

LA-UR-23-23237

Accepted Manuscript

Multiscale Reactive Model for 1,3,5-Triamino-2,4,6-trinitrobenzene Inferred by Reactive MD Simulations and Unsupervised Learning

Lafourcade, Paul
Maillet, Jean-Bernard
Roche, Jerome
Sakano, Michael N.
Hamilton, Brenden William
Strachan, Alejandro

Provided by the author(s) and the Los Alamos National Laboratory (2023-08-14).

To be published in: The Journal of Physical Chemistry C

DOI to publisher's version: 10.1021/acs.jpcc.3c02678

Permalink to record:

<https://permalink.lanl.gov/object/view?what=info:lanl-repo/lareport/LA-UR-23-23237>



Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Triad National Security, LLC for the National Nuclear Security Administration of U.S. Department of Energy under contract 89233218CNA000001. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

Multiscale reactive model for 1,3,5-triamino-2,4,6-trinitrobenzene inferred by reactive MD simulations and unsupervised learning

P. Lafourcade,^{1,2,*} J.-B. Maillet,^{1,2} J. Roche,³ M. Sakano,⁴ B. W. Hamilton,⁵ and A. Strachan⁶

¹CEA, DAM, DIF, F-91297 Arpajon, France

²Université Paris-Saclay, LMCE, 91680 Bruyères-le-Châtel, France

³Institut Polytechnique de Paris, 91190 Palaiseau, France

⁴Sandia National Laboratories, Albuquerque, New Mexico, 87185, USA

⁵Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico, 87548, USA

⁶School of Materials Engineering and Birck Nanotechnology Center, Purdue University, West Lafayette, Indiana 47907, USA

When high-energy-density materials are subjected to thermal or mechanical insults at extreme conditions (shock loading), a coupled response between the thermo-mechanical and chemical behavior is systematically induced. We develop a reaction model for the fast chemistry of 1,3,5-triamino-2,4,6-trinitrobenzene (TATB) at the mesoscopic scale where the chemical behavior is determined by underlying microscopic reactive simulations. The slow carbon clusters formation is not discussed in the present work. All-atom reactive MD simulations are performed with the ReaxFF potential and a reduced-order chemical kinetics model for TATB is fitted on isothermal and adiabatic simulations of single crystal chemical decomposition. Unsupervised machine learning techniques based on the non-negative matrix factorization are applied to MD trajectories to model the decomposition kinetics of TATB in terms of a four components model. The associated heats of reaction are fit to the temperature evolution from adiabatic decomposition trajectories. Using a chemical species analysis, we show that NMF captures the main chemical decomposition steps of TATB, and provides an accurate estimation of their evolution with temperature. The final analytical formulation, coupled to a diffusion term, is incorporated into a continuum formalism and simulation results are compared one-to-one against MD simulations of a 1D reaction propagation along different crystallographic directions and with different initial temperatures. A good agreement is found for both temporal and spatial evolution of the temperature field.

I. INTRODUCTION

The shockwave induced initiation of energetic materials is thought to be governed by energy localizations called hotspots[1, 2]. Hotspots can be formed from a wide range of mechanisms including the collapse of porosity[3–5], shear band formation[6, 7], interfacial friction[8, 9], and jetting[10, 11].

These mechanisms lead to regions of excess energy that react significantly faster than in the bulk, shocked crystal[12, 13]. If above a critical size, these hotspots can result in a self-propagating reaction wave that can eventually lead to explosion and detonation[14]. Hotspot formation and compression initiation can also lead to complex chemical events such as mechanochemistry[15–17] and metalization[18, 19]. Additionally, hotspot formation from

mechanisms such as pore collapse is highly sensitive to initial conditions such as pore size and shape[3, 4] and shock-wave velocity[10, 20]. This leads to hotspot formation and chemical initiation being a highly coupled and complex series of thermo-mechanical and chemical processes.

While atomistic simulations are a powerful tool for probing these phenomena, molecular dynamics is too computationally costly to probe hotspot formation and shock to detonation transition for realistic length and time scales, which ranges from the nanometer and picosecond regime to the micrometer and microsecond regime. Hence, a valid multiscale approach is necessary to map the physics of these processes onto mesoscale and continuum methods. Previous works have utilized a variety of ways to up-scale MD into continuum models for both mechanics[4, 21–23] and chemistry[24–27].

Such approaches have already been used for high explosives materials. The ultrafast dynamic response and shear localization processes in β -HMX were studied using

* E-mail: paul.lafourcade@cea.fr

a Lagrangian-Eulerian code [5, 28] in which both chemistry and mechanics are taken into account. In addition, the hotspots dynamics of porous TATB and HMX have been investigated using mesoscale simulations in order to build up surrogate models of ignition and growth models [29, 30]. In addition, a lot of efforts have been provided in taking into account the microstructure at the grain scale to use numerical twin approaches that allow direct comparisons between meso-scale simulations and experiments [31].

In the case of TATB, which exhibits peculiar thermo-mechanical behavior due to its strong anisotropy, the deformation mechanisms at the single crystal scale under dynamical or shock loading strongly depend on the direction of solicitation. These mechanisms range from elastic instabilities (when the loading is parallel to the molecular layers plane), pure elastic buckling or twinning, basal homogeneous shear and layer sliding, and out-of-plane dislocations [32–34]. Efforts has been devoted to the construction of a mesoscopic thermo-mechanical model [35], where the necessary ingredients for proper coarse graining have been shown to be entirely contained into the directional yield stress surface [34]. The present constitutive law for TATB at the meso-scale includes non-linear anisotropic hyperelasticity [36] depending on an EOS and a phase field by reaction pathway formalism to model twinning. While crystal plasticity is not yet included and will be the object of future work, the model has been validated against one-to-one comparison with atomistic simulations using a variety of deformation setups. To properly study hotspot formation and deflagration-to-detonation transition, a matching mesoscopic chemistry model is needed. Here we focus on the chemical reactions aspect of hotspots, using all-atom molecular dynamics simulations to create a continuum scale reaction model for 1,3,5-triamino-2,4,6-trinitrobenzene (TATB) that properly predicts kinetics and energetics for a range of temperatures relevant to hotspot physics.

In this work, we extend the methodology used for RDX by Sakano et. al. [24] to the fast chemistry in TATB decomposition, which excludes the long time process of carbon cluster formation. We develop a 4 component (3 step) model based on classical molecular dynamics simulations using the ReaxFF forcefield[4, 37], and we utilize a fully temperature dependent specific heat model. From the initial model fitting procedure[24], we add a step to correct for ReaxFF's failure to conserve total energy at high temperatures, leading to energy drift. We also introduce a new algorithm for defining the time window of the non-

negative matrix factorization (NMF) analysis. We present comparisons between this model and MD simulations of a quasi-1D hotspot propagating through material over time.

This paper is organized as follows: we present, in detail, the different simulations methods used here, Section II, i.e. reactive molecular dynamics (MD) and continuum simulations as well as the corresponding equations and models. Then, in Section III, we present our unsupervised learning of the reduced-order model for TATB kinetics from isothermal reactive MD simulations. Section IV is dedicated to the calibration of heats of reaction using adiabatic reactive MD simulations. Section V contrasts the results of the different models generated here and provides an analysis of the characteristic timescales of the chemical decomposition steps. Finally, simulations of a quasi-1D hotspot propagation under different conditions are performed with both MD and continuum simulations, which are compared against each other in Section VI.

II. ATOMISTIC AND CONTINUUM SIMULATION SETUPS

In this work, we adopt and extend the approach previously used for RDX [24] that is based on reactive MD simulations of isothermal and adiabatic chemical decomposition using unsupervised learning methods to calibrate the kinetics of decomposition. We do it for a TATB single crystal. The MD simulation details are described in Section II A while Section II B is dedicated to continuum formulation of the problem.

A. Reactive molecular dynamics simulations

1. Reactive force field

Assessing the chemical decomposition of energetic materials through MD simulations necessitates the use of a reactive force field which is able to account for bond-breaking and formation phenomena over time. The reactive force field ReaxFF [37] is used in the present work. Covalent interactions are characterized using partial bond orders and electrostatics are described by computing at each time step new polarizable charges [38]. While multiple ReaxFF parametrizations are available in the literature, we use utilize 'ReaxFF-2018' [4] here. This force field has been previously employed to study shock initiation in TATB [13, 39], as well as also to widely investigate the

thermochemistry of various energetic materials at extreme conditions in temperature and pressure [40, 41]. We however emphasize the inherent limitations of using a classical reactive force field [42] for which the underlying functional forms limit its accuracy. Machine learning potentials offer a promising alternative to classical reactive force fields for the decomposition of high energy (HE) materials [43, 44], but our primary aim is to develop and validate a multiscale modeling approach. All the MD simulations were performed using the parallel MD software LAMMPS from Sandia National Laboratories [45, 46] with 3D periodic boundary conditions.

2. TATB single crystal preparation

The single crystal TATB supercell was obtained by using the same procedure as in a previous study [39]. The generalized crystal-cutting method (GCCM) [47] framework was used to create a quasi-orthorhombic supercell with cell vectors **A**, **B** and **C**, starting from the triclinic unit cell reported by Cady and Larson [48] with lattice vectors **a**, **b** and **c**. The resulting super-cell vectors are displayed in Figure 1, along with a snapshot of the super-cell for the homogeneous simulations considered in this work.

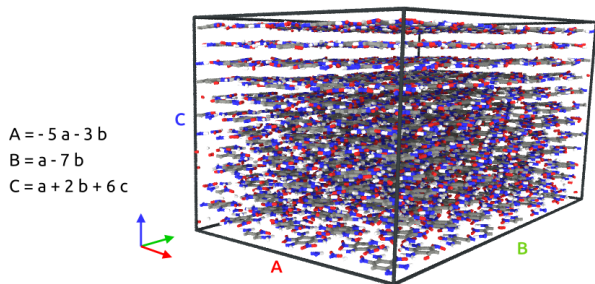


FIG. 1. Reactive MD simulation super-cell for the homogeneous simulations in the isothermal and adiabatic conditions.

All three cell angles lie within 0.5° from 90° . The resulting super-cell contained 456 molecules of TATB (10944 atoms) and was subsequently relaxed via atomic structure minimization. Finally, the cell is equilibrated in the isothermal-isobaric NPT ensemble at ambient conditions to ensure that the pressure is hydrostatic. Damping constants were set to 0.1 ps and 1.0 ps for both thermostat [49, 50] and barostat [51]. The final lengths along the periodic directions were $L_a=40.26 \text{ \AA}$, $L_b=69.58 \text{ \AA}$ and $L_c=35.87 \text{ \AA}$ with angles $\alpha=89.46^\circ$, $\beta=89.67^\circ$ and $\gamma=90.04^\circ$, leading to a crystal density of 1.94 g cm^{-3} . This super-cell was used for both isothermal and adiabatic

decomposition simulations.

3. Isothermal and adiabatic chemical decomposition simulations

Simulation of isothermal decomposition were performed for temperatures 1600-4000 K at 100 K intervals. An initial velocity distribution matching the desired temperature is applied over all atoms. Then, the system is coupled to a Nosé-Hoover [50] thermostat for the duration of the simulation. The thermostat damping constant and the simulation timestep were set to 100 fs and 0.1 fs, respectively. Simulations are continued until the perceived end of chemical reactions, at which point the system is considered at chemical equilibrium. Slow carbon clusters formation is not considered in this study, and we make the assumption that its associated characteristic time (spanning the microsecond domain) can be decoupled from the fast chemistry. It was found that initializing velocities to twice the targeted temperature (due to equipartition of energy) limited spurious oscillations in the potential energy at small times. This then allowed the target temperature to be reached in less than 0.1 ps, ensuring that very few chemical reactions would occur prior to thermal equilibrium. During reactions, we first observe a decrease of the potential energy of the system as chemical reactions are globally exothermic, followed by a plateau associated to the equilibrium state when the fast reactions are mostly complete. The total time of the simulations depends strongly on the temperature, and ranges between 30 ps and 2 ns. The evolution of the chemical bonding environment is analysed during the simulations, and provides the chemical fingerprints from which reduced-order terms for the kinetics are constructed.

The chemical decomposition under adiabatic conditions was studied in a second batch of reactive MD simulations. The samples were initialized at similar temperatures ranging between 1200 and 3000 K. Similarly, to reach the target initial temperature, atomic velocities are drawn from a distribution corresponding to twice the desired temperature. After a temperature dependent time delay (induction time), the system begins to react. The heat release from the reaction leads to a temperature increase, accelerating the chemical decomposition. The completion of chemical reaction corresponds to a plateau in the temperature evolution. However, it was found that the reactive force field used in the present work, even with a 0.1 fs, do not conserve the total energy accurately. Instead of reaching a temperature plateau once chemical reactions are complete,

a linear increase in temperature over time was systematically observed, and a linear drift in the time evolution of the total energy is also observed. As the heat of reaction of the reduced-order model is fit to the temperature evolution, this systematic drift would imply a bias in the optimization procedure. In order to avoid this bias, a procedure based on the conservation of total energy, described in Section IV B, is applied to re-scale the temperature-time evolution.

4. Thermal conductivity

Much work has been done in the past decades in predicting the thermal conductivity of TATB at the grain scale. Anisotropic thermal conductivity coefficients for both perfect and defective single crystals were calculated using classical MD simulations [52, 53] as well as its temperature-dependent form [54], which was used to study the anisotropic relaxation of idealized hotspots in crystalline TATB [55]. More recently, differences and implications of considering isotropic vs anisotropic thermal conductivity coefficients were discussed in a study [56] in which the authors use the shock trapping internal boundaries (STIB) method [15] to obtain the relaxation characteristics of hotspots generated by non-reactive MD simulations of shock-induced void collapse. In addition, experimental thermal conductivity measurements might not concern the single crystal but rather macroscopic samples [57–59]. Due to the large amount of data available in the literature, the thermal conductivity was not computed here, and both directional and temperature dependent descriptions are discussed in Section VI when comparing reaction propagation simulations.

5. Specific Heat

Modeling the temperature evolution using a diffusion-reaction model requires a specific heat for the material, and its variation with temperature. MD simulations, utilizing classical statistical mechanics, leads to a value of $\frac{1}{2}k_B T$ per degree of freedom. For molecular crystals, governed by quantum statistical mechanics, this classical value is a large over estimate at low to moderate temperatures, which can result in reduced accuracy in the continuum model. We used the same approach as for RDX [24] and computed the specific heat for each of the components of our reduced-order model based on quantum statistical mechanics. The contribution to the total specific heat C_v of the differ-

ent species belonging to each component was calculated from the roto-vibrational density of states of the corresponding atoms from NVT simulations using the following equation [60]:

$$C_v = k_B \int_0^\infty d\omega \cdot S(\omega) \cdot \frac{\hbar^2 \omega^2 \exp^{\hbar\omega/k_B T}}{(k_B T)^2 (1 - \exp^{\hbar\omega/k_B T})^2} \quad (1)$$

with k_B and \hbar being the Boltzmann and Planck constants, and $S(\omega)$ is the phonon density of states. $S(\omega)$ is calculated from the Fourier transform of the velocity auto-correlation function [61]. The specific heat for each chemical component was fit with the following analytic formulation:

$$C_{v,\text{component}i}(T) = \frac{1}{(1 + T^a)^b} \quad (2)$$

with a and b as the adjustable parameters. The corresponding evolution with temperature is provided in the Supplementary Material. It was found that their evolution was very close to the one of TATB. For this reason, we used the temperature-dependent TATB specific heat in every calculation that involves C_v : the heat evolved and the diffusion-reaction equations.

6. Simulation of a propagating 1D hotspot

Hotspot propagation simulations were performed using a larger supercell. We aim to compare the analytical diffusion-reaction model to reactive MD simulations for two directions of propagation, i.e. within the TATB layers and transverse to them. The supercell containing 456 molecules used for isothermal and adiabatic decomposition simulations is replicated 12 times along the **A** and 14 times along **C** direction for the within and transverse to layers, respectively, which leads to systems containing 5472 molecules and 6384 molecules. The length of the supercell in the **A** and **C** directions are 48.31 nm and 50.22 nm, respectively. The cross-section lengths remain unchanged. Periodic boundary conditions are applied in three directions.

In order to initiate the hot spot, a 10 nm portion of these samples is heated across 2 ps using a linear temperature ramp from 300 K to the target temperature, and then held for an additional 2 ps in the NVT ensemble. The remaining part of the sample is held at 300 K. The thermodynamic ensemble is then switched to the microcanonical ensemble (NVE) to assess the hot spot evolution under adiabatic

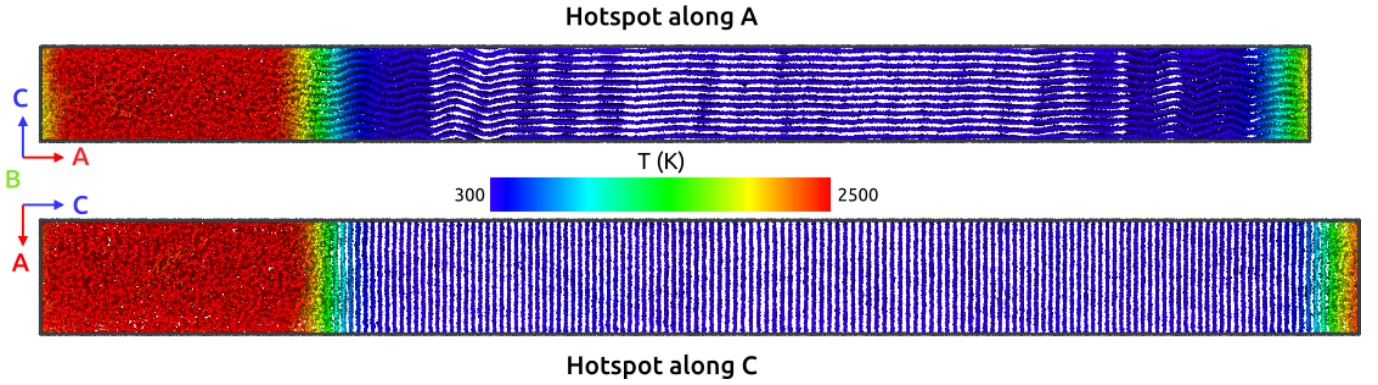


FIG. 2. MD supercell for hotspot simulations along the **A** and **C** directions. The snapshots are picked at the end of the NVT ramp + equilibration process ($t=4$ ps) for the case $T=2500$ K. Atoms are colored w.r.t. local temperature.

conditions. Snapshots of the two oriented samples described above with $T_0=2500$ K are displayed in Figure 2, taken at $t=4$ ps, i.e. at the end of the hotspot initialization process. It can be seen that for the sample with **A** parallel to the hotspot propagation direction, the non-zero pressure induced by hotspot thermal expansion is sufficient to activate the onset of a buckling instability, which requires very low deviatoric stress [34, 35, 62]. This may have implications on the value of the directional thermal conductivity coefficient as well as induce mechanochemical effects [15, 16], which will be discussed in Section VI.

B. Continuum model for diffusion coupled with chemical reactions problems

We describe in the following section the different systems of equations that are used to build the full model of diffusion coupled with chemical reactions at the continuum scale. We first present the Arrhenius laws for kinetics of chemical decomposition, then their coupling with the heat evolved equation for adiabatic conditions, and finally their coupling with the thermal diffusion equation.

1. Arrhenius laws for kinetics of chemical decomposition

We present a set of kinetics equations for modeling the evolution of a four component system that corresponds to a three steps decomposition mechanism $C_1 \mapsto C_2 \mapsto C_3 \mapsto C_4$ with C_i being the molar fraction of the i^{th} component. We choose the three chemical reactions to be first-order and irreversible. While C_1 and C_4 can be roughly described as TATB and reaction products, respectively, C_2 and C_3 are interpreted as various reaction intermediates. The cor-

responding kinetics equations are as follows:

$$\dot{C}_1 = -C_1 Z_a \exp\left(-\frac{E_a}{RT}\right) \quad (3a)$$

$$\dot{C}_2 = C_1 Z_a \exp\left(-\frac{E_a}{RT}\right) - C_2 Z_b \exp\left(-\frac{E_b}{RT}\right) \quad (3b)$$

$$\dot{C}_3 = C_2 Z_b \exp\left(-\frac{E_b}{RT}\right) - C_3 Z_c \exp\left(-\frac{E_c}{RT}\right) \quad (3c)$$

$$\dot{C}_4 = C_3 Z_c \exp\left(-\frac{E_c}{RT}\right) \quad (3d)$$

with R as the universal gas constant, Z_a, Z_b, Z_c are the kinetic prefactors, and E_a, E_b, E_c are the activation energies associated with the three chemical reactions. The calibration procedure for the kinetic prefactors and the activation energies is detailed in Section III.

2. Adiabatic conditions: heat evolution equation

Each chemical reaction absorbs or release heat depending on its endothermic or exothermic nature. The kinetics equations presented above determine the time evolution of the concentration of the different components. Using these concentrations, we can write the heat balance equation controlling the evolution of the temperature with time (assuming adiabatic conditions):

$$\rho C_v \dot{T} = -Q_1 \dot{C}_1 - Q_2 (\dot{C}_1 + \dot{C}_2) + Q_3 \dot{C}_4, \quad (4)$$

with ρ as the material density, C_v as the classical specific heat, and Q_1, Q_2, Q_3 as the heats of reaction. For RDX, the authors [24] used a temperature-dependent expression for the heats of reaction. In the present study this dependence is not as relevant, allowing for the use of constant heats of reaction.

3. Diffusion-reaction analytical formulation

Lastly, the reduced-order kinetic model is coupled with the diffusion equation, leading to:

$$\rho C_v \dot{T} = k \nabla^2 T - Q_1 \dot{C}_1 - Q_2(\dot{C}_1 + \dot{C}_2) + Q_3 \dot{C}_4, \quad (5)$$

with k as the thermal conductivity of the material. This model is implemented in a 1D (in-house) code and solved numerically using finite differences over a regular grid. The generalization of this model to the 2D and 3D cases is the purpose of further investigations. The simulations at the continuum scale are performed using the exact same dimensions as in MD simulations so as to allow direct comparison. To ensure that the temperature profile in the bar is strictly the same in both simulations, we fit the MD temperature profile obtained at the end of the hotspot formation time (i.e. 4 ps) using the following expression:

$$T_{\text{bar}}(\mathbf{x})|_{t=4 \text{ ps}} = \frac{(T_{\text{hotspot}} - 300)}{2} \left[\frac{\tanh(\mathbf{x} - a)}{\delta} - \frac{\tanh(\mathbf{x} - b)}{\delta} \right] + 300 \quad (6)$$

with a , b and T_{hotspot} as free parameters. This expression is then used as the initial temperature input for the continuum simulation. Finally, the temperature ramp and hold processes are performed in the continuum model using two isothermal steps:

$$T_{\text{MESO}}(\mathbf{x}, t) = \frac{t}{2} T_{\text{bar}}(\mathbf{x})|_{t=4 \text{ ps}} \text{ for } t \leq 2 \text{ ps}, \quad (7a)$$

$$T_{\text{MESO}}(\mathbf{x}, t) = T_{\text{bar}}(\mathbf{x})|_{t=4 \text{ ps}} \text{ for } 2 \leq t \leq 4 \text{ ps}, \quad (7b)$$

allowing for similar temperature evolution to MD while letting the concentrations of the chemical constituents evolve accordingly over the first 4 ps. Chemical reactions can then occur from the very beginning of the simulation during the ramp and equilibration period as they would in the MD. At the end of the 4 ps process, before removing the thermostat coupling and switching to adiabatic conditions, the concentrations may already have evolved in the hot and even in the adjacent, non-heated material.

III. UNSUPERVISED LEARNING OF ARRHENIUS LAWS FROM REACTIVE MD SIMULATIONS IN NVT

In this section, we describe the unsupervised learning approach adopted to identify a reduced-order chemical kinetics model for TATB using reactive MD simulations of isothermal decomposition over a wide range of temperatures. For each trajectory the analysis is performed over a particular time interval, and the following section describes

the procedure to extract this characteristic time from the evolution of the potential energy. The dimensionality reduction procedure is applied to the evolution of chemical environments and is detailed in Section III B. Finally, Section III C presents the calibration of Arrhenius parameters (pre-exponential factor and activation energies) for the kinetics of decomposition of TATB under isothermal conditions.

A. Extraction of characteristic time from the evolution of the potential energy

Isothermal decomposition simulations were performed for temperatures between 1200 K and 4000 K at every 100 K. These conditions led to a full chemical decomposition of the sample over time, involving a systematic decrease in potential energy (PE) of the system. These simulations were run over a sufficiently long time such that the PE reaches a plateau, as shown in Figure 3. The analysis of

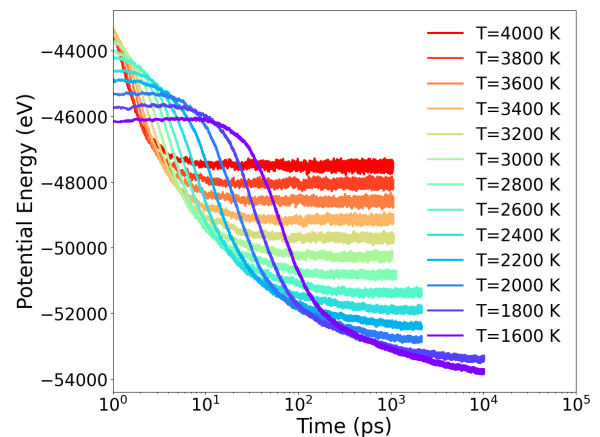


FIG. 3. Temporal evolution of the potential energy (in logscale) during the chemical decomposition of TATB single crystal under isothermal conditions for $T \in [1600; 4000 \text{ K}]$. Each color corresponds to different temperatures. It is seen that at $t=10 \text{ ns}$ the plateau is not completely reached for $T=1600 \text{ K}$.

the bonding environment is performed over a time interval that needs to be consistently defined so it can apply to any trajectory. The general procedure is presented in Figure 4 for one NVT trajectory and can be described over the following four steps:

- 1- The PE vs. time curve is log-transformed so that we now consider the evolution of PE with the logarithm of time.
- 2- The PE vs t is changed to $|\Delta PE|$ vs $\log(t - (1 - t_{cut}))$

where t_{cut} corresponds to the time before the data is considered, and $|\Delta PE|$ is the difference between the PE at $t = t_{cut}$ and the current value. t_{cut} corresponds to the induction time of the system during the reactive MD simulations and approximately equals 1 ps. The new curve appears as a tilted sigmoid as shown in Figure 4. A Richard’s curve with the following expression is fit to the data: $R(x) = \frac{A}{(1+e^{-B(x-M)})^\nu}$. This allows for the removal of noise in the MD trajectory.

- 3- In the third step, we use the analytical expression of the Richard’s curve and compute its first and second derivatives. The area of maximum rate of change (AMRC), defined as the interval over which an exponential growth is happening, is set as between the local extrema of the signal second derivative. In the present case, the AMRC represents the interval over which 90 % of the heat of reaction is released.
- 4- Finally, the data is transformed back to the initial units of Step 1, the AMRC now defining a characteristic time interval for subsequent analysis.

The aim of the process described above is to consistently identify a characteristic time t_{AMRC} for each temperature. We then define the total analysis time window as $t_{induction} + k t_{AMRC}$, where $t_{induction}$ equals the time interval between t_{cut} and the beginning of the AMRC window, highlighted in yellow in Figure 4. Using $k=2.5$ appears to be a good compromise between getting enough information at low temperatures (that require longer simulation times) and computational cost. The identification of the time window over which the chemical reactions monitoring should be performed allows to define a frequency of analysis. This frequency was set to lead to 100 snapshots over the entire time window previously defined.

B. Dimensionality reduction of the time evolution of bonding environments

At the mesoscale, thermo-chemistry models generally use a small number of components for describing the chemical evolution of the system. Thus, the key point in the reduced-order model resides in the ability to identify as few components as possible that reproduces faithfully the full evolution of the chemical environment. The approach previously adopted for RDX [24] involves three steps:

- (i) use appropriate descriptors of the chemical environment,

- (ii) use data science methods to perform the dimension reduction while retaining a high physical meaning of the components evolution,

- (iii) identify kinetics equation to describe the chemical decomposition simulated using reactive MD simulations

which are all described in detail in the following subsections.

1. Chemical environment descriptors

The use of physically sound descriptors is highly valuable for machine learning applications, in particular for building Machine Learning Potentials [63–65]. However, these descriptors generally rely on a geometrical analysis of atomic environments whereas here we look at describing their chemical nature, i.e. their bonding properties. Sakano *et al* [24] chose a descriptor able to generally reproduce the molecules that decompose and form during reactive MD simulations. However, describing the entire set of molecules in the system would lead to very large dimension spanned by the decomposition of CHNO systems. The approach used in the present work, similar to the work on RDX [24] is a Coordination Geometry Analysis (CGA), that involves each system atom’s bonding environment. It derives from the classic group additivity method developed by Benson to calculate heats of formation and various properties [66, 67]. In order to overcome the difficulty of analysing short life of radicals that appear at high temperatures, the authors proposed to retain only the number and types of atoms bonded to a given central atom as a fingerprint of its local chemical environment. This simplified descriptor has the additional advantage of possessing a finite and constant dimension, a recommended property for subsequent machine learning treatment.

In the case of TATB for example, the initial structure of a gas-phase single molecule would be described as 6 C[CCN], 3 N[CHH], 3 N[COO], 6 H[N], 6 O[N] where the atom outside of brackets is the center. As stated in the previous work[24], conventional organic chemistry indicates that each atom would not possess more than four bonds and that H would preferentially be single bonded. For CHNO systems, these assumptions lead to 280 possible combinations of coordination geometries. We used the same convention here for the evolution of bonding environment populations, that is, its representation through a matrix \mathbf{X} with 280 columns by N_t rows, with N_t corresponding to the number of time frames at which the bond-

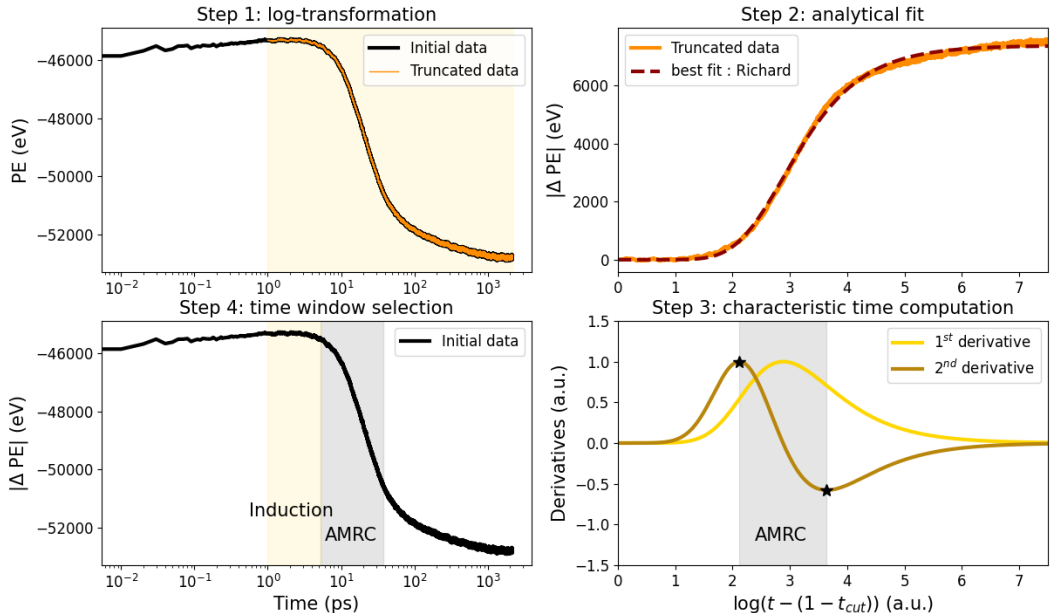


FIG. 4. Extraction procedure of a characteristic time in a consistent way for NVT simulations.

ing environments are assessed. ReaxFF intrinsically offers the ability to track the bond order (BO) between atoms. Here, as for RDX, we determine bonds as atom pairs with a $BO \geq 0.5$. A fine tracking of bonding changes, bond network, or reaction descriptions could also be investigated to provide some insight into the elementary steps of chemical reactions occurring at the atomic level [25], however, this is not the aim of the present study.

2. Reduced-order model using non-negative matrix factorization

Assuming the time evolution of the bonding environment is known (through the \mathbf{X} matrix), we aim at reducing the number of components from 280 to a low number tractable for a continuum formulation. Various dimensionality reduction techniques are available in the literature, from linear methods like principal component analysis [68, 69] to non linear methods like auto-encoders [70, 71]. Linear dimensionality reduction techniques are faster and more averse to overfitting than nonlinear methods, but they are not able to identify components that are nonlinear functions of the input variables. However, they can be used to project the original data into a linear subspace that contains the relevant components. As both methods reproduce the axis of principal variance of the input data, they can lead to neg-

ative values among the corresponding eigenvalues. However, physical processes often have non-negative contributions, that can be viewed as additional constraints in the dimensionality reduction procedure. Moreover, they cannot often be represented by linear combinations of input variables. Non-negative matrix factorization (NMF) [72] is a mathematical technique that decomposes a given matrix into two non-negative matrices, one of which represents the approximate values of the original matrix’s components, and the other represents the approximate weights of the components in the original matrix. The non-negativity constraint of the factorization allows the method to capture the additive nature of the data and to model the parts-based representation of the data. This method is particularly useful in physics because it allows for the identification of underlying physical processes that are characterized by non-negative contributions.

In the present case, the initial bonding information is contained in the \mathbf{X} matrix with dimensions $(N_t \times 280)$. Using the NMF technique, \mathbf{X} can be reduced into a product of two sub-matrices \mathbf{W} and \mathbf{H} of dimensions $(N_t \times N_c)$ and $(N_c \times 280)$ respectively, with N_c being the dimension of the projection space. This way, \mathbf{W} rows represent the time evolution of the N_c components while the rows of matrix \mathbf{H} maps the population distribution of the 280 initial components into the reduced components N_c . The

number of dimensions of the reduced space, i.e. the number of components N_c , is fixed a priori. The efficiency of the dimensionality reduction procedure is quantitatively measured by the root mean square error (RMSE) criterion between the original data \mathbf{X} and the reconstructed data from its low-dimensional representation $\tilde{\mathbf{X}} = \mathbf{W} \times \mathbf{H}$. Increasing N_c leads to a decrease in the RMSE, as less data is lost during compression, thus a wise choice of N_c would result in the best compromise between a compact and an accurate model.

We show in Figure 5 the RMSE of the reduced model as a function of the number of components N_c and temperature. The total error of the model is computed between the raw data of the reactive MD simulations, i.e. the actual population of bonding environments \mathbf{X} , and the reconstructed population $\tilde{\mathbf{X}} = \mathbf{W} \times \mathbf{H}$, averaged over time and 280 components. Interestingly, the dependence on tem-

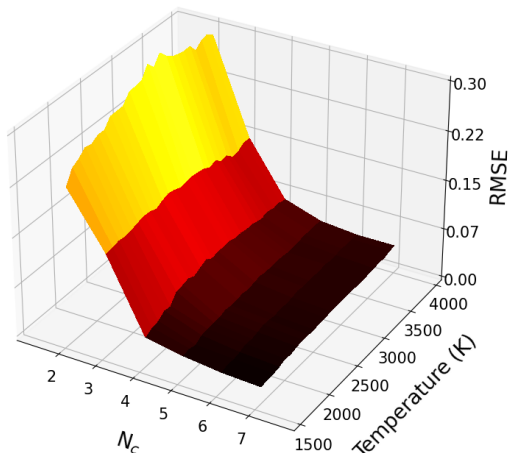


FIG. 5. Total RMSE of the fitted data using NMF as a function of the number of components N_c and temperature on the 280 bonding environments.

perature of the RMSE is much less pronounced compared to its dependence on the number of components. Not surprisingly, increasing the number of reduced components N_c drastically reduces the total RMSE. While the total RMSE with $N_c=3$ seems moderately acceptable, it drops considerably when $N_c=4$. Increasing N_c to larger values does not significantly improve the reconstruction error. Hence, here we will calibrate and develop the full thermochemistry mesoscopic model with 4 components.

We recall that each component obtained using the NMF are a combination of the 280 bonding environments; this combination defines their chemical identity, which remains

fixed over time. Only the relative concentration between the components can evolve with time.

After the NMF is performed, one important additional step is still required. Since we aim to reproduce the time evolution of the concentrations of the different components along the chemical decomposition, as concentrations are usually expressed as proportions, their total sum should be normalised to one. However, the matrix \mathbf{W} , that contains the time evolution of the N_c components is not normalized and can exhibit values greater than 1. Normalisation is achieved by computing for each component the sum of its population over time:

$$\mathbf{Pop}[i] = \sum_J^{280} \mathbf{H}[i, J] \quad (8)$$

The next step consists in propagating this normalization into matrix \mathbf{W} :

$$\mathbf{W}[:, i] = \mathbf{Pop}[i] \times \mathbf{W}[:, i] \quad (9)$$

allowing for both a constant sum over time of the concentrations and the same order of magnitude in the different components. Finally, the data is re-scaled between 0 and 1 through:

$$\mathbf{W}[:, i] = \frac{\mathbf{W}[:, i]}{\sum_c \mathbf{W}[:, c]} \quad (10)$$

which ensures that all the reduced components sum to 1. As an example in Figure 6 we present the time evolution of component concentrations, as well as the time evolution of the raw data extracted from the reactive MD simulations of an isothermal decomposition of TATB at $T=1600$ K. The time has been truncated to 500 ps to focus on the time window where the intermediates exhibit their maximum value. The different components obtained through NMF exhibit a time evolution that is physically consistent with concentrations evolution of reactants, intermediates, and products for chemical decomposition processes.

C. Kinetics calibration from the reduced-order model

The rate equations corresponding to a 4 components model are given in Equations 3. They contain parameters associated with the kinetics prefactors and activation energies of the corresponding chemical equations. In order to calibrate these parameters on the MD trajectories, we

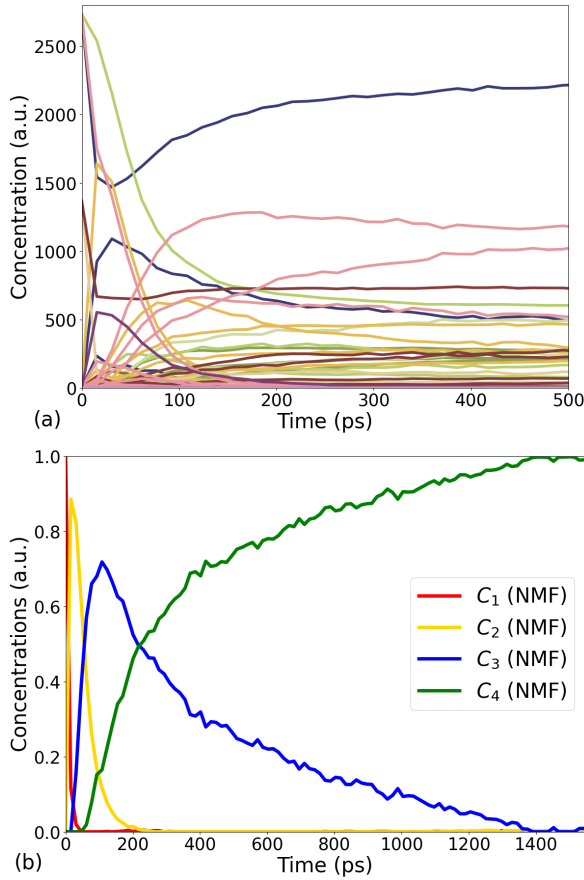


FIG. 6. Example of dimensionality reduction using NMF. The raw data from the reactive MD simulation is presented in (a) while (b) represents the normalized concentrations evolution obtained with 4 components. Reactants and final products are represented with red and green lines respectively while intermediates are represented with blue and yellow lines.

first define the following Arrhenius laws:

$$\beta(T) = Z_a \exp\left(\frac{-E_a}{RT}\right) \quad (11a)$$

$$\gamma(T) = Z_b \exp\left(\frac{-E_b}{RT}\right) \quad (11b)$$

$$\delta(T) = Z_c \exp\left(\frac{-E_c}{RT}\right) \quad (11c)$$

where Z_a , Z_b , Z_c and E_a , E_b , E_c are the prefactors and activation energies. For each isothermal simulation, the four components concentrations evolution obtained via the NMF procedure are fitted by the following analytical solu-

tions of Equation 3:

$$C_1(t) = \exp(-\beta t) \quad (12a)$$

$$C_2(t) = \frac{\beta}{\gamma - \beta} [\exp(-\beta t) - \exp(-\gamma t)] \quad (12b)$$

$$C_3(t) = \frac{\beta\gamma}{(\delta - \beta)(\delta - \gamma)(\gamma - \beta)} \frac{[(\delta - \gamma)\exp(-\beta t) + (\gamma - \beta)\exp(-\delta t) - (\delta - \beta)\exp(-\gamma t)]}{(\delta - \beta)(\delta - \gamma)(\gamma - \beta)} \quad (12c)$$

$$C_4(t) = 1 - C_1(t) - C_2(t) - C_3(t). \quad (12d)$$

These equations are obtained by solving the differential equations 3. The fit is performed through mean-square minimization using the differential evolution function of the scipy python package, leading to the optimization of β , γ and δ for each temperature. An example of the resulting fits is given in Figure 7-a (in dotted lines) for a decomposition at $T = 1600$ K. Applying the same procedure independently for each temperature allows to extract the evolution of the three quantities, displayed in Figure 7-b. Reformulating the Arrhenius laws of Equation 11 clarifies the temperature dependence of these quantities:

$$T \ln \beta(T) = T \ln Z_a - \frac{E_a}{R} \quad (13a)$$

$$T \ln \gamma(T) = T \ln Z_b - \frac{E_b}{R} \quad (13b)$$

$$T \ln \delta(T) = T \ln Z_c - \frac{E_c}{R} \quad (13c)$$

Indeed, $T \ln \beta(T)$, $T \ln \gamma(T)$ and $T \ln \delta(T)$ are linear functions of temperature, with slopes $\ln Z_a$, $\ln Z_b$ and $\ln Z_c$ and intercepts $-\frac{E_a}{R}$, $-\frac{E_b}{R}$ and $-\frac{E_c}{R}$. These linear expressions are fitted to the discrete data from individual fits, as shown in Figure 7-b, and the values obtained for the four components model are presented in Table II. Hence, these expressions can now be used to predict the values of β , γ and δ at any temperature, and reconstruct the concentration evolution along an isothermal trajectory. This prediction is displayed in Figure 7-c using dashed lines, and exhibits a good agreement with MD data. The computed RMSE is equal to 0.1257, slightly larger than 0.1139 obtained from a direct fit of the data. Thus, we can consider that the kinetics laws from Equation 3 with kinetics parameters and activation energies given in Table II reproduce accurately the concentration evolution of the different constituents of TATB decomposition under isothermal conditions.

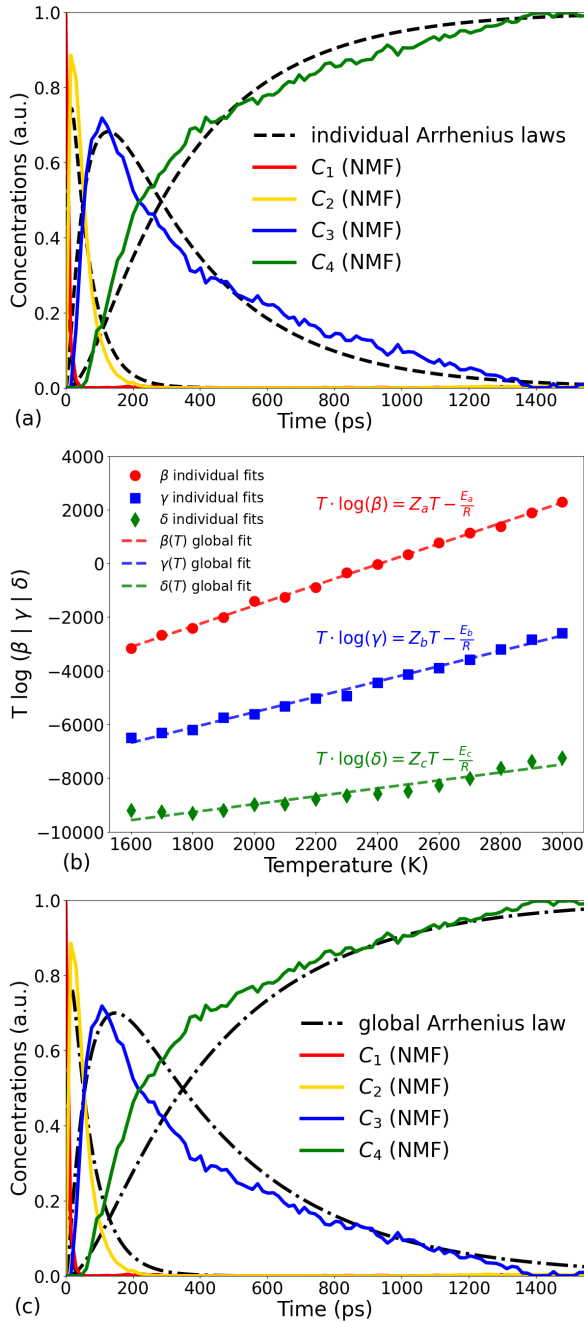


FIG. 7. Individual to global fits procedure with a focus on the case $T = 1600$ K. (a) Individual fits of 4 components. (b) Identification of global expression for the Arrhenius laws. (c) Comparison of the 4 components evolution with time using the global Arrhenius laws.

IV. HEATS OF REACTION CALIBRATION OVER REACTIVE MD ADIABATIC DECOMPOSITION SIMULATIONS

A. Calibration and evaluation

We focus in this section on the calibration of heats of reaction, Q_i , needed for our reduced-order chemical kinetics model (see equation 4). Each considered chemical reaction is associated to some heat release (possibly positive or negative), the sum of all heats of reaction being equal to the total exothermicity - or endothermicity - of the global decomposition (in the case of TATB the global decomposition is exothermic). The heat of reaction, which is the amount of heat absorbed or released during a chemical reaction, then plays a crucial role in determining the temperature change that occurs during adiabatic decomposition. We can then calibrate the heat of reaction so as to reproduce the temperature evolution along MD trajectories. Adiabatic decomposition reactive MD simulations were performed with initial temperatures ranging from 1200 K to 3000 K in 50 K steps. Using the kinetics prefactors and activation energies determined previously on isothermal decomposition, and integrating the evolution of the concentration of each component using the kinetic model, the heats of reaction are calibrated on the entire set of adiabatic decomposition simulations. The optimization is conducted so as to minimize a loss function defined as the normalized RMSE between the MD observed temperatures and the predicted ones:

$$\mathcal{L}_{\mathcal{T}} = \frac{1}{M} \sum_{j=1}^M \text{RMSE}(\tilde{\mathbf{T}}_j^{\text{model}}, \tilde{\mathbf{T}}_j^{\text{MD}}) \quad (14)$$

where $\tilde{\mathbf{T}}_j = \mathbf{T}_j / (\max \mathbf{T}_j^{\text{MD}} - \min \mathbf{T}_j^{\text{MD}})$ is the normalized vector representing all temperatures at the different times of the simulation and M is the total number of trajectories. This corresponds to **model 1** in the remainder of this paper.

Instead of considering separately the optimizations for the kinetic model and of the heat of reaction, both sets of parameters can be calibrated at the same time on adiabatic simulations only. The loss function $\mathcal{L}_{\mathcal{T}}$ of the optimization procedure is once again defined as the normalized RMSE between the observed and predicted temperature evolution. This corresponds to **model 2** for the rest of the article.

Finally, a third model is proposed, **model 3**, for which all sets of parameters are calibrated on adiabatic decomposition using a loss function defined by the normalized

RMSE on temperature evolution \mathcal{L}_T and on the RMSE on the concentration evolution. The latter corresponds to the difference between the predicted concentrations and the concentrations given by the NMF reduction of the MD data:

$$\mathcal{L}_C^{NVE} = \frac{1}{N_c} \sum_c \frac{1}{M} \sum_{j=1}^M \text{RMSE}(\mathbf{C}_{c,j}^{\text{model}}, \mathbf{C}_{c,j}^{\text{NMF}}) \quad (15)$$

where $\mathbf{C}_{c,j}$ is the vector representing all concentrations of component c at different times of the simulation, and M is the total number of trajectories. In all cases the optimization is performed using a differential evolution algorithm. A summary of the three models used in this paper is proposed in Table I. The three models are then evaluated using

Model	Calibration of the		Calibration of
	kinetic model parameters	heats of reaction	
model 1	NVT	NVE	\mathcal{L}_T
model 2	NVE	NVE	\mathcal{L}_T
model 3	NVE	NVE	$\mathcal{L}_T + \mathcal{L}_C^{NVE}$

TABLE I. Summary of the three models used in this paper.

the same metric as the ones used in the loss function, i.e. the normalized RMSE on the temperatures and concentrations. Results are presented in the next section. Prior to calibrating the different models, a procedure is proposed to correct the temperature obtained during the adiabatic decomposition trajectories. Indeed, we observe a drift in the evolution of the total energy with time during and after the chemical reaction. The amplitude of the drift depends on the value of the integration timestep but never vanishes, even for timestep as low as 0.05 fs. This energy drift artificially heats up the system, hence biasing our optimization procedure. It is necessary, in order to develop accurate heats of reaction, to un-bias the MD simulations by applying a temperature correction.

B. Temperature correction of reactive MD involving energy non conservation

We present a simple procedure to correct the total energy and the temperature evolution during reactive NVE simulations. The total energy is increasing linearly with time, as evidenced in figure 8-a. This total energy drift has consequences on the temperature of the system (see figure 8-b). The long time behavior of the system still exhibits a continuous increase of temperature whereas chemical reactions are supposedly completed. Under the assumption that the drift in total energy follows the equipartition principle, we apply the following temperature correction at each

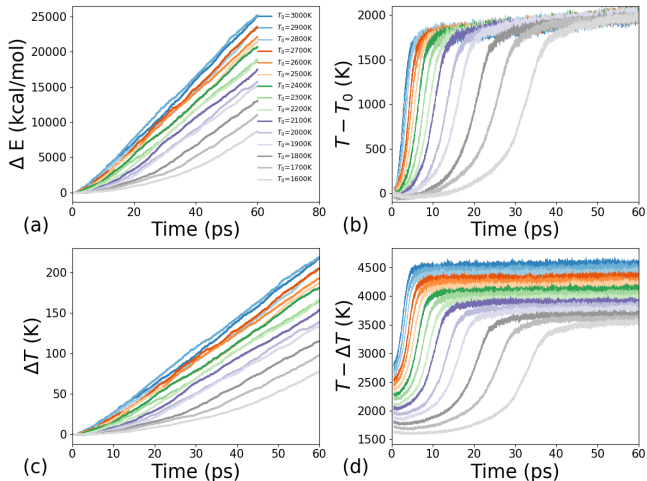


FIG. 8. Correction process of temperature vs time from the total energy drift using ReaxFF for reactive MD simulations of adiabatic decomposition in the NVE ensemble.

time t :

$$\Delta T(t) = 0.5\Delta E(t)/(\rho C_v) \quad (16)$$

where $\Delta E(t)$ is the energy drift at time t . This correction in temperature, to be removed from the MD temperature, is displayed in Figure 8-c. This correction may be negligible during the first steps of the reaction but can reach a few hundreds of kelvin after 100 picoseconds. Finally, the corrected temperature profiles are shown in Figure 8-d, exhibiting a plateau once chemistry is completed. We consider these temperature curves as reference data for the calibration of the three models.

V. RESULTS OF THE DIFFERENT MODELS

We discuss the accuracy of the 3 models in this section. In addition, an analysis of the chemical constituents of the intermediates of reaction is provided.

A. Performances of the models

The performances of the models is evaluated using three loss functions: the loss of the temperature evolution on adiabatic simulations \mathcal{L}_T , the loss of the relative concentrations of the components along adiabatic decompositions \mathcal{L}_C^{NVE} , and an equivalent loss along isothermal decomposition simulations \mathcal{L}_C^{NVT} . As these losses are dimensionless, they can be summed to form a total loss $\mathcal{L} = \frac{1}{3}(\mathcal{L}_T + \mathcal{L}_C^{NVE} + \mathcal{L}_C^{NVT})$. The parameters of the ki-

netics part ($\{Z_i, E_i\}$) and the heats of reaction ($\{Q_i\}$) are optimized using a differential evolution algorithm with results provided in table II.

In order to provide a visual comparison between the models, error results are displayed as bar plots in figure 9. As expected, **Model 2** (fitted only on \mathcal{L}_T) performs well on \mathcal{L}_T but appears unable to predict the relative concentrations of the constituents in either NVT or NVE simulations. Its total loss remains greater than the two others models. **Model 1** and **Model 3** exhibit equivalent performances on the three losses, and their overall score \mathcal{L} is equivalent, although their optimization is performed using different loss functions. It is remarkable to note that **Model 3**, fit only on adiabatic simulations, is able to predict accurately the relative concentration of the components along isothermal simulations. Its transferability is indeed ensured by including both temperature and concentration in its loss function. Additionally, the performances of the models also need to

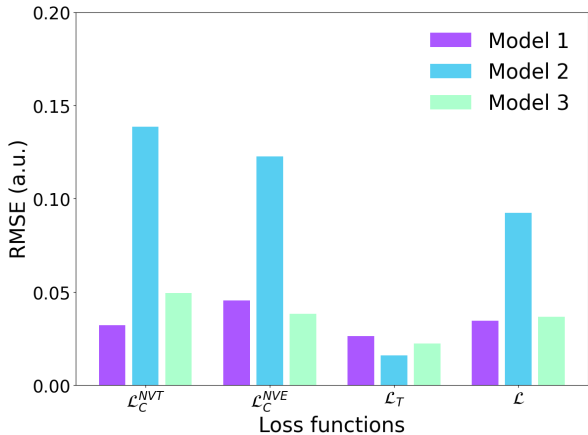


FIG. 9. Individual RMSE’s computed using the different loss functions of Equations 14 and 15, for the three models considered in this work.

be balanced with their associated CPU cost, in particular to the cost of the data required for their optimization. In this context, **Model 3** would offer the best ratio performance over cost, as adiabatic simulations are shorter than isothermal simulations. Indeed, NVE simulations would typically be run over a few hundred picoseconds, whereas isothermal decomposition at low temperature would require a few nanoseconds of simulation to be completed. We present in Figure 10 the result of **Model 3** for representative simulations of decomposition under isothermal and adiabatic conditions. The relative concentrations of the components are compared to the NMF results on separate rows corresponding to NVT (first row) and NVE (last row) simulations. The middle graphs represents the temperature

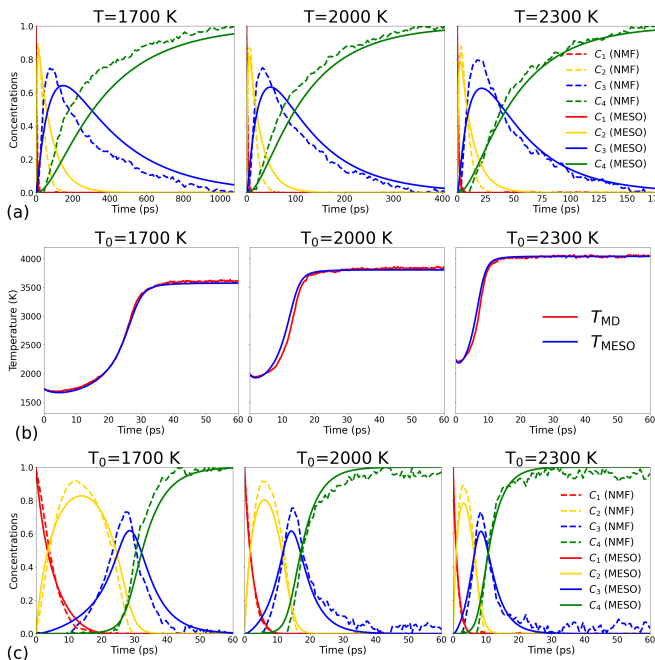


FIG. 10. Concentrations evolution in NVT (a) and NVE (c) as well as temperature evolution in NVE (b) for different initial temperatures : (left) $T=1700$ K, (center) $T=2000$ K, (right) $T=2300$ K.

evolution along NVE simulations predicted by the model and computed by MD simulations. The overall agreement between the model and MD simulations is good for both the temperature and relative concentrations. Moreover, the characteristic timescale of the intermediate reactions are well reproduced by the model. The entire set of curves is provided in the Supplementary Material.

B. Characteristic timescales of the chemical decomposition

Beyond the capabilities of the model to accurately reproduce MD simulations, it is also interesting to evaluate the ability of the NMF method to extract the salient feature of the chemical decomposition mechanism. Indeed, NMF is a mathematical method, agnostic to chemistry, and the separation into components is somehow arbitrary. We provide below a more detailed chemical analysis of the components given by the NMF method and compared them with an analysis of molecular population. The implementation of ReaxFF in LAMMPS allows for the explicit extraction of the number of current molecules over time, along the MD simulation, which we display in figure 11. On the other hand, the NMF decomposition gathers all the bond

Model	step	Z_i [ps^{-1}]	E_i [kcal/mol]	Q_i [kcal]	\mathcal{L}_C^{NVT}	\mathcal{L}_C^{NVE}	\mathcal{L}_T	\mathcal{L}
1	$C_1 \mapsto C_2$	44.23	18.19	-2.8719×10^5	0.0321	0.0455	0.0263	0.0346
	$C_2 \mapsto C_3$	17.46	22.38	1.7054×10^6				
	$C_3 \mapsto C_4$	04.40	23.72	2.6597×10^5				
2	$C_1 \mapsto C_2$	48.80	26.76	-3.4907×10^5	0.1386	0.1227	0.0159	0.0924
	$C_2 \mapsto C_3$	26.19	20.59	1.8214×10^6				
	$C_3 \mapsto C_4$	01.32	13.92	2.1019×10^5				
3	$C_1 \mapsto C_2$	58.18	19.09	-1.4687×10^5	0.0494	0.0383	0.0224	0.0367
	$C_2 \mapsto C_3$	19.98	24.75	1.7157×10^6				
	$C_3 \mapsto C_4$	06.59	25.88	1.1136×10^5				

TABLE II. Kinetics parameters and heats of reaction obtained for the different models considered in the present work along with the corresponding computed losses. \mathcal{L}_C^{NVT} and \mathcal{L}_C^{NVE} correspond to the loss function of Equation 15 applied on concentrations of NVT and NVE trajectories, respectively while \mathcal{L}_T corresponds to the loss function of Equation 14 applied on temperature of NVE trajectories. The global loss function \mathcal{L} corresponds to the average of the three losses previously defined.

environments in only four components. The first and last component of the NMF decomposition are associated with the TATB molecule and to reaction products, respectively. The NMF has also allocated a number of bond environments to the second and third component C_2 and C_3 , associated with different characteristic timescales. In the following we tentatively correlate the components of the NMF decomposition to a few common molecular species, showing that the NMF decomposition indeed corresponds to a chemically distinct processes. For this purpose, we compute the first time at which 80% of the maximum concentration is reached as t_{min} . Similarly, t_{max} is defined as the last time at which 90% of the maximum concentration is observed. As an example, figure 11 displays the evolution of the concentration of the NMF components superimposed with the concentration of some molecular species. t_{min} and t_{max} are obtained as the abscissa of the intersection of the 90% horizontal line with the different curves.

A detailed analysis of the TATB chemical decomposition mechanisms is beyond the scope of this article, we select only a few molecular species to illustrate the different characteristic timescales of the decomposition process. As evidenced in figure 11, the early production of water molecules occurs on the same timescale as C_2 . It can then be deduced that the first step of the decomposition mechanism[73], i.e. the release of water molecules, is indeed encoded in the second component of the NMF. Similarly, the appearance of the CHNO molecule, resulting from the breaking of the aromatic carbon ring, occurs on a different timescale and correlates to C_3 . Finally, the production of CO_2 molecules, a final product of decomposition, is clearly associated to the increase of C_4 . We

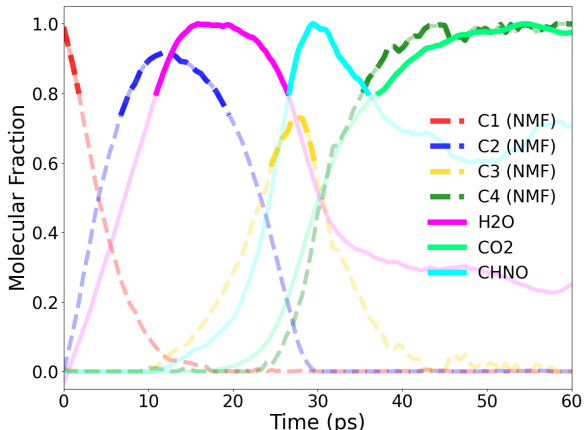


FIG. 11. Evolution of the normalized molecules number (divided by their maximum value) as a function of time. Superimposed are the NMF concentrations. This corresponds to an adiabatic decomposition with an initial temperature set to 1700 K.

conclude from these observations that the NMF splitting is related to the main steps of the decomposition mechanisms of TATB. NMF not only provides a reduced framework for modeling the chemical decomposition of TATB but also can be used to some extent as an analysis tool at a coarser scale. NMF provides access to the population of each of its constituents through matrix **Pop** of Equation 8. The population of the 10 predominant molecular species present in the reactants, intermediates and products for an adiabatic decomposition with $T_0=1700$ K are reported in Table **S1** of the Supplementary Material where the results are in good agreement with the data presented in Figure 11.

In the following we extend the previous analysis to different temperatures, in order to confirm if the overall scheme

of decomposition remains unchanged. We extract the characteristic timescales as t_{min} and t_{max} for the different molecular species and NMF components, and follow their evolution with temperature. Results are displayed in figure 12. The timescales of the NMF components evolution

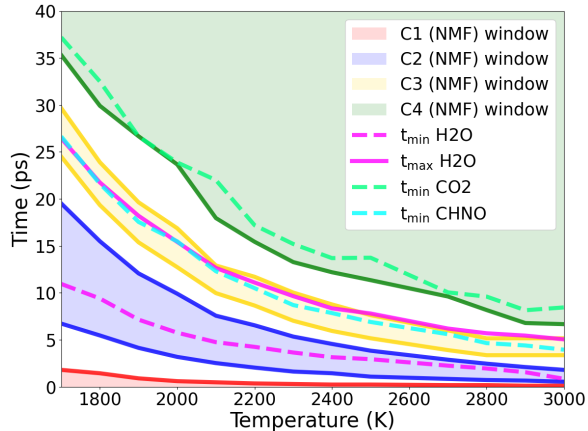


FIG. 12. Evolution of individual components concentrations and molecular species characteristic times with temperature. Continuous lines and colored areas correspond to time intervals of individual concentrations, i.e., red, blue, gold and green for the reactants, two intermediates and final products. Dashed lines correspond to minimum and maximum (when applicable) characteristic times of molecular species, i.e. purple, green and cyan for H₂O, CO₂ and CHNO species.

τ^{C_i} are represented as colored envelopes defined by the time at which the concentration is above 90% of its maximum value : $\tau^{C_i} = t_{max}^{C_i} - t_{min}^{C_i}$. For clarity, the characteristic timescales of the molecular species are only reported as t_{min} (solid curves) and t_{max} (dashed curves).

From this representation, a clear picture of the different timescales and their evolution with temperature emerges: firstly, at every temperature, the chemical steps of the TATB decomposition, given by the evolution of some representative molecular species corresponding to the steps analysed at T=1700K: there is no abrupt change in the evolution of characteristic timescales with temperature. Besides, the NMF timescales evolves comparably to the ones given by the molecular analysis: the NMF, performed independently at each temperature, is a robust analysis tool with respect to different thermodynamic conditions. Finally, the overall characteristic timescales exhibit a non linear monotonic decrease as the temperature increases, highlighting the role of temperature in the chemical rates of decomposition.

We conclude this part recalling the two main observa-

tions obtained through the comparison between the NMF splitting and a molecular analysis. Firstly, NMF indeed captures the key steps of the chemical decomposition and their associated characteristic timescales. Secondly, the decomposition mechanism remains unchanged as the temperature increases, but the chemical reaction rates increase monotonically.

VI. MODEL VALIDATION: HOT SPOT EVOLUTION

Simulations of one dimensional hot spot using the set up described in section II have been performed for several initial hot spot temperatures and for two directions of propagation, namely along the **A** and **C** directions, as defined in Figure 2. This corresponds to an infinite hot spot in the two transverse directions. The total simulation time was set to 360 ps, with a timestep of 50 fs and the sample was discretized with 100 elements each approximately equal to 25 Å. The quantum corrected specific heat of TATB was used in all continuum simulations, as for the calibration process. In the continuum model, the thermal conductivity of TATB is fixed to $k_C = 0.0.52$ for the simulations with a hotspot propagating along **C**. This value corresponds to the value linearly extrapolated to an infinite-length supercell as explained in [53]. However, the anisotropic coefficient along the **A** direction was not set to 1.07, the computed value of [53], due to multiple reasons. Indeed, using the anisotropic thermal conductivity coefficient for **A** led to worse results in comparison to MD simulations. This is a consequence of the compressive waves emitted by the hot spot inducing large deformation in the material, that trigger the elastic buckling instability. The large out-of-plane displacement of the TATB layers ultimately induce a rotation of the crystal, leaving a more isotropic structure, both in terms of elastic and transport properties. Additional work is in progress to analyse this effect in detail, and will be the subject of a coming article. For the propagation along **A**, the thermal conductivity coefficient was fitted to reproduce at best the average temperature profile for a hotspot set up at 1900 K and the value obtained was $k_A = 0.61$, within both experimental and computational ranges. This means that pure one-to-one comparisons are done for the two hot spots with initial temperatures of 1900 and 2500 K along **C**, in a predictive way. Concerning the direction **C**, the case at 1900 K helps to calibrate the conductivity, while the case at 2500 K is for comparison to the MD profile.

Comparisons are provided for representative adiabatic

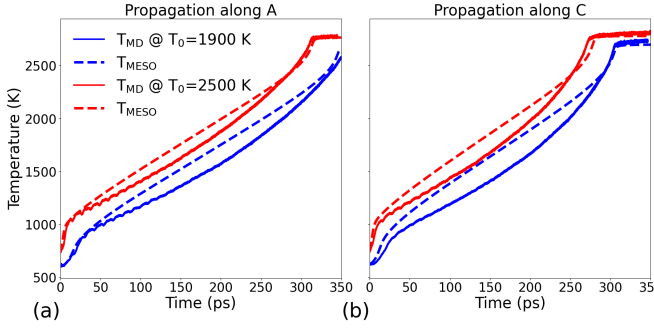


FIG. 13. Calibration of heats of reactions for the 4 components model. (a) and (b) panels stand for hotspots propagating along **A** and **C** directions, respectively.

simulations at initial temperatures equal to 1900 and 2500 K along the two directions. Note that the MD average temperatures have been corrected with the same procedure as described above. Figure 13 displays the time evolution of the temperature averaged on the whole system (the first 4 ps where the system is heated are not shown). In each case we observe a good agreement between the predictions of the mesoscopic model and the MD results. The overall shape of the curves exhibit a similar behavior, and the plateaus of the temperature are well reproduced. However, after the hot spot becomes critical, the continuum model predicts a linear evolution of temperature with time, which is due to the inability of the model to take into account volumetric compression and non linearity.

In order to get a deeper comparison between the model and MD simulations, the evolution of the temperature profile with time is computed and displayed on a time-space diagram (see figure 14). In MD simulations temperatures are averaged along slices perpendicular to the hot spot propagation, and displayed with a colormap. The conic shape is the signature of the propagation of a reactive wave forward and backward in the periodic material (the two waves are equivalent). The slope of these lines, i.e. the deflagration velocity, is a result of the competition between chemical reactions that tend to increase the temperature of the material, and the diffusion term. In each case we see an almost linear shape which can be associate with the stationary behavior of the deflagration wave. Nevertheless some tiny differences persist: in the initial hot spot the temperature is systematically larger than the temperature predicted by the model. This is a consequence of the non conservation of energy in MD simulations, as the rescaling procedure cannot be applied to a non homogeneous system (the correction should depend on the local temper-

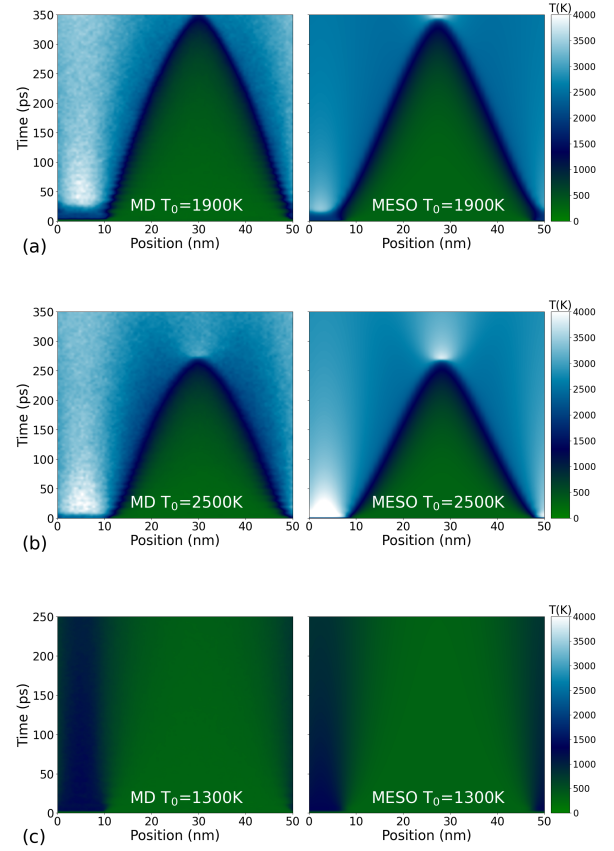


FIG. 14. Time-space diagrams of temperature. (a), (b), (c) concern the directions **C**, **A**, **C** with hotspot initial temperature set to 1900 K, 2500 K and 1300 K respectively. The MD results are displayed on the left side while the right side concerns the continuum model.

ature). The error committed by MD predictions is of the order of a few hundreds kelvins. It has a small impact on the early behavior of the deflagration wave which exhibits a concave shape instead of a pure linear behavior. Overall the behavior between the atomistic simulations and the continuum model is quantitatively similar for every conditions of temperature and every direction of propagation. While Figure 14a and Figure 14b concern conditions where the hotspots go critical, an additional one-to-one comparison was done under non-critical conditions. Figure 14c concerns a hotspot propagating along **C** at an initial temperature of 1300 K. While MD leads to a small error on temperature evolution with time, both simulation methods indicate that the hotspot quenches and does not lead to a deflagration wave. This last comparison gives confidence in the mesoscopic model for being able to catch the hotspot criticality which will be the object of future work.

VII. CONCLUSION AND PERSPECTIVES

We have developed a continuum model describing both chemistry and heat transport of TATB single crystal. The chemical part is based on a classical Arrhenius framework and the heat transport is described by the heat diffusion equation. We employed atomistic reactive simulation to extract the parameters of the model, and proposed different loss functions for their optimization. The best model is obtained for a loss function including the temperature and concentration evolution along adiabatic simulations. As the total energy is not strictly conserved during adiabatic simulations, we proposed a rescaling procedure based on the equipartition of energy. The rescaled temperatures are subsequently used as reference data in the loss function. The concentration of reactants, intermediates and final products are obtained by means of a dimensionality reduction technique, the non-negative matrix factorization. This method allows to extract the minimum number of components to reproduce the atomistic decomposition trajectories. In the case of TATB, four components are needed for a quantitative rendering of concentration evolution along adiabatic and isothermal decompositions. The parameters of the kinetic part as well as the heats of reaction are all optimized using a differential evolution algorithm so as to minimize the multi-objective loss function. Then the predictions of the model are compared to direct reactive simulations on a large temperature interval, and show a very good agreement for temperature and concentrations evolution during both isothermal and adiabatic decompositions. This demonstrated that the mesoscale thermochemistry model is able to reproduce the decomposition of TATB under different temperature conditions.

Besides, a comparison of the NMF splitting with a molecular species analysis reveals that NMF indeed captures the main steps of the decomposition mechanisms and their associated timescales. The chemical rates increase with temperature (i.e. the characteristic timescales decrease), but the overall decomposition mechanisms remain unchanged.

The kinetic model is coupled to a diffusion equation, and one to one comparisons with atomistic simulations are performed for one dimensional hot spot propagation along different crystalline directions and at different initial temperatures. We show that the behavior of the hot spot is well predicted by the continuum model, the tiny reported differences being attributed to the non conservation of energy in MD simulations and the simplicity of the model

which does not consider either spherical or deviatoric mechanical deformations. Finally, as only adiabatic simulations are needed to calibrate the model, an extension to a large pressure range would be straightforward in order to be able to consider non equilibrium conditions such as shock loadings. Further work will also include the effect of dimensionality on the hot spot formation, investigating the relation between the mechanical distortions of the structure and its chemical reactivity.

ACKNOWLEDGMENTS

Paul Lafourcade would like to personally thank Dr. Alizée Dubois for numerous discussions and for her help concerning analytical aspects of the study.

This research was partially supported by the US Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-20-2-0189. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory, or the U.S. Government.

Approved for unlimited release: LA-UR-23-23237.

REFERENCES

-
- [1] CA Handley, BD Lambourn, NJ Whitworth, HR James, and WJ Belfield. Understanding the shock and detonation response of high explosives at the continuum and meso scales. *Applied Physics Reviews*, 5(1):011303, 2018.
- [2] Brenden W Hamilton, Michael N Sakano, Chunyu Li, and Alejandro Strachan. Chemistry under shock conditions. *Annual Review of Materials Research*, 51:101–130, 2021.
- [3] Brenden W Hamilton and Timothy C Germann. Energy localization efficiency in 1, 3, 5-trinitro-2, 4, 6-triaminobenzene pore collapse mechanisms. *Journal of Applied Physics*, 133(3):035901, 2023.
- [4] M. A. Wood, D. E. Kittell, C. D. Yarrington, and A. P. Thompson. Multiscale modeling of shock wave localization in porous energetic material. *Phys. Rev. B*, 97:014109, Jan 2018.
- [5] Ryan A Austin, Nathan R Barton, John E Reaugh, and Laurence E Fried. Direct numerical simulation of shear localization and decomposition reactions in shock-loaded hmx crystal. *Journal of Applied Physics*, 117(18):185902, 2015.
- [6] Matthew P Kroonblawd and Laurence E Fried. High explosive ignition through chemically activated nanoscale shear bands. *Physical Review Letters*, 124(20):206002, 2020.
- [7] MJ Cawkwell, Thomas D Sewell, Lianqing Zheng, and Donald L Thompson. Shock-induced shear bands in an energetic molecular crystal: Application of shock-front absorbing boundary conditions to molecular dynamics simulations. *Physical Review B*, 78(1):014107, 2008.
- [8] Md Mahbulul Islam and Alejandro Strachan. Role of dynamical compressive and shear loading on hotspot criticality in rdx via reactive molecular dynamics. *Journal of Applied Physics*, 128(6):065101, 2020.
- [9] Chunyu Li, Brenden W Hamilton, Tongtong Shen, Lorena Alzate, and Alejandro Strachan. Systematic builder for all-atom simulations of plastically bonded explosives. *Propellants, Explosives, Pyrotechnics*, 47(8):e202200003, 2022.
- [10] Chunyu Li, Brenden W Hamilton, and Alejandro Strachan. Hotspot formation due to shock-induced pore collapse in 1, 3, 5, 7-tetranitro-1, 3, 5, 7-tetraoctane (hmx): Role of pore shape and shock strength in collapse mechanism and temperature. *Journal of Applied Physics*, 127(17):175902, 2020.
- [11] Yen T Nguyen, Dilki Perera, Puhao Zhao, Tommy Sewell, and HS Udaykumar. Head-to-head comparison of molecular and continuum simulations of shock-induced collapse of an elongated pore in an energetic molecular crystal. *Propellants, Explosives, Pyrotechnics*, 47(8):e202200016, 2022.
- [12] Mitchell A Wood, Mathew J Cherukara, Edward M Kober, and Alejandro Strachan. Ultrafast chemistry under nonequilibrium conditions and the shock to deflagration transition at the nanoscale. *The Journal of Physical Chemistry C*, 119(38):22008–22015, 2015.
- [13] Brenden W Hamilton, Matthew P Kroonblawd, and Alejandro Strachan. Extemporaneous mechanochemistry: Shock-wave-induced ultrafast chemical reactions due to intramolecular strain energy. *The Journal of Physical Chemistry Letters*, 13(29):6657–6663, 2022.
- [14] Craig M Tarver, Steven K Chidester, and Albert L Nichols. Critical conditions for impact-and shock-induced hot spots in solid explosives. *The Journal of Physical Chemistry*, 100(14):5794–5799, 1996.
- [15] Brenden W. Hamilton, Matthew P. Kroonblawd, Chunyu Li, and Alejandro Strachan. A hotspot’s better half: Non-equilibrium intra-molecular strain in shock physics. *The Journal of Physical Chemistry Letters*, 12(11):2756–2762, 2021. PMID: 33705143.
- [16] Brenden Hamilton and Alejandro Strachan. Many-body mechanochemistry: Intra-molecular strain in condensed matter chemistry. *chemrxiv preprint chemrxiv-2022-bvdmp*, 2022.
- [17] Brenden W Hamilton and Timothy C Germann. Interplay of mechanochemistry and material processes in the graphite to diamond phase transformation. *arXiv preprint arXiv:2302.04684*, 2023.
- [18] Evan J Reed, JD Joannopoulos, and Laurence E Fried. Electronic excitations in shocked nitromethane. *Physical Review B*, 62(24):16500, 2000.
- [19] Evan J Reed, M Riad Manaa, Laurence E Fried, Kurt R Glaesemann, and JD Joannopoulos. A transient semimetallic layer in detonating nitromethane. *Nature physics*, 4(1):72–76, 2008.
- [20] H Keo Springer, Sorin Bastea, Albert L Nichols III, Craig M Tarver, and John E Reaugh. Modeling the effects of shock pressure and pore morphology on hot spot mechanisms in hmx. *Propellants, Explosives, Pyrotechnics*, 43(8):805–817, 2018.
- [21] Matthew P Kroonblawd and Ryan A Austin. Sensitivity of pore collapse heating to the melting temperature and shear viscosity of hmx. *Mechanics of Materials*, 152:103644, 2021.
- [22] Puhao Zhao, Sangyup Lee, Tommy Sewell, and HS Udaykumar. Tandem molecular dynamics and continuum studies of shock-induced pore collapse in tatb. *Propellants, Explosives, Pyrotechnics*, 45(2):196–222, 2020.
- [23] Camilo A Duarte, Chunyu Li, Brenden W Hamilton, Alejandro Strachan, and Marisol Koslowski. Continuum and molecular dynamics simulations of pore collapse in shocked

- β -tetramethylene tetranitramine (β -hmx) single crystals. *Journal of Applied Physics*, 129(1):015904, 2021.
- [24] Michael N. Sakano, Ahmed Hamed, Edward M. Kober, Nicolo Grilli, Brenden W. Hamilton, Md Mahbulul Islam, Marisol Koslowski, and Alejandro Strachan. Unsupervised learning-based multiscale model of thermochemistry in 1,3,5-trinitro-1,3,5-triazinane (rdx). *The Journal of Physical Chemistry A*, 124(44):9141–9155, 2020. PMID: 33112131.
- [25] Edward M. Kober. Developing reaction chemistry models from reactive molecular dynamics: Tatb. *Propellants, Explosives, Pyrotechnics*, 47(8):e202100386, 2022.
- [26] Kibaek Lee, Kaushik Joshi, Santanu Chaudhuri, and D Scott Stewart. Mirrored continuum and molecular scale simulations of the ignition of high-pressure phases of rdx. *The Journal of chemical physics*, 144(18):184111, 2016.
- [27] Jean-Bernard Maillet, Laurent Soulard, and Gabriel Stoltz. A reduced model for shock and detonation waves. ii. the reactive case. *Europhysics Letters*, 78(6):68001, 2007.
- [28] Joseph M. Zaug, Ryan A. Austin, Michael R. Armstrong, Jonathan C. Crowhurst, Nir Goldman, Louis Ferranti, Cheng K. Saw, Raymond A. Swan, Richard Gross, and Laurence E. Fried. Ultrafast dynamic response of single-crystal β -hmx (octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine). *Journal of Applied Physics*, 123(20):205902, 2018.
- [29] Yen Nguyen, Pradeep Seshadri, Oishik Sen, D. Barrett Hardin, Christopher D. Molek, and H. S. Udaykumar. Multi-scale modeling of shock initiation of a pressed energetic material i: The effect of void shapes on energy localization. *Journal of Applied Physics*, 131(5):055906, 2022.
- [30] Nirmal Kumar Rai, Oishik Sen, and H. S. Udaykumar. Macro-scale sensitivity through meso-scale hotspot dynamics in porous energetic materials: Comparing the shock response of 1,3,5-triamino-2,4,6-trinitrobenzene (tatb) and 1,3,5,7-tetranitro-1,3,5,7-tetrazocane (hmx). *Journal of Applied Physics*, 128(8):085903, 2020.
- [31] Shobhan Roy, Belinda P. Johnson, Xuan Zhou, Yen T. Nguyen, Dana D. Dlott, and H. S. Udaykumar. Hot spot ignition and growth from tandem micro-scale simulations and experiments on plastic-bonded explosives. *Journal of Applied Physics*, 131(20):205901, 2022.
- [32] Paul Lafourcade, Christophe Denoual, and Jean-Bernard Maillet. Dislocation core structure at finite temperature inferred by molecular dynamics simulations for 1,3,5-triamino-2,4,6-trinitrobenzene single crystal. *The Journal of Physical Chemistry C*, 121(13):7442–7449, 2017.
- [33] Paul Lafourcade, Christophe Denoual, and Jean-Bernard Maillet. Detection of plasticity mechanisms in molecular dynamics: Application to tatb single crystal. *AIP Conference Proceedings*, 1979(1):180005, 2018.
- [34] Paul Lafourcade, Christophe Denoual, and Jean-Bernard Maillet. Irreversible deformation mechanisms for 1,3,5-triamino-2,4,6-trinitrobenzene single crystal through molecular dynamics simulations. *J. Phys. Chem. C*, 122:14954–14964, 2018.
- [35] Paul Lafourcade, Christophe Denoual, and Jean-Bernard Maillet. Mesoscopic constitutive law with nonlinear elasticity and phase transformation for the twinning-buckling of tatb under dynamic loading. *Phys. Rev. Mater.*, 3:053610, 2019.
- [36] Paul Lafourcade, Jean-Bernard Maillet, Nicolas Bruzy, and Christophe Denoual. Elastic anisotropy of 1,3,5-triamino-2,4,6-trinitrobenzene as a function of temperature and pressure: A molecular dynamics study. *Propellants, Explosives, Pyrotechnics*, 47(8):e202100384, 2022.
- [37] Adri C. T. van Duin, Siddharth Dasgupta, Francois Lorant, and William A. Goddard. Reaxff: A reactive force field for hydrocarbons. *The Journal of Physical Chemistry A*, 105(41):9396–9409, 2001.
- [38] H. M. Aktulga, J. C. Fogarty, S. A. Pandit, and A. Y. Grama. Parallel reactive molecular dynamics: Numerical methods and algorithmic techniques. *Parallel Computing*, 38:245–259, 2012.
- [39] Brenden W. Hamilton, Matthew P. Kroonblawd, Md Mahbulul Islam, and Alejandro Strachan. Sensitivity of the shock initiation threshold of 1,3,5-triamino-2,4,6-trinitrobenzene (tatb) to nuclear quantum effects. *The Journal of Physical Chemistry C*, 123(36):21969–21981, 2019.
- [40] Brenden W Hamilton, Brad A Steele, Michael N Sakano, Matthew P Kroonblawd, I-Feng W Kuo, and Alejandro Strachan. Predicted reaction mechanisms, product speciation, kinetics, and detonation properties of the insensitive explosive 2, 6-diamino-3, 5-dinitropyrazine-1-oxide (llm-105). *The Journal of Physical Chemistry A*, 125(8):1766–1777, 2021.
- [41] M. M. Islam and A. Strachan. A reactive molecular dynamics simulations to investigate the shock response of liquid nitromethane. *J. Phys. Chem. C*, 123:2613, 2019.
- [42] Thomas P Senftle, Sungwook Hong, Md Mahbulul Islam, Sudhir B Kylasa, Yuanxia Zheng, Yun Kyung Shin, Chad Junkermeier, Roman Engel-Herbert, Michael J Janik, Hasan Metin Aktulga, et al. The reaxff reactive force-field: development, applications and future directions. *npj Computational Materials*, 2(1):1–14, 2016.
- [43] P Yoo, M. Sakano, S Desai, M.M. Islam, P Liao, and A. Strachan. Neural network reactive force field for c, h, n, and o systems. *npj Comput. Mater.*, 7:9, 2021.
- [44] Brenden W. Hamilton, Pilsun Yoo, Michael N. Sakano, Md Mahbulul Islam, and Alejandro Strachan. High-pressure and temperature neural network reactive force field for energetic materials. *The Journal of Chemical Physics*, 158(14):144117, 04 2023.
- [45] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1–19, 1995.

- [46] Aidan P. Thompson, H. Metin Aktulga, Richard Berger, Dan S. Bolintineanu, W. Michael Brown, Paul S. Crozier, Pieter J. in 't Veld, Axel Kohlmeyer, Stan G. Moore, Trung Dac Nguyen, Ray Shan, Mark J. Stevens, Julien Tranchida, Christian Trott, and Steven J. Plimpton. Lammps - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications*, 271:108171, 2022.
- [47] Matthew P. Kroonblawd, Nithin Mathew, Shan Jiang, and Thomas D. Sewell. A generalized crystal-cutting method for modeling arbitrarily oriented crystals in 3d periodic simulation cells with applications to crystal-crystal interfaces. *Computer Physics Communications*, 207:232–242, 2016.
- [48] H. H. Cady and A. C. Larson. The crystal structure of 1,3,5-triamino-2,4,6-trinitrobenzene. *Acta Crystallographica*, 18(3):485–496, Mar 1965.
- [49] Shuichi Nosé. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.*, 81(1):511–519, 1984.
- [50] William G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31(3):1695–1697, 1985.
- [51] Glenn J. Martyna, Douglas J. Tobias, and Michael L. Klein. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.*, 101(5):4177–4189, 1994.
- [52] Matthew P. Kroonblawd and Thomas D. Sewell. Theoretical determination of anisotropic thermal conductivity for crystalline 1,3,5-triamino-2,4,6-trinitrobenzene (tatb). *The Journal of Chemical Physics*, 139(7):074503, 2013.
- [53] Matthew P. Kroonblawd and Thomas D. Sewell. Theoretical determination of anisotropic thermal conductivity for initially defect-free and defective tatb single crystals. *The Journal of Chemical Physics*, 141(18):184501, 2014.
- [54] Matthew P. Kroonblawd and Thomas D. Sewell. Predicted anisotropic thermal conductivity for crystalline 1,3,5-triamino-2,4,6-trinitrobenzene (tatb): Temperature and pressure dependence and sensitivity to intramolecular force field terms. *Propellants, Explosives, Pyrotechnics*, 41(3):502–513, 2016.
- [55] Matthew P. Kroonblawd and Thomas D. Sewell. Anisotropic relaxation of idealized hot spots in crystalline 1,3,5-triamino-2,4,6-trinitrobenzene (tatb). *The Journal of Physical Chemistry C*, 120(31):17214–17223, 2016.
- [56] Matthew P. Kroonblawd, Brenden W. Hamilton, and Alejandro Strachan. Fourier-like thermal relaxation of nanoscale explosive hot spots. *The Journal of Physical Chemistry C*, 125(37):20570–20582, 2021.
- [57] B.D. Faubien. Thermal conductivity of tatb and tatb blends by differential scanning calorimetry. *Report No. MHSMP-76=30C*, 1976.
- [58] R.H. Cornell and G.L. Johnson. Measuring thermal diffusivities of high explosives by the flash method. [lx-04, lx-07, lx-09, lx-10, lx-14, and rx-03-bb]. 10 1978.
- [59] J F Baytos. Specific heat and thermal conductivity of explosives, mixtures, and plastic-bonded explosives determined experimentally. 9 1979.
- [60] J P Hague. Determining the phonon density of states from specific heat measurements via maximum entropy methods. *Journal of Physics: Condensed Matter*, 17(15):2397, apr 2005.
- [61] Peter H. Berens, Donald H. J. Mackay, Gary M. White, and Kent R. Wilson. Thermodynamics and quantum corrections from molecular dynamics for liquid water. *The Journal of Chemical Physics*, 79(5):2375–2389, 1983.
- [62] Paul Lafourcade, Christophe Denoual, and Jean-Bernard Maillet. Elastic instability in graphite single crystal under dynamic triaxial compression: Effect of strain-rate on the resulting microstructure. *J. Appl. Phys.*, 128(4):045101, 2020.
- [63] Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.*, 98:146401, Apr 2007.
- [64] Albert P. Bartók, Mike C. Payne, Risi Kondor, and Gábor Csányi. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.*, 104:136403, Apr 2010.
- [65] A.P. Thompson, L.P. Swiler, C.R. Trott, S.M. Foiles, and G.J. Tucker. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *Journal of Computational Physics*, 285:316–330, 2015.
- [66] N. Cohen and S. W. Benson. Estimation of heats of formation of organic compounds by additivity methods. *Chemical Reviews*, 93(7):2419–2438, 1993.
- [67] Leonidas Constantinou and Rafiqul Gani. New group contribution method for estimating properties of pure compounds. *AIChE Journal*, 40(10):1697–1710, 1994.
- [68] N. Halko, P. G. Martinsson, and J. A. Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review*, 53(2):217–288, 2011.
- [69] Per-Gunnar Martinsson, Vladimir Rokhlin, and Mark Tygert. A randomized algorithm for the decomposition of matrices. *Applied and Computational Harmonic Analysis*, 30(1):47–68, 2011.
- [70] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006.
- [71] Matthias Scholz, Martin Fraunholz, and Joachim Selbig. Nonlinear principal component analysis: neural network models and applications. In *Principal manifolds for data visualization and dimension reduction*, pages 44–67. Springer Berlin Heidelberg, 2008.
- [72] Cédric Févotte and Jérôme Idier. Algorithms for Nonnegative Matrix Factorization with the β -Divergence. *Neural Computation*, 23(9):2421–2456, 09 2011.
- [73] Christine J Wu and Laurence E Fried. Ring closure medi-

ated by intramolecular hydrogen transfer in the decomposition of a push-pull nitroaromatic: Tatb. *The Journal of*

Physical Chemistry A, 104(27):6447–6452, 2000.