the training are purely elastic with no path dependence.

3. Thermo-mechanical and rate-dependence effects on the elasticity are neglected.

4. A filter is used to reduce the high-frequency responses.

The stored energy functional $\bar{\psi}$ can be written as a function of the deformation gradient $F$. The first Piola-Kirchhoff stress $P$ is conjugate to the deformation gradient $F$ and can be obtained from the following relation,

$$ P(F) = \frac{\partial \bar{\psi}(F)}{\partial F}. $$

Notice that a necessary condition for this energy functional to be correct is the material-frame indifference. Here the deformation gradient is not sensitive to rigid-body translation. However, to ensure the the $SO(3)$ equivalence, the machine learning generated energy functional must satisfy the following constraint,

$$ \bar{\psi}(F) = \bar{\psi}(QF), \quad \forall Q \in SO(3) \quad (3) \quad $n$.

A possible way to bypass the need to introduce additional constraints in the loss function is to derive the energy functional as a function of the Green strain tensor $E$ for which:

$$ E' = \frac{1}{2}(C' - I) = \frac{1}{2}(F^T \cdot F' - I) = \frac{1}{2}(F^T \cdot Q^T \cdot Q \cdot F - I) = \frac{1}{2}(F^T \cdot F - I) = \frac{1}{2}(C - I) = E, \quad (4) \quad $n$.

so we then acquire an equivalent expression:

$$ \bar{\psi}(F) = \psi(E). \quad (5) \quad $n$

The second Piola-Kirchhoff stress $S$ is conjugate to the Green strain $\dot{E}$, which is derived as:

$$ S(E) = \frac{\partial \psi(E)}{\partial E}. \quad (6) \quad $n$

In addition to the frame invariance, another major benefit of expressing the energy functional in terms of the Green strain tensor is that the resultant stress measure is symmetric and the elastic tangential operator possesses both major and minor symmetries. These symmetries may reduce the dimension of the input parametric space 9 to 6 and hence simplify the training. Furthermore, while $C$ and $E$ can both be used as the input for the inherently frame-indifferent energy functional that yields $S$ as the first derivative, $E = 0$ implies the energy functional becomes zero. Meanwhile, training $\bar{\psi}(F)$ as the learned function can be more convenient for implicit PDE solver for large deformation where the tangent corresponding to $P - F$ is required to solve the linearized system of equation.

As such, we will train two hyperelasticity functionals, one takes the deformation gradient as input and another one takes the Green strain tensor as input respectively and we will compare the results obtained from numerical experiments. The relationships among elasticity tangential tensors corresponding to different stress-strain conjugate pairs will also be discussed in Section 5.

4 Stress-based Sobolev training for stored-energy function

We introduce a neural network training technique that constructs the hyperelasticity energy functional using solely the stress data and a single reference configuration where $F = I$. Recall that a feed-forward neural network can be trained to approximate an energy functional $\psi$ that takes the Green-Lagrange deformation tensor $E$ as input. This energy function is parametrized by weights $W$ and biases $b$. The supervised learning that minimizes the inner product of the difference between the true $\psi$ and the approximated $\hat{\psi}$ for $N$ samples can be written as
\[ W', b' = \arg\min_{W,b} \left( \frac{1}{N} \sum_{i=1}^{N} \| \psi_i - \hat{\psi}_i \|_2^2 \right), \]

where \( \psi_i = \psi(E_i) \) and \( \hat{\psi}_i = \hat{\psi}(E_i) \) accordingly. While this approach could reduce the discrepancy of the predicted and true free energy values—if available, it would not necessarily improve the performance of the stress predictions. However, true energy functional value samples were not available from the MD simulations, thus, we designed a variation of the free energy functional loss function that only uses stress data. From this point forward, we refer to free energy simply as energy.

### 4.1 Sobolev constraints for the hyperelastic energy functional

To introduce a hyperelasticity model suitable to incorporate into numerical solvers for boundary value problems, the accuracy, stability, robustness, smoothness, and uniqueness of the hyperelasticity responses are all important to consider. Unlike neural networks that directly generate stress predictions, a hyperelasticity model must be sufficiently smooth and differentiable to avoid discontinuity in the predicted stress and elastic tangent (Vlassis et al., 2020; Vlassis and Sun, 2021; Le et al., 2015).

Consider the stored-energy functional solely constructed via (1) a reference configuration where the Green strain tensor equals to \( E_0 \), and (2) the Cauchy stress measured in the MD simulations. The corresponding loss function reads,

\[ W', b' = \arg\min_{W,b} \left( \| \psi_0 - \hat{\psi}_0 \|_2^2 + w_s \left\| \frac{\partial \psi_0}{\partial E_0} - \frac{\partial \hat{\psi}_0}{\partial E_0} \right\|_2^2 + \frac{w_s}{N} \sum_{i=1}^{N} \left\| \frac{\partial \psi_i}{\partial E_i} - \frac{\partial \hat{\psi}_i}{\partial E_i} \right\|_2^2 \right), \]

where \( \psi_0 = \psi(E_0) \) and \( \hat{\psi}_0 = \hat{\psi}(E_0) \) are the true and approximated values of the energy functional at strain \( E_0 \), \( N \) is the number of non-trivial stress data, and \( w_s \) is the weighting factor for the multi-objective optimization. In this work, we use the configuration at (300 K, 1 atm) as the reference and assume this configuration is stress-free.

The corresponding loss function for the energy conjugate pair \( P - F \) hyperelastic model is:

\[ W', b' = \arg\min_{W,b} \left( \| \psi_0 - \hat{\psi}_0 \|_2^2 + w_p \left\| \frac{\partial \psi_0}{\partial F_0} - \frac{\partial \hat{\psi}_0}{\partial F_0} \right\|_2^2 + \frac{w_p}{N} \sum_{i=1}^{N} \left\| \frac{\partial \psi_i}{\partial F_i} - \frac{\partial \hat{\psi}_i}{\partial F_i} \right\|_2^2 \right), \]

where \( \psi_0 = \psi(F_0) \) and \( \hat{\psi}_0 = \hat{\psi}(F_0) \) are the true and approximated values of the energy functional at the reference configuration \( F_0 \) at (300 K, 1 atm).

### 4.2 Transfer learning to enforce frame invariance

A hyperelastic model described by the conjugate pair \( P - F \) tensors is expected to satisfy the frame invariance conditions described in Eq. (3). To ensure that the frame invariance is preserved during training, we re-use a previously trained neural network but modifying the loss function by introducing a number \( M \) of random rotations \( Q^m, i = 1, 2, ..., M \) and penalizing the violation of the objectivity by adding the following weighted objectives:

\[ w_p \frac{1}{M} \sum_{m=1}^{M} \| P(Q^m F) - Q^m P(F) \|_2^2 + w_c \frac{1}{M} \sum_{m=1}^{M} \| A(Q^m F) - Q^m A(F) \|_2^2. \]

\[ 0 \leq w_p, w_c \leq 1 \]
4.3 Transfer learning to enforce crystal symmetries

The monoclinic unit cell of the single crystal $\beta$-HMX in the $P2_1/n$ space group setting is shown in Figure 4. The covariant crystal basis vectors $M_1, M_2,$ and $M_3$ represent the crystal axis in the reference configuration, with corresponding contravariant basis vectors $M^1, M^2,$ and $M^3$ such that $M^i \cdot M_j = \delta^i_j$. Furthermore, the covariant crystal basis vectors in the current configuration are denoted as $m_1, m_2,$ and $m_3$, where $m_i = FM_i$. A general form of the deformation gradient $F$ that maintains the monoclinic unit cell reads:

$$V_F = \{ F | RU, U = \sum_{i=1}^{3} (a_i M_i \otimes M^i) + a_4 (M_1 \otimes M^3 + M_3 \otimes M^1), R \in SO(3), a_j \in \mathbb{R} \}.$$  \hspace{1cm} (11)

![Figure 4: Monoclinic unit cell of $\beta$-HMX in the $P2_1/n$ space group setting. The lattice constants are $a = 6.53\,\AA$, $b = 11.03\,\AA$, $c = 7.35\,\AA$, $\alpha = \gamma = 90^\circ$, and $\beta = 102.689^\circ$ (at 295 K) (Eiland and Pepinsky, 1954). The vectors $e_1, e_2, \text{and} e_3$ denote the basis vectors of the global Cartesian coordinate system.](image)

Under an imposed deformation gradient under the constraint in Equation (11), the symmetry group of the monoclinic unit cell in the current configuration reads

$$V_Q = \{ Q | Q = \exp \left[ k \pi \frac{\text{spn}(m_2)}{\| m_2 \|} \right], \quad m_2 = FM_2, \quad F \in V_F, \quad k \in \mathbb{Z} \}.$$  \hspace{1cm} (12)

Here, the infinitesimal rotation map and the finite rotation map are defined as (Simo et al., 1989)

$$\text{spn}(\theta) = - \varepsilon \cdot \theta, \quad \exp \left[ \frac{\text{spn}(\theta)}{\| m_2 \|} \right] = I + \frac{\sin(\theta)}{\theta} \text{spn}(\theta) + \frac{1 - \cos(\theta)}{\theta^2} \text{spn}(\theta)^2,$$

where $\varepsilon$ is the permutation tensor and $\theta = \| \theta \|$ is the rotation angle.

Due to material symmetry, the first elasticity tensor $C^{PF}$ has the following symmetric property:

$$C^{PF}(FQ) = C^{PF}(F), \quad \forall F \in V_F, Q \in V_Q,$$  \hspace{1cm} (13)

where the tensor components are expressed in the global Cartesian frame for convenience.

To ensure that the crystal symmetry is preserved, we can again re-use the previous trained function $(9)$ and $(10)$, and modify the loss function by introducing $M$ number of rotations $Q^k_{\text{sym}} \in V_Q, k = 1, 2, ..., M$ to penalize the violation of the material symmetry by adding the following weighted objectives:

$$\sum_{k=1}^{M} \left( w_\psi \frac{1}{N} \sum_{i=1}^{N} \left\| \psi(FQ^i_{\text{sym}}) - \psi(F) \right\| + w_P \frac{1}{N} \sum_{i=1}^{N} \left\| P(FQ^i_{\text{sym}}) - P(F)Q^k_{\text{sym}} \right\| + w_A \frac{1}{N} \sum_{i=1}^{N} \left\| A(FQ^i_{\text{sym}}) - A(F)Q^k_{\text{sym}} \right\| \right).$$  \hspace{1cm} (14)
5 Post-training validation of the predicted elastic tangential operators

In this section, we introduce numerical tests to determine whether the predicted constitutive responses are thermodynamically admissible, preserve the symmetry, and lead to unique and stable elastic responses. A subset of these criteria are required to constitute a correct constitutive law (e.g. material frame invariance), while others such as the convexity and the strong ellipticity are not necessary conditions but are desirable properties for stability and uniqueness of the boundary value problem. While in principle many of these physics constraints/laws can be incorporated into the loss function in the supervised learning process, putting all the constraints explicitly into the loss function is not necessarily always ideal, as the multiple constraints may alter the landscape of the loss function and thus complicate the search for the optimal energy functional (Mavrotas, 2009).

As such, our goal is to introduce a suite of necessary conditions which the learned hyperelasticity constitutive law must fulfill. These necessary conditions, along with the fact that the hyperelasticity constitutive law must be capable of generating predictions within a threshold error, are necessary but not sufficient to guarantee the safety of using the machine learning model for high-consequence high-risk predictions (such as those for explosives).

5.1 Mapping between finite and infinitesimal kinematics

To examine the admissibility of the hyperelasticity model and compare the finite strain model with other published results based on the infinitesimal strain assumption, the connections among the tangents of different energy-conjugate pairs are provided below for completeness. Here our first goal is to obtain an underlying small-strain tangent of the finite-strain counterpart by using the logarithmic and exponential mappings, such that the elasticity tensors predicted here and those from the literature can be compared. Recall that the logarithmic elastic strain $\epsilon$ can be defined as (Cuitino and Ortiz, 1992),

$$\epsilon = \ln U^2 = \frac{1}{2} \ln C,$$  \hspace{1cm} (15)

where $U$ is the right-stretch tensor and $C$ is the right Cauchy-Green strain tensor. The small-strain elastic tensor $C^{\sigma - \epsilon}$ can be obtained from the chain rule,

$$C^{\sigma - \epsilon} = \frac{\partial \sigma}{\partial \epsilon} = \frac{\partial \sigma}{\partial S} : \frac{\partial E}{\partial \epsilon} = \frac{1}{2} \frac{\partial \sigma}{\partial S} : \frac{\partial C}{\partial \epsilon} = \frac{1}{2} \frac{\partial \sigma}{\partial S} : \frac{\partial \exp 2\epsilon}{\partial \epsilon}, \hspace{1cm} (16)$$

where $\sigma = J^{-1} F : S \cdot F^T$ is the Cauchy stress. To compute the small-strain elasticity tensor, one first rewrites Eq. (15) in an infinite series representation,

$$C = \exp 2\epsilon = \sum_{n=0}^{\infty} \frac{1}{n!} (2\epsilon)^n.$$  \hspace{1cm} (17)

As such, the Cartesian component of the derivative $\partial C / \partial \epsilon$ reads (Miehe, 1998)

$$[\frac{\partial C}{\partial \epsilon}]_{ijkl} = \sum_{n=1}^{\infty} \frac{2^n}{n!} \sum_{m=1}^{\infty} [\epsilon_{ik}^{n-1}] [\epsilon_{lj}^{n-m}]. \hspace{1cm} (18)$$

The first tangential tensor $C^{P - F}$ can be related to the second derivative of the hyperelastic energy functional $\psi(E)$,

$$C^{P - F} = \frac{\partial P}{\partial F} = \frac{\partial S}{\partial E} : F \cdot F + S \otimes \delta,$$  \hspace{1cm} (19)

where $g$ is the metric tensor. This expression is derived from Marsden and Hughes (1994) (see page 215), where we simply use the chain rule to link the tangents $\partial S / \partial E$ with $\partial S / \partial C$. Note that this tensor
corresponds to the first Piola-Kirchhoff stress and the deformation gradient, and does not possess minor symmetry.

In both Eq. (16) and Eq. (19), the derivative $\partial S / \partial E$ is obtained from the neural elastic stored energy, while the rest of the terms can be obtained via either analytical solution or automatic differentiation.

### 5.2 Strong ellipticity

While many works are dedicated to training elastic constitutive laws (Ghaboussi et al., 1998; Pernot and Lamarque, 1999; Le et al., 2015; Hoerig et al., 2018; Fuhg et al., 2021; Huang et al., 2020; Vlassis et al., 2020; Vlassis and Sun, 2021), surprisingly few among these analyze the stability and uniqueness of the learned neural network constitutive laws or provide any evidence of the well-posedness for the trained model. Recent work by (Klein et al., 2021) address this issue by enforcing polyconvexity via invariants (cf. Hartmann and Neff (2003)).

Consider $A$ to be the acoustic tensor corresponding to $C^{PF}$ and that $C^{PF}$ is the elastic tangential operator for the energy conjugate pairs $(P, F)$, that is,

$$ A(N) = N \cdot C^{PF} \cdot N $$

The Legendre-Hadamard condition requires that for any pair of vectors $N$ and $m$, the following condition holds:

$$ m \cdot A \cdot m \geq 0, $$

where $N$ is a Lagrangian unit vector and $m$ is an Eulerian vector. Because we assume that $\beta$-HMX is a Green-elastic material, the necessary and sufficient conditions for strong ellipticity are (cf. Ogden (1997) page 392)

$$ A_{ii}(N) > 0, \quad i \in \{1, 2, 3\} $$

$$ A_{ii}(N)A_{jj}(N) - A_{ij}(N)^2 > 0, \quad j \neq i \in \{1, 2, 3\} $$

$$ \det A(N) > 0 $$

for any $N \in \mathbb{R}^3$. Notice that the material response is nonlinear and the ellipticity may vary according to the Eulerian vector $m$. A simple way to ensure the conditions (22)-(24) are satisfied is to create the worst-case scenario, that is, find the infimum, and the unit vectors $N$ that minimize $A_{ii}(N)$, $A_{ii}(N)A_{jj}(N) - A_{ij}(N)^2$, and $\det A(N)$ accordingly and check whether the three terms remain positive. Depending on the parameterization, the corresponding minimization problems can be written as

$$ f(q) = A_{ii}(N(q)), \quad \arg\min_q f(q), \quad N(q) \in S^2 $$

$$ g(q) = A_{ii}(N(q))A_{jj}(N(q)) - A_{ij}(N(q))^2, \quad \arg\min_q g(q), \quad N(q) \in S^2 $$

$$ d(q) = \det A(N(q)), \quad \arg\min_q d(q), \quad N(q) \in S^2 $$

where $q$ represents a parametrization of the unit vector $N(q)$. Mota et al. (2016) provide a comprehensive review of how different parameterizations, namely the spherical, stereographic, projective and tangent parameterizations, may lead to different local mininizers of the acoustic tensor in the parametric space. For spherical parameterization, a unit vector $N$ is an element of the unit sphere $S^2$ which can be parameterized by the spherical coordinates, that is, the polar angle $\phi \in [0, \pi]$ and the azimuthal angle $\theta \in [0, \pi]$:

$$ N(\phi, \theta) = \sin \phi \cos \theta e_1 + \sin \phi \sin \theta e_2 + \cos \phi e_3, $$

12
where \( \{e_1, e_2, e_3\} \) is the the orthogonal basis for \( \mathbb{R}^3 \).

To ensure stability for any given admissable deformation, we must ensure that Eqs. (22)-(24) are valid for any \( F \). While this can be, in principle, determined analytically for hand-crafted energy functionals, the expression of the neural network energy functional would likely be too complicated to analyze. As such, we again resort to constructing a test to check the hypothesis that the material demonstrates strongly ellipticity, via an attempt to find the minima, that is,

\[
f'(q, F) = A_{ii}(N(q), F), \quad \arg\min_{q,F} f'(q, F), \quad N(q) \in S^2, F \in GL^+(3)
\]

\[
g'(q, F) = A_{ii}(N(q), F) A_{jj}(N(q), F) - A_{ij}(N(q), F)^2,
\]

\[
\arg\min_{q,F} g'(q), \quad N(q) \in S^2, F \in GL^+(3)
\]

\[
d'(q, F) = \det A(N(q), F), \quad \arg\min_{q,F} d'(q, F), \quad N(q) \in S^2, F \in GL^+(3).
\]

It is impossible to test all the possible deformation gradients in the MD simulations while maintaining the path independence of the constitutive responses, so we instead construct a test where we only consider a range of possible deformation gradients and search for the minima within this range.

The numerical strong ellipticity test is conducted via the following three steps.

1. We create two sets of point clouds in the parametric space with uniform spacing, \( V_q = \{q_1, q_2, q_3, \ldots\} \) and \( V_F = \{F_1, F_2, F_3, \ldots\} \), and select the combination of \((q, F)\) that minimizes \( f', g', d' \). If there exist other \((q, F)\) combinations that yield a value sufficiently close to the minimum (say within 5\% difference), then the additional coordinates will be stored as the candidate position(s) for the gradient-free search. This treatment is to ensure that more local optimal points can be identified and compared and to avoid the issues exhibited in Mota et al. (2016).

2. We then use the candidate position determined from the previous step as the starting point and apply a gradient-free optimizer via the third-party gradient-free optimizer library (cf. Simon Blanke (since 2020)) to examine whether we can find new coordinates for which the functions \( f'(q, F), g'(q, F), \) and \( d'(q, F) \) are smaller than the candidate position(s) identified in Step 1.

3. If Eqs. (22) - (24) are not violated in the worst case obtained from the Step 2, then we consider the neural network functional to have passed the strong ellipticity test.

5.3 Convexity and growth conditions

In nonlinear elasticity in the finite strain regime, convexity is not necessary and can be over-restrictive for physical phenomena that involve instability or buckling (Clayton, 2010). Nevertheless, the convexity condition has to be satisfied to predict stable elastic responses under large deformation. The convexity condition can be stated as (cf. Ogden, 1997),

\[
\psi(F') - \psi(F) - \text{tr}(P \cdot (F - F')) \geq 0
\]

Because convexity is not a requirement for realistic simulations (although it might be expected for HMX), we do not incorporate this criterion in the training of the neural network. However, the uniqueness and stability of the elasticity model are not only important for predicting realistic elastic responses but crucial if the model will be deployed as the underlying elasticity model for crystal plasticity and damage models.

Another important condition to prevent degenerated elastic behavior is from Rosakis and Simpson (1994) which requires

\[
\psi(F) \rightarrow \infty \quad \text{as} \quad \det F \rightarrow 0^+.
\]
Recall that \( \det F \to 0 \) only happen if the distance between two material points that was non-zero in the reference configuration vanishes in the current configuration. Note that it is unlikely a material would remain elastic if the volumetric deformation is extremely large. Furthermore, enforcing these constraints explicitly in the loss function is difficult due to the infinity. Nevertheless, the constraint may provide a helpful indicator of the admissibility of the machine learning extrapolated predictions. As a result, we suggest a post-training validation test where we generate the response for deformation gradients with \( \det F \) approaching zero and observe whether the resultant energy is monotonically increasing.

5.4 Material Anisotropy

A predictive elasticity model must preserve the overall crystal symmetry while capturing how the anisotropy of the elasticity tensor evolves under arbitrary deformation. The degree of anisotropy of the elastic response can be measured by various metrics available in the literature (cf. (Li and Bradt, 1987; Kube, 2016; Ranganathan and Ostoja-Starzewski, 2008)). Many of these anisotropy metrics (or indices) are intended for components of the elasticity tensor. Typically, the distinction between the secant and tangential elastic tensors is not taken into account. This can be confusing for materials undergoing finite deformation where both material and geometrical non-linearities play important roles in the anisotropy of the constitutive response. More importantly, the impacts of the former and latter types of non-linearity should be distinguished properly such that a meaningful evaluation can be conducted.

5.4.1 Ledbetter and Miglioli general anisotropy index

Here, we use the idea from previous work due to Ledbetter and Miglioli (2006), where the ratio between the maximum and minimum shear-wave speed is used to define an anisotropy measure. Interestingly, this method can also be used to detect instability as the vanishing of the slowest wave speed is accompanied by divergence of the Ledbetter-Miglioli index.

This measure can be easily extended to the finite strain regime by replacing the infinitesimal elasticity tangent with the elasticity tensor corresponding to the first Piola-Kirchhoff stress and deformation gradient (Ogden, 1997). This idea can be summarized into the following steps.

1. Generate as many unit vectors \( N \) as possible.
2. Solve the Christoffel equation for each unit vector \( N \), that is,
   \[
   \det \left( N \cdot C(F) \cdot N - \rho v^2 I \right) = 0
   \]  
(34)
3. Pick the largest solution \( v_2 \) and the smallest solution \( v_1 \). Then, the anisotropy index is simply
   \[
   A_I = \frac{v_2^2}{v_1^2}
   \]  
(35)

Here, instead of a Monte Carlo search, we can leverage the search formulated in Section 5.2 to obtain the smallest eigenvalue \( v_1 \) and largest eigenvalue \( v_2 \) of the acoustic tensor. Again, the optimization is conducted by using a uniformly spaced point cloud to search for the initial guess, then a gradient-free optimizer is used to find the normal vectors that maximize and minimize \( v \).

6 Results

In this section, we discuss the performance of neural network models for discovering the hyperelastic energy functional from the \( \beta \)-HMX MD simulation data. We describe the training setup of the
networks and compare the performance of the architectures. We then demonstrate the predictive capabilities of the models against the present MD simulation data and elastic constants taken from the literature for the same MD force field used here. Finally, we investigate the energy functional models in terms of how well they satisfy desired properties from the hyperelasticity literature.

6.1 Training performance and learning capacity

In this section, we discuss the performance of the neural network architectures for the Sobolev constraints described in Section 4. We first demonstrate how we trained the neural networks to generate a hyperelastic energy functional data from the MD simulation data. We use two different architectures to discover the hyperelastic energy functional for $\beta$-HMX. The first architecture is based on the energy conjugate pair $S = E$ (Model $M_1$). The input and output variables are symmetric tensors and, thus, can be described by six components. The second architecture is based on the energy conjugate pair $P = F$ (Model $M_2$). In addition, we also re-train Model $M_2$ with an additional material frame indifference constraint (Eq. (10)) in the loss function (model $M_3$). As the difference in the predictions obtained from Models $M_2$ and $M_3$ is minor, we did not enforce the Eq. (10) explicitly in the last model we trained (Model $M_4$). Instead, only monoclinic symmetry is enforced as an additional term for the weighted loss function in the re-training step to ensure that the material symmetry is preserved.

Table 2: Summary of the trained models.

| Model | Description |
|-------|-------------|
| $M_1$ | Energy conjugate pair $S = F$ model trained via the loss function described in Eq. 8. |
| $M_2$ | Energy conjugate pair $P = F$ model trained via the loss function described in Eq. 9. |
| $M_3$ | Energy conjugate pair $P = F$ model trained with pre-trained model $M_2$ and additional loss function Eq. (10) to enforce material frame indifference. |
| $M_4$ | Energy conjugate pair $P = F$ model trained pre-trained model $M_2$ and additional loss function Eq. (14) to enforce monoclinic symmetry. |

The energy functional neural networks have a feed-forward architecture consisting of a hidden dense layer (100 neurons / ReLU), followed by two multiply layers (cf. Vlassis and Sun (2021)), then another hidden dense layer (100 neurons / ReLU), and finally an output dense layer (Linear). The training and validation procedures of the neural network are implemented in Python with machine learning libraries Keras (Chollet et al., 2015) and Tensorflow (Abadi et al., 2015). The kernel weight matrix of the layers was initialized with a Glorot uniform distribution and the bias vector with a zero distribution. The models were trained on 163400 MD simulation data points and validated on 70030 data points. All the models were trained for 1000 epochs with a batch size of 512, using the Nadam optimizer (Dozat, 2016) initialized with default values in the Keras library.

The loss function training curves for the architectures $M_1$ and $M_2$ are demonstrated in Fig. 5. The two architectures appear to have similar accuracy so they will be used interchangeably below. The predictive capabilities of $M_1$ and $M_2$ are further demonstrated in Section 6.2.1.

To check and, if necessary, enforce the frame invariance of the neural network hyperelastic models as described in Section 4.2, we conduct a transfer learning experiment by retraining the neural network model $M_2$. We first train the energy conjugate pair $P = F$ model ($M_2$) for 1000 epochs without any frame invariance constraints in the loss function (i.e., Eq. (9)). We record the frame invariance metrics during training by applying random rotation $Q$ tensors on the input deformation gradient.
tensors and examine whether the material response is frame invariant; that is, whether the predicted energy remains the same before and after rotation and whether the predicted stress tensor rotates accordingly. The trained model $M_2$ is then retrained with the additional frame invariance constraints in Eq. (10) for another 1000 epochs (model $M_3$). The comparison of the training curves for $M_2$ and $M_3$ is shown in Fig. 6. Model $M_2$ appears to already satisfy well the frame invariant properties, with the additional constraints of model $M_3$ mostly improving the frame invariance energy constraints.

We also perform a transfer learning experiment by retraining the neural network model $M_2$ to ensure it retains the observed $\beta$-HMX crystal symmetries as described in Section 4.3. We first train the energy conjugate pair $P - F$ model ($M_2$) for 1000 epochs without any symmetry constraints in the loss function and record the symmetry metrics during training. By applying a rotation $Q_{\text{sym}}$ on the input deformation gradient tensors, we check for the material response to retain the expected monoclinic symmetry behavior. The check includes the constraints up to the first-order derivatives of the network. The trained model $M_2$ is then retrained with the additional symmetry constraints in Eq. (14) for another 1000 epochs (model $M_4$). The results for the two training experiments are shown
Figure 7: Comparison of the training loss curves for (a) the energy and (b) stress symmetry constraints for the energy conjugate pair \( P - F \) model (\( M_2 \)) without any symmetry constraints in the loss function and the energy conjugate pair \( P - F \) model (\( M_4 \)) trained with the additional symmetry-constraint loss function Eq. (14).

Remark 1. Rescaling of the training data. As a pre-processing step, we have normalized all data to avoid the vanishing or exploding gradient problem that may occur during the back-propagation process (Bishop et al., 1995). The \( X_i \) sample of a measure \( X \) is scaled to a unit interval via,

\[
X_i := \frac{X_i - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}},
\]

where \( X_i \) is the normalized sample point. \( X_{\text{min}} \) and \( X_{\text{max}} \) are the minimum and maximum values of the measure \( X \) in the training data set such that all different types of data used in this paper (e.g. strain, stress, etc) are all normalized within the range \([0, 1] \).

6.2 Validation of the constitutive responses

In this section, we validate the neural network predicted constitutive response against MD simulation data as well as \( \beta \)-HMX elastic coefficients from the literature. We also monitor the learned physical properties for the trained models, such as the strong ellipticity, the energy growth, and the anisotropy index.

6.2.1 Validation against unseen MD simulations

We validate the predictive performance of the learned models against unseen MD simulation loading paths. The neural network architectures considered in this section are the energy conjugate pair \( S - E \) model (\( M_1 \)) and the energy conjugate pair \( P - F \) model (\( M_2 \)).

The stress predictions of the networks against three uniaxial strains along the axes \( x_1 \), \( x_2 \), and \( x_3 \) are demonstrated in Fig. 8 and Fig. 11. All the symmetric stress tensor components are plotted against the main loading direction of the MD simulation experiment. The predictions are compared against the raw MD simulation data before the filtering pre-processing described in Section 2.4. The stress predictions for three pure shear MD experiments in the positive and negative \( e_1 \otimes e_2 \), \( e_2 \otimes e_3 \), and \( e_1 \otimes e_3 \) directions are shown in Fig. 9 and Fig. 12. Finally, the stress predictions for three biaxial compression tests along the \( x_1 \) and \( x_2 \), \( x_2 \) and \( x_3 \), and \( x_1 \) and \( x_3 \) axes are shown in Fig. 10 and Fig. 13.
It is noted that the stress fluctuations in the MD shear data appear to have a larger magnitude than those of the axial simulations. However, the magnitude of the fluctuations of the stress components is
similar across all simulations; it appears to be larger in the shear simulations due to the smaller scale of the stress response.

Both models are able to accurately capture the shear behavior of $\beta$-HMX, which differs greatly in the positive vs. negative directions as seen in Fig. 9 and Fig. 12. The shear stress response of the material appears to be highly non-linear and exhibits directional dependence. This behavior is not expected to be captured by a material model with an invariant formulation, as it requires specific treatment of the shear response along different directions to replicate the directional dependent behavior even qualitatively. Here, however, a more general representation of the material using the full second-order stress and strain tensors allows for the neural network to automatically recover this behavior and rather precisely.

![Graphs showing stress response](image)

Figure 11: Comparison of the predicted stress response against three uniaxial MD simulations for the energy conjugate pair $P - F$ models $M_2$ and $M_4$. (a,d) Uniaxial compression and extension along the $x_1$ axis for models $M_2$ and $M_4$ respectively. (b,e) Uniaxial compression and extension along the $x_2$ axis for models $M_2$ and $M_4$ respectively. (c,f) Uniaxial compression and extension along the $x_3$ axis for models $M_2$ and $M_4$ respectively.

Remark 2. As seen in Fig. 11, the predictions of the models $M_2$ and $M_4$ are very close. We have also examined the other predictions and the discrepancies of the stress predictions inferred from $M_2$ and $M_4$ are also very minor. Hence, we do not include those comparisons in the paper for brevity. In the following sections, the validation tests of the energy conjugate pair $P - F$ models’ properties will be performed on the model $M_2$ as the behavior of the models $M_2$, $M_3$, and $M_4$ was observed to be similar.
6.2.2 Validation of Strong ellipticity

In this section, we perform the strong ellipticity tests as described in Section 5.2 on the trained neural network. The neural network architecture used in this comparison is model $\mathcal{M}_2$, which uses the energy conjugate pair $P - F$. This model was chosen for the convenience of obtaining the fourth-order elasticity tensor needed for the acoustic tensor checks.

This check is performed by initially predicting the fourth-order elasticity tensor at a specific deformation gradient level. We sample 1000 unit vectors $N$ on the unit sphere $S^2$ in spherical coordinates by sampling the polar angle $\phi \in [0, \pi]$ and the azimuthal angle $\theta \in [0, \pi]$ in a uniform grid, following Eq. (28). These vectors can be used to construct 1000 initial acoustic tensors following Eq. (20). The acoustic tensors will be used as the initial grid landscape for a gradient-free optimizer set to discover the minimum values of the three strong ellipticity tests described in Eqs. (29), (30), and (31). We use a Hill Climbing gradient-free optimizer search, using the library implemented by Simon Blanke (since 2021).
Figure 14: Validation of strong ellipticity conditions (a), (b), (c) with the criteria in Eq. (25), Eq. (26), and Eq. (27), respectively, for an elasticity tensor close to the reference strain state. The unit vectors were sampled from the surface of a unit sphere to perform the validation via a Hill Climbing gradient-free optimizer search (d, e, f). The minimum value of the condition value discovered by the optimizer is marked.

2020), to find the pair of $(\phi, \theta)$ that minimizes the strong ellipticity check values. The Hill Climbing algorithm performs 10000 iterations of the search per test to discover the minimum value of the check, which in most tests was obtained within the first 5000 iterations of the search.

The predicted strong ellipticity test and the corresponding optimizer search for the minimum values are demonstrated in Fig. 14 and Fig. 15 for two different elasticity tensors. In Fig. 14, we show the ellipticity test results for the elasticity tensor close to the relaxed reference state, that is when the deformation gradient is the identity tensor. The neural network passes all three ellipticity tests, discovering the minimum of all tests to be greater than zero in the unit vector search space. In Fig. 15, we show the first strain state of a biaxial compression simulation along the $x_1$ and $x_2$ axes where the strong ellipticity test fails – the acoustic tensor determinant for Eq. (29) is found to be less than zero for the first time (compression of approximately 8% along the $x_1$ and $x_2$ axes).

Given that the machine learning generated constitutive responses match very well with the filtered MD simulations (as shown in Figs. 8-13), the acoustic tensor losing positive definiteness is an indication of unstable elastic responses corresponding to the shear mode along the $N$ direction which could be potentially physical (C. Picu, personal communication, 2021).

6.2.3 Validation of energy growth for extrapolated predictions

In this section, we perform the validation check described in Section 5.3 to monitor if the behavior of the predicted energy functional degenerates for very large deformations. To test that, we impose deformation gradients on the neural network model spanning several orders of magnitude with the $\det F$ decreasing towards zero. The test is performed on the neural network architecture for the en-
Figure 15: (a) Loss of the strong ellipticity condition in the prediction of a biaxial compression test along the $x_1$ and $x_2$ axes. The unit vectors were sampled from the surface of a unit sphere to perform the validation via a Hill Climbing gradient-free optimizer search (b). The minimum value of the condition value discovered by the optimizer is marked.

Energy conjugate pair of $P - F$ (model $M_2$). As the Jacobian decreases, the energy functional values are expected to increase monotonically (Eq. (33)). Therefore, we apply a sequence of volumetric compression deformation gradients with the Jacobian approaching zero and plot the energy against the increasing pure volumetric deformation. The results are shown in Fig. 16.

Note that the $\beta$-HMX may exhibit plastic yielding or damage under high pressure. When this occurs, it is not physically feasible to have an elastic response. Yet the continuum mechanics theory validation does require that the stored energy approach infinity as a finite volume of HMX crystal collapses into a point (Rosakis and Simpson, 1994). This does not happen in our trained neural network model even though the growth rate within the training data interval seems reasonable.

A similar extrapolation issue has been investigated previously in Versino et al. (2017) in which the symbolic regression requires an additional artificial data point added in order to prevent an incorrect prediction of softening. Presumably, a similar treatment can also be applied either by adding a very large artificial data point with a very large energy at the supposedly singular point or by rigorously enforcing the singularity in the learned energy functional. At this point, robust ways to introduce singular data into the neural network and the formulation of the loss function are not clear, but we intend to examine it in future studies. Nevertheless, the results do reveal that the energy functional trained by the neural network may only be valid within the interval of the data and that any extrapolated results outside of the data interval must be used with caution, even if a significant number of physical constraints (e.g. material symmetry) have already been applied as auxiliary objectives for the supervised learning.
6.2.4 Stress-dependent anisotropy of HMX crystal

In this section, we recover the predicted material-response anisotropy index as described in Section 5.4.1 to monitor the evolution of material anisotropy. The anisotropy index is acquired for the neural network architecture of the energy conjugate pair of $P - F$ (model $M_2$). To obtain the index, the fourth-order elasticity tensor is predicted at different deformation gradients along a prescribed loading path. We then sample 1000 unit vectors $N$ per deformation gradient in a uniform grid, following Eq. (28), by sampling the polar angle $\phi \in [0, \pi]$ and the azimuthal angle $\theta \in [0, \pi]$. For each elasticity tensor, 1000 initial acoustic tensors are constructed according to Eq. (20). The Hill Climbing algorithm performs 10000 iterations of the search deformation gradient sample to discover the minimum $v^2_1$ and the maximum $v^2_2$ values of each acoustic tensor. The anisotropy index $A_I$ is then calculated using Eq. (35). The anisotropy index calculated for three loading paths is demonstrated in Fig. 17.

Interestingly, for all three uniaxial compression cases, the Ledbetter-Migliori anisotropy index, which is the ratio of the fastest and slowest shear wave speeds of the $\beta$-HMX crystal, all tend to increase significantly. The most significant changes occur when the uniaxial deformation is more than 8%. In all three cases, the elastic anisotropy is not very profound when the deformation is small. However, in all three cases, the anisotropy index jumps from less than 10 to more than 40 in the deformation along the $x_1$ direction and more than two orders in the $x_2$ and $x_3$ direction. These results
signify the importance of capturing the evolving anisotropy of the HMX materials.

6.2.5 Comparisons with literature calculations on elasticity

We now provide the coefficients of the elastic tangents for the \( S - E \) and \( P - F \) conjugated pairs obtained from the trained neural network energy functionals, models \( \mathcal{M}_1 \) and \( \mathcal{M}_2 \) and compare them with the elasticity tangent for the \( \sigma - \epsilon \) conjugated pairs previously reported by Pereverzev and Sewell (2020).

In the present work, the strain measure is obtained differently in the sense that the models in these papers introduce only one reference configuration such that \( F = I \) when the Cauchy pressure is at \( 10^{-4} \) GPa. Meanwhile, the strain measure in the \( \beta \)-HMX in (Pereverzev and Sewell, 2020) is reset at different reference pressure where the MD simulation begins. This difference is minor for the atmospheric pressure case, for which the geometrical nonlinearity is insignificant, but may lead to significant differences in the values of the elastic tangent coefficients for high-pressure cases. Note that, due to the anisotropic nature of the elastic responses, the imposed Cauchy pressure may also lead to isochoric deformation due to volumetric-deviatoric coupling. As such, the coordinates of the reference and current configurations \( x_i \) and \( X_I \) are not necessarily co-axial. Hence, a direct comparison of the values of the coefficient is not productive.

Furthermore, discrepancies may also be caused by the different data de-noising processes employed in Pereverzev and Sewell (2020). In this paper, we employ a de-noising algorithm to filter out the high-frequent oscillation in the constitutive responses before the supervised learning is conducted whereas Pereverzev and Sewell (2020) employs a finite-difference approximation with a sufficiently large strain increment to calculate the elasticity tensor.

Nevertheless, a comparison of elasticity tangent operators from previous MD simulations, as well as those obtained for different conjugate stress-strain pairs, does indicate the significance of geometric nonlinearity in the material responses and the importance of taking it into consideration in numerical simulations.

In the MD simulation, the crystal cell of \( \beta \)-HMX is first equilibrated at a target temperature and pressure through an isochoric-isothermal (NVT) simulation and then strains are imposed at different directions to obtain the stress information for the differentiation. The comparison between the predicted and literature reported elastic coefficients at 300 K temperature at pressure \( 10^{-4} \) GPa and 5 GPa is demonstrated in Tables 3 and 4, respectively, for the neural network models \( \mathcal{M}_1 \) and \( \mathcal{M}_2 \) as described in Section 6.1. It is noted that for the energy conjugate pair of \( P - F \) Model \( \mathcal{M}_2 \), the full tangent requires a \( (9 \times 9) \) matrix to represent it in the Voigt notation.

Table 3 shows the results of different elastic tangents obtained from the neural network calculation and those obtained from Pereverzev and Sewell (2020). While there are differences among the three tangents, they are relatively minor. This is expected as the geometrical nonlinearity is not significant.

Table 4, on the other hand, shows a more significant difference in the numerical values of the coefficients for different energy-conjugated pairs. This is consistent with the derivation in Section 5.1 where the deformation gradient at this point is no longer infinitesimal and the incorporation of the geometrical nonlinearity is necessary to capture the elastic constitutive responses properly.

7 Conclusions

This paper introduce a mechanistic machine learning framework to infer anisotropic hyperelasticity energy functional from \( \beta \)-HMX. Conventionally, machine learning constitutive laws are often formulated to match experimental data. As such, the discrepancy between experimental data and the predictions is often the only term in the loss function for training and validation. Here we attempt to formulate the training of hyperelastic model not only to minimizing the discrepancy of data but also introduce additional objectives to ensure that the learned hyperelastic model obey the physics constraints. To ensure the robustness of the predictions, we also
Table 3: Comparison of the predicted $\beta$-HMX elastic coefficients (GPa) at pressure $10^{-4}$ GPa for the energy conjugate pair $S - E$ (Model $\mathcal{M}_1$) and the energy conjugate pair $P - F$ (Model $\mathcal{M}_2$) to the ones reported by Pereverzev and Sewell (2020) for (300 K, $10^{-4}$ GPa). Note that $C^{P-F}$ is not symmetric but the additional terms are not shown for brevity.

| $D_{ij}$ | $C_{ijkl}$ | Model $\mathcal{M}_1$ | Model $\mathcal{M}_2$ | Pereverzev and Sewell (2020) |
|---------|------------|-----------------------|-----------------------|-------------------------------|
| $D_{11}$ | $C_{1111}$ | 21.354                | 25.861                | 22.97                         |
| $D_{22}$ | $C_{2222}$ | 22.149                | 18.092                | 22.62                         |
| $D_{33}$ | $C_{3333}$ | 21.314                | 21.627                | 21.67                         |
| $D_{44}$ | $C_{1212}$ | 8.616                 | 5.8335               | 8.645                         |
| $D_{55}$ | $C_{2323}$ | 10.982                | 8.225                 | 10.407                        |
| $D_{66}$ | $C_{1313}$ | 9.497                 | 10.078                | 9.527                         |
| $D_{12}$ | $C_{1122}$ | 8.789                 | 6.898                 | 9.2                           |
| $D_{13}$ | $C_{1133}$ | 12.348                | 12.7828               | 12.32                         |
| $D_{23}$ | $C_{2233}$ | 15.913                | 13.375                | 12.37                         |
| $D_{15}$ | $C_{1123}$ | -0.998                | -0.584                | -0.43                         |
| $D_{25}$ | $C_{2223}$ | 4.247                 | -0.877                | 4.47                          |
| $D_{35}$ | $C_{3323}$ | 2.192                 | -0.792                | 1.84                          |
| $D_{46}$ | $C_{1213}$ | 2.484                 | 1.571                 | 2.248                         |

introduce a set of validation tests to examine the admissibility (e.g., preserving material symmetry, obeying growth conditions) and stability (convexity, strong ellipticity) of the constitutive responses generated from the trained neural networks. With the usage of Sobolev training and automatic differentiation to facilitate the training of constitutive laws, the resultant model exhibit highly accurate predictions within the training data range. These treatments are shown to be effective in improving the accuracy and robustness of the predictions, while the theoretical validation may provide the much-needed post hoc interpretability of the neural network constitutive laws to understand the properties of the machine learning models. More importantly, the validation exercise may provide a reliable way to reveal the weakness of the models and safeguard against cherry-picking interpretation, which could be a key ingredient to make black-box neural network predictions more trustworthy.

8 Data availability statements

The code used to conduct the validation tests will be available in a Github repository upon publication of this manuscript. The datasets generated and/or analyzed during the current study are available from the authors upon reasonable request.

9 Acknowledgements

Fruitful discussions with Andrey Pereverzev and Bahador Bahmani are gratefully acknowledged. The efforts and labor hours are primarily supported by the Air Force Office of Scientific Research under grant contracts FA9550-19-1-0318, with additional support provided to WCS and NNV from the the NSF CAREER grant at National Science Foundation under grant contracts CMMI-1846875 and OAC-1940203.
Table 4: Comparison of the predicted $\beta$-HMX elastic coefficients (GPa) at pressure 5 GPa for the energy conjugate pair $S - E$ (model $M_1$) and the energy conjugate pair $P - F$ (model $M_2$) with the ones reported in Pereverzev and Sewell (2020). Note that $C^{p-F}$ is not symmetric but the additional terms are not shown for brevity.

| $D_{ij}$ | $C_{ijkl}$ | Model $M_1$ | Model $M_2$ | Pereverzev and Sewell (2020) |
|---------|------------|--------------|--------------|-------------------------------|
| $D_{11}$ | $C_{1111}$ | 80.157       | 87.556       | 87.71                         |
| $D_{22}$ | $C_{2222}$ | 68.666       | 53.441       | 62.08                         |
| $D_{33}$ | $C_{3333}$ | 71.453       | 72.011       | 62.11                         |
| $D_{44}$ | $C_{1212}$ | 0.358        | 0.033        | 19.461                        |
| $D_{55}$ | $C_{2323}$ | 3.71         | 2.813        | 34.08                         |
| $D_{66}$ | $C_{1313}$ | -2.999       | 1.736        | 19.662                        |
| $D_{12}$ | $C_{1122}$ | 44.048       | 26.828       | 36.93                         |
| $D_{13}$ | $C_{1133}$ | 46.187       | 32.423       | 52.95                         |
| $D_{23}$ | $C_{2233}$ | 55.267       | 52.603       | 46.49                         |
| $D_{15}$ | $C_{1123}$ | -0.939       | 0.639        | -11.32                        |
| $D_{25}$ | $C_{2223}$ | 10.358       | -6.108       | 11.1                          |
| $D_{35}$ | $C_{3323}$ | -0.421       | -6.120       | 2.48                          |
| $D_{46}$ | $C_{1213}$ | 5.546        | -4.082       | 6.06                          |

References

Lammps molecular dynamic simulator. LAMMPS is available at http://lammps.sandia.gov.

D. Bedrov, G. D. Smith, and T. D. Sewell. Thermal conductivity of liquid octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (HMX) from molecular dynamics simulations. Chem. Phys. Lett., 324:64, 2000.

Christopher M Bishop et al. Neural networks for pattern recognition. Oxford university press, 1995.

Ronaldo I Borja. Plasticity: modeling & computation. Springer Science & Business Media, 2013.

Frank Philip Bowden, Frank Philip Bowden, and AD Yoffe. Initiation and growth of explosion in liquids and solids, CUP Archive, 1985.

Eric C Bryant and WaiChing Sun. A mixed-mode phase field fracture model in anisotropic rocks with consistent kinematics. Computer Methods in Applied Mechanics and Engineering, 342:561–584, 2018.

Howard H Cady, Allen C Larson, and Don T Cromer. The crystal structure of $\alpha$-hmx and a refinement of the structure of $\beta$-hmx. Acta crystallographica, 16(7):617–623, 1963.
R. Chitsazi, M. P. Kroonblawd, A. Pereverzev, and T. D. Sewell. A molecular dynamics simulation study of thermal conductivity anisotropy in \( \beta \)-octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (\( \beta \)-HMX). Model. Simul. Mater. Sc., 28:025008, 2020.

François Chollet et al. Keras. https://keras.io, 2015.

John D Clayton. Nonlinear mechanics of crystals, volume 177. Springer Science & Business Media, 2010.

Alberto Cuitino and M Ortiz. A material-independent method for extending stress update algorithms from small-strain plasticity to finite plasticity with multiplicative kinematics. Engineering computations, 1992.

Wojciech Marian Czarnecki, Simon Osindero, Max Jaderberg, Grzegorz Świrszcz, and Razvan Pascanu. Sobolev training for neural networks. In 31th Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, USA, 2017.

P. Das, P. Zhao, D. Perera, T. Sewell, and H. S. Udaykumar. Molecular dynamics-guided material model for the simulation of shock-induced pore collapse in \( \beta \)-octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (\( \beta \)-HMX). J. Appl. Phys., 130(8):085901, 2021.

Timothy Dozat. Incorporating nesterov momentum into adam, 2016.

Phillip Frank Eiland and Ray Pepinsky. The crystal structure of cyclotetramethylene tetranitramine. Zeitschrift für Kristallographie-Crystalline Materials, 106(1-6):273–298, 1954.

Ari L Frankel, Reese E Jones, Coleman Alleman, and Jeremy A Templeton. Predicting the mechanical response of oligocrystals with deep learning. Computational Materials Science, 169:109099, 2019.

Jan Niklas Fuhg, Michele Marino, and Nikolaos Bouklas. Local approximate gaussian process regression for data-driven constitutive laws: Development and comparison with neural networks. arXiv preprint arXiv:2105.04554, 2021.

J Ghaboussi, JH Garrett Jr, and Xiping Wu. Knowledge-based modeling of material behavior with neural networks. Journal of engineering mechanics, 117(1):132–153, 1991.

Jamshid Ghaboussi, David A Pecknold, Mingfu Zhang, and Rami M Haj-Ali. Autoprogressive training of neural network constitutive models. International Journal for Numerical Methods in Engineering, 42(1):105–126, 1998.

Stefan Hartmann and Patrizio Neff. Polyconvexity of generalized polynomial-type hyperelastic strain energy functions for near-incompressibility. International journal of solids and structures, 40(11):2767–2791, 2003.

Yousef Heider, Kun Wang, and WaiChing Sun. So (3)-invariance of informed-graph-based deep neural network for anisotropic elastoplastic materials. Computer Methods in Applied Mechanics and Engineering, 363:112875, 2020.

R. W. Hockney and J. W. Eastwood. Computer Simulation Using Particles. Hilger, New York, NY, 1988.

Cameron Hoerig, Jamshid Ghaboussi, and Michael F Insana. Data-driven elasticity imaging using cartesian neural network constitutive models and the autoprogressive method. IEEE transactions on medical imaging, 38(5):1150–1160, 2018.

Gerhard A Holzapfel and Ray W Ogden. On planar biaxial tests for anisotropic nonlinearly elastic solids. a continuum mechanical framework. Mathematics and mechanics of solids, 14(5):474–489, 2009.
Gerhard A Holzapfel, Gerhard Sommer, and Peter Regitnig. Anisotropic mechanical properties of tissue components in human atherosclerotic plaques. J. Biomech. Eng., 126(5):657–665, 2004.

Daniel E Hooks, Kyle J Ramos, Cynthia Anne Bolme, and Marc J Cawkwell. Elasticity of crystalline molecular explosives. Propellants, Explosives, Pyrotechnics, 40(3):333–350, 2015.

W. G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. Phys. Rev. A, 31:1695, 1985.

Daniel Z Huang, Kailai Xu, Charbel Farhat, and Eric Darve. Learning constitutive relations from indirect observations using deep neural networks. Journal of Computational Physics, 416:109491, 2020.

Dominik Klein, Mauricio Fernández, Robert J. Martin, Patrizio Neff, and Oliver Weeger. Polyconvex anisotropic hyperelasticity with neural networks, 2021.

M. P Kroonblawd and L. E. Fried. High explosive ignition through chemically activated nanoscale shear bands. Phys. Rev. Lett., 124:206002, 2020.

M. P Kroonblawd, N. Mathew, S. Jiang, and T. D. Sewell. A generalized crystal-cutting method for modeling arbitrarily oriented crystals in 3D periodic simulation cells with applications to crystal-crystal interfaces. Comput. Phys. Commun., 207:232, 2016.

Christopher M Kube. Elastic anisotropy of crystals. AIP Advances, 6(9):095209, 2016.

Marcos Latorre and Francisco Javier Montáns. Anisotropic finite strain viscoelasticity based on the sidoroff multiplicative decomposition and logarithmic strains. Computational Mechanics, 56(3):503–531, 2015.

BA Le, Julien Yvonnet, and Q-C He. Computational homogenization of nonlinear elastic materials using neural networks. International Journal for Numerical Methods in Engineering, 104(12):1061–1084, 2015.

Hassel Ledbetter and Albert Migliori. A general elastic-anisotropy measure. Journal of applied physics, 100(6):063516, 2006.

Marek Lefik and Bernhard A Schrefler. Artificial neural network as an incremental non-linear constitutive model for a finite element code. Computer methods in applied mechanics and engineering, 192(28-30):3265–3283, 2003.

Z Li and Richard C Bradt. The single-crystal elastic constants of cubic (3c) sic to 1000 c. Journal of materials science, 22(7):2557–2559, 1987.

Ran Ma, WaiChing Sun, and Catalin R Picu. Atomistic-model informed pressure-sensitive crystal plasticity for crystalline hmx. International Journal of Solids and Structures, 232:111170, 2021.

Jerrold E Marsden and Thomas JR Hughes. Mathematical foundations of elasticity. Courier Corporation, 1994.

N. Mathew and T. Sewell. Pressure-dependent elastic coefficients of β-HMX from molecular simulations. Prop., Explos., Pyrotech., 43:233, 2018.

George Mavrotas. Effective implementation of the ε-constraint method in multi-objective mathematical programming problems. Applied mathematics and computation, 213(2):455–465, 2009.

Ch Miehe. Comparison of two algorithms for the computation of fourth-order isotropic tensor functions. Computers & structures, 66(1):37–43, 1998.
Alejandro Mota, Qiushi Chen, James W Foulk III, Jakob T Ostien, and Zhengshou Lai. A cartesian parametrization for the numerical analysis of material instability. International Journal for Numerical Methods in Engineering, 108(2):156–180, 2016.

Damien Muti and Salah Bourennane. Multidimensional filtering based on a tensor approach. Signal Processing, 85(12):2338–2353, 2005.

S. Nosé. A unified formulation of the constant-temperature molecular-dynamics methods. J. Chem. Phys., 81:511, 1984.

Raymond W Ogden. Non-linear elastic deformations. Courier Corporation, 1997.

Andrey Pereverzev and Tommy Sewell. Elastic coefficients of β-hmx as functions of pressure and temperature from molecular dynamics. Crystals, 10(12):1123, 2020.

Stéphane Pernot and C-H Lamarque. Application of neural networks to the modelling of some constitutive laws. Neural Networks, 12(2):371–392, 1999.

S. Plimpton. Fast parallel algorithms for short-range molecular dynamics. J. Comp. Phys., 117:1, 1995.

Shivakumar I Ranganathan and Martin Ostoja-Starzewski. Universal elastic anisotropy index. Physical Review Letters, 101(5):055504, 2008.

Phoebus Rosakis and Henry C Simpson. On the relation between polyconvexity and rank-one convexity in nonlinear elasticity. Journal of elasticity, 37(2):113–137, 1994.

JC Simo, DD Fox, and MS Rifai. On a stress resultant geometrically exact shell model. part ii: The linear theory; computational aspects. Computer Methods in Applied Mechanics and Engineering, 73(1):53–92, 1989.

Simon Blanke. Gradient-Free-Optimizers: Simple and reliable optimization with local, global, population-based and sequential techniques in numerical search spaces. https://github.com/SimonBlanke, since 2020.

G. D. Smith and R. K. Bharadwaj. Quantum chemistry based force field for simulations of HMX. J. Phys. Chem. B, 103:3570, 1999.

W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. J. Chem. Phys., 76:637, 1982.

Gregory H Teichert, AR Natarajan, A Van der Ven, and Krishna Garikipati. Machine learning materials physics: Integrable deep neural networks enable scale bridging by learning free energy functions. Computer Methods in Applied Mechanics and Engineering, 353:201–216, 2019.

L. Verlet. Computer “Experiments” on classical fluids. I. thermodynamical properties of lennard-jones molecules. Phys. Rev., 159:98, 1967.

Daniele Versino, Alberto Tonda, and Curt A Bronkhorst. Data driven modeling of plastic deformation. Computer Methods in Applied Mechanics and Engineering, 318:981–1004, 2017.

Nikolaos N Vlassis and WaiChing Sun. Sobolev training of thermodynamic-informed neural networks for interpretable elasto-plasticity models with level set hardening. Computer Methods in Applied Mechanics and Engineering, 377:113695, 2021.

Nikolaos N Vlassis, Ran Ma, and WaiChing Sun. Geometric deep learning for computational mechanics part i: Anisotropic hyperelasticity. Computer Methods in Applied Mechanics and Engineering, 371:113299, 2020.

P. Zhao, S. Lee, T. Sewell, and H. S. Udaykumar. Tandem molecular dynamics and continuum studies of shock-induced pore collapse in TATB. Propellants, Explos. Pyrotech., 45:1, 2020.