Figure 1. ReaxFF predictions of the probabilities of $O_2 + H_2 + (M) \rightarrow H + HO_2 + (M)$ (dashed lines) and $2O_2 + H_2 \rightarrow 2HO_2$ (dotted lines) occurring as the initial reaction for composition ratios (a) $[H_2]/[O_2] = 0.25$, (b) $[H_2]/[O_2] = 0.5$, (c) $[H_2]/[O_2] = 1$, and (d) $[H_2]/[O_2] = 2$ at 3,000 K (blue), 4,000 K (green), and 5,000 K (red). The lines are drawn to guide the eye. The error was calculated using $error = Z \sqrt{\frac{[Pr(1 - Pr)]}{N}}$, where $Z$ was set to 1.645 to ensure 90% level of confidence, $Pr$ is the reaction probability, and $N$ is the number of trajectories for a given set of initial conditions. In most cases ensembles of 50 trajectories were run, but in some cases an additional 20 trajectories were run to lower the error. The average error is ± 0.09.
Figure 2. (a) Pressure-density isotherms ($P$-$\rho$, $T$) for $H_2$ predicted by ReaxFF (solid curves) compared to those obtained using the empirical equation reported by Lemmon et al. (dashed curves) at 300 K (red), 500 K (green) and 1,000 K (blue). (b) The relative errors of the ReaxFF results for the $H_2$ pressure-density isotherms at 300 K (red), 500 K (green) and 1,000 K (blue).
Figure 3. (a) Pressure-density isotherms ($P$-$\rho$, $T$) for O$_2$ predicted by ReaxFF (solid curves) compared to those obtained using the empirical equation reported by Kuznetsov et al. (dashed curves) 300 K (blue), 500 K (green), 1,000 K (red), 2,000 K (black), 3,000 K (purple), and 4,000 K (orange). (b) The relative errors of the ReaxFF pressure-density isotherms, with the same color scheme for the temperatures used in frame (a).
Figure 4. The average frequencies of O₂⋯H₂ (red curves and circle points) and O₂⋯H₂⋯O₂ (purple curves and square points) collisions as functions of density for the composition ratio of [H₂]/[O₂] = 0.25 at 3,000 K (dashed curves and blue points) and 4,000 K (solid curves and green points). The red lines are linear least-square fits of the O₂⋯H₂ points and the purple curves are nonlinear least-square fits of the O₂⋯H₂⋯O₂ data.
Figure 5. The ratios of termolecular to bimolecular collision frequencies as functions of density for composition ratio [H₂]/[O₂] = 0.25 at 3,000 K (blue) and 4,000 K (green). The curves are nonlinear least square fits of the points to Eq. (25). The equations of the curves are 

\[ \frac{ZT}{ZB} = -(2.121 \times 10^{-5})^2 \rho + (0.0106)^2 \rho \] for 3,000 K and 4,000 K, respectively.
Figure 6. ReaxFF predictions of the density dependence of the average collision time for O₂⋯O₂ (red) and O₂⋯H₂ (purple) collisions for composition ratio [H₂]/[O₂] = 0.25 at 3,000 K (dashed lines and blue squares) and 4,000 K (solid lines and green squares). The curves are linear least-squares fits to the points. The slopes of the lines correspond to τ_{H₂⋯O₂} are -0.0016 and -0.0017 at 3,000 K and 4,000 K, respectively; and that for τ_{O₂⋯O₂} are -0.0829 and -0.0899 at 3,000 K and 4,000 K, respectively. All units are in (fs m³/kg). The average collisions durations monotonically decrease with density. This is because at elevated density collision pairs are more likely to be interrupted by a third species, thus lowering the collision duration for the pair.
Figure 7. The characteristic collision frequency ratios for O$_2$····H$_2$···O$_2$ and O$_2$···H$_2$ collisions as functions of density for composition ratio [H$_2$]/[O$_2$] = 0.25 at 3,000 K (blue) and 4,000 K (green).

The points here were obtained by dividing the points in Fig. (5) by $\sqrt{T}$. The curves are nonlinear least-squares fits of the points to Eq. (26).
On Identifying Collisions of Various Molecularities in Molecular Dynamics Simulations

Homayoon Rafatijo, M. Monge-Palacios,* and Donald L. Thompson*

Department of Chemistry, University of Missouri-Columbia, Columbia, Missouri 65211-7600

ABSTRACT

We present a method based on kinetic molecular theory that identifies reactions of various molecularities in molecular dynamics (MD) simulations of bulk gases. The method allows characterization of the thermodynamic conditions at which higher than bimolecular reactions are a factor in the mechanisms of complex gas-phase chemistry. Starting with Bodenstein’s definition of termolecular collisions we derive analytical expressions for the frequency of higher molecularity collisions. We have developed a relationship for the ratio of the frequencies of termolecular to bimolecular collisions in terms of the temperature, density, and collision times. To demonstrate the method we used ReaxFF in LAMMPS to carry out MD simulations for NVT ensembles of mixtures of H₂:O₂ over the density range 120.2 kg m⁻³ to 332.7 kg m⁻³ and temperature range 3,000 K to 5,000 K. The simulations yield ReaxFF-based predictions of the relative importance of termolecular collisions O₂⋯H₂⋯O₂ and bimolecular collisions O₂⋯H₂ in the early chemistry of hydrogen combustion.

*Current address: King Abdullah University of Science and Technology (KAUST), Clean Combustion Research Center (CCRC), Physical Science and Engineering (PSE), Thuwal 23955-6900, Saudi Arabia
1. INTRODUCTION

The phase space data produced in molecular dynamics (MD) simulations of many-atom systems contain complete classical descriptions of the processes determined by the force field for the initial conditions. The usefulness of the simulation results comes from gleaning physically meaningful information from the mass of phase space points. This is particularly challenging for simulations of combustion chemistry, and especially so for conditions at high pressures where the usual concepts of elementary reactions may not apply. This is an area of practical and fundamental interest, and one where new approaches are needed.

Recently we proposed using a *reaction cluster* construct in MD simulations to identify “independent” chemical reactions in the chemical soup that evolves as processes such as hydrocarbon combustion progress.\(^1\) Reaction clusters are defined in bulk simulations using proximity criteria to identify those species (atoms, molecules, and radicals) within distances of the atoms of a breaking bond such that they can potentially participate chemically or even physically in the reaction. This identifies participants of reactions of any molecularity. We have illustrated application of reaction clusters in a demonstration of a generalization of Tolman’s concept\(^2\) of activation energy for short-time thermal reactions\(^1\) of hydrogen combustion.

In 1922 Bodenstein\(^3\) proposed a definition of termolecular collisions. He assumed that they are indirect collisions in which a bimolecular complex BC is formed before the arrival of a third species A. He assumed that the key factor determining the frequency is the average lifetime of the BC complex, which he estimated to be twice the time needed for a hard-sphere impact. In 1927 Tolman\(^4\) introduced a more realistic model. Assuming that there is a region that is impenetrable for real B···C collisions due to repulsive forces he proposed a length parameter that defines the interaction range. However, he did not explicitly define the length parameter in terms
of forces. In 1932 Kassel\textsuperscript{5} presented a model that took into account complexes due to near-orbiting trajectories that are interrupted by a third-body species. These early definitions of termolecular collisions have been the basis of work on processes such as radical recombination, but they are not explicitly used because of the primitive treatment of dynamics. The basic aspects of the models are substantiated by realistic MD simulations of reactions.

We present a method to compute the fundamental kinetic parameters such as collision frequency and duration from MD simulations of bulk gases. We use the cluster representation\textsuperscript{6} for gases to generalize the concepts of Bodenstein,\textsuperscript{3} Tolman,\textsuperscript{4} and Kassel\textsuperscript{5} to define collisions of various molecularities. The collision frequency and complex lifetimes depend on the intermolecular forces, and can be computed directly from MD simulations for collisions of any molecularity. The method we present can be used to determine the role of high molecularity collisions in the macroscopic behavior of overall reactions at given thermodynamic conditions.

We use the approximate force field ReaxFF\textsuperscript{7-10} as a convenient tool to carry out MD simulations of the early chemistry of hydrogen combustion to illustrate the method. The simulations yield ReaxFF-based predictions of the relative importance of termolecular collisions $\text{O}_2 \cdots \text{H}_2 \cdots \text{O}_2$ and bimolecular collisions $\text{O}_2 \cdots \text{H}_2$ that promote initial reactions $2\text{O}_2 + \text{H}_2 \rightarrow 2\text{HO}_2$ and $\text{O}_2 + \text{H}_2 \rightarrow \text{H} + \text{HO}_2$, respectively. Because our purpose was to illustrate a theoretical method we have not attempted to validate ReaxFF for this chemistry.

2. EXTENSION OF BODENSTEIN’S DEFINITION OF TERMOLECULAR COLLISIONS TO HIGHER MOLECULARITIES

We have extended Bodenstein’s\textsuperscript{3} definition of termolecular collisions based on the cluster representation\textsuperscript{6} for gases to define collisions of various molecularities. In the cluster
representation a gaseous system is considered as a mixture of clusters of different molecularities at any instant of time. An $n$-species cluster results from an $n$-species collision. A collision occurs when any two species interact, which begins when the atoms of the two species approach to within “interaction range.”

According to Bodenstein’s definition the frequency of $A\cdots B\cdots C$ collisions ($A$, $B$, and $C$ can be atoms, molecules, or radicals) is the sum of three bimolecular collision frequencies; that is,

$$Z_{ABC} = Z_{AB,C} + Z_{BC,A} + Z_{CA,B} = \sum_{ij,k} Z_{ij,k} ; \quad (1)$$

where $i$, $j$, and $k$ represent the three species involved in the collision, and $Z_{ij,k}$, is the frequency of collisions between one of the species, $k$, and a pair, $ij$. The frequency of bimolecular collisions between species $i$ and $j$ can be expressed as the product of a temperature dependent factor $f_{ij}(T)$ and concentrations; that is,

$$Z_{ij} = f_{ij}(T)[i][j]. \quad (2)$$

Therefore, Eq. (1) becomes

$$Z_{ABC} = f_{AB,C}(T)[AB][C] + f_{BC,A}(T)[BC][A] + f_{CA,B}(T)[CA][B] = \sum_{ij,k} f_{ij,k}(T)[ij][k] . \quad (3)$$

Assuming that the $ij$ pair is in equilibrium with $i$ and $j$ the termolecular collision frequency is


$$= \left( \sum_{ij,k} f_{ij,k}(T)K_{i,j} \right) [A][B][C] , \quad (4)$$

where $K_{i,j} = [ij]/[i][j]$.
More general models than Bodenstein’s can be devised by assuming that three clusters of species of any molecularity do not collide simultaneously. We do not see how to write a general expression for the frequency, thus we write an expression for each molecularity. We develop the model for tetra-species collisions, but the extension to higher molecularity collisions is straightforward.

Consider for example a tetra-species collision involving A, B, C, and D. Consider the collisions of two bimolecular clusters (pairs), say, AB and CD, and a termolecular cluster and a single species, say, ABC and D; the frequency of tetra-species collisions is

\[ Z_{ABCD} = \sum_{ij,kl} Z_{ij,kl} + \sum_{ijk,l} Z_{ijk,l} \tag{5} \]

where \( i, j, k, \) and \( l \) represent the species A, B, C, and D; and \( Z_{ij,kl} \) is the collision frequency for collisions between bimolecular clusters of \( ij \) and \( kl \) and \( Z_{ijk,l} \) is that for collisions between a termolecular cluster of \( ijk \) and a species \( l \).

Assuming Eq. (2) is valid to describe the frequency of bimolecular collisions involving two clusters, we have for \( Z_{ij,kl} \) and \( Z_{ijk,l} \)

\[ Z_{ij,kl} = f_{ij,kl}(T)[ij][kl], \tag{6} \]
\[ Z_{ijk,l} = f_{ijk,l}(T)[ijk][l], \tag{7} \]

where \( f_{ij,kl}(T) \) and \( f_{ijk,l}(T) \) are temperature dependent factors. Assuming that bimolecular clusters \( ij \) and \( kl \) are in equilibrium with their components we can write

\[ Z_{ij,kl} = f_{ij,kl}(T)K_{i,j}[i][j]K_{k,l}[k][l] = f_{ij,kl}(T)K_{i,j}K_{k,l}[i][j][k][l], \tag{8} \]

where \( K_{i,j} = [ij]/[i][j] \), and \( K_{k,l} = [kl]/[k][l] \). Similarly, assuming that termolecular cluster \( ijk \) is in equilibrium with its components the concentrations of the clusters can be related to the products of concentrations of \( i, j, \) and \( k \):
\[ [ijk] = K_{ij,k}[i][j] + K_{ik,j}[j] + K_{kj,i}[k] , \]  

(9)

where \(K_{ij,k}, K_{ik,j}, \text{ and } K_{kj,i}\) are the equilibrium constants. Collisions in which the three species collide simultaneously are neglected. Assuming that bimolecular clusters \(ij, ik, \text{ and } kj\) are in equilibrium with their components, Eq. (9) becomes

\[ [ijk] = K_{ij,k}K_{i,j} + K_{ik,j}K_{i,k} + K_{kj,i}K_{k,j} . \]  

(10)

Substitution of Eq. (10) in Eq. (7) for \([ijk]\) yields

\[
Z_{ijk,l} = f_{ijk,l}(T) \left( K_{ij,k}K_{i,j} + K_{ik,j}K_{i,k} + K_{kj,i}K_{k,j} \right) [i][j][l] \\
= f_{ijk,l}(T) \left( \sum_{xy,z} K_{xy,z}K_{x,y} \right) [i][j][l].
\]  

(11)

Substitution of Eqs. (8) and (11), respectively, for the first and second terms in Eq. (5) gives the expression for the frequency of tetra-species collisions involving A, B, C, and D as

\[
Z_{ABCD} = \left\{ \sum_{ij,kl} f_{ij,kl}(T)K_{i,j}K_{k,l} + \sum_{ijk,l} f_{ijk,l}(T) \left( \sum_{xy,z} K_{xy,z}K_{x,y} \right) \right\} [A][B][C][D].
\]  

(12)

The term inside the braces depends on temperature, thus Eq. (12) can be written in the simpler form

\[
Z_{ABCD} = f_{\text{tetra}}(T)[A][B][C][D].
\]  

(13)

The frequency of \(n\)-species collisions involving species A, B, \ldots, N can be expressed as the product of a \(n\)-molecular temperature dependent factor \(f_n(T)\) and the concentrations of A, B, \ldots, N; that is,

\[
Z_n = f_n(T)[A][B][C]...[N].
\]  

(14)
For example, the temperature dependent factor for \( n = 5 \), \( f_{\text{penta}} (T) \), is

\[
 f_{\text{penta}} = \sum_{ijklm} f_{ijklm}(T) \left( \sum_{xyz} K_{xyz} K_{x,y} \right) K_{l,m}
 + \sum_{ijklm} f_{ijklm}(T) \left( \sum_{xyz} K_{xyz} K_{x,y} \sum_{xyz} K_{xyz} K_{x,y} \right).
\]

The temperature dependent factors for \( n = 3, 4, \) and 5 are listed in Table 1.

The calculation of the temperature dependent factors, \( f_n (T) \), requires the equilibrium constants, \( K \)'s. We estimated the equilibrium constant \( K_{ij} \) in Eq. (4) as the ratio of the average collision duration for \( i \cdots j \) collisions, \( \tau_{ij} \), and mean free time \( t \) for species \( i \) between sequential \( i \cdots j \) collisions divided by concentration of \( j \); that is,

\[
 K_{i,j} = \frac{[ij]}{[i][j]} = \frac{\tau_{i,j}}{t[j]},
\]

where \( t \) is given by

\[
 t = \frac{1}{\sigma_{i,j} \bar{v}[j]},
\]

where \( \bar{v} = \sqrt{\frac{8kT}{\pi \mu_{bi}}} \) is the average relative velocity ( \( \mu_{bi} \) is the \( i,j \) reduced mass) and \( \sigma_{i,j} \) is the \( i,j \) collision cross section. Substituting Eq. (17) into Eq. (16) gives

\[
 K_{i,j} = \sigma_{i,j} \bar{v} \tau_{i,j},
\]

where \( \sigma_{i,j} \bar{v} \) is the temperature dependent factor,

\[
 f_{i,j}(T) = \sigma_{i,j} \bar{v}.
\]

Therefore, the equilibrium constant for the formation of bimolecular cluster \( ij \) is

\[
 K_{i,j} = f_{i,j}(T) \tau_{i,j}.
\]
Thus, Eq. (4) becomes

\[ Z_{ABC} = \left( \sum_{ij,k} f_{ij,k}(T) f_{i,j}(T) \tau_{i,j} \right) [A][B][C] = f_{\text{ter}}(T)[A][B][C]. \]  

(21)

Assuming that Eq. (19) describes the temperature dependent factors for collisions involving a bimolecular cluster and a single species, say, \( ij \) and \( l \), yields

\[ Z_{ABC} = \left( \frac{8kT}{\pi \mu_{\text{ter}}} \right) \left( \sum_{ij,k} \sigma_{ij,k} \tau_{ij,k} \right) [A][B][C], \]  

(22)

where \( \mu_{\text{ter}} \) is the termolecular reduced mass.

Dividing Eq. (22) by the bimolecular collision frequency \( Z_{AB} = \frac{8kT}{\pi \mu_{\text{bi}}} \sigma_{AB}[A][B] \), gives the ratio of the frequencies of termolecular to bimolecular collisions (\( ZT/ZB \)) as

\[ \frac{ZT}{ZB} = \frac{\sum_{ij,k} \sigma_{ij,k} \tau_{ij,k}}{\pi \mu_{\text{ter}}^2 \sigma_{AB}} [C], \]  

(23)

where \( i, j, \) and \( k \) represent species \( A, B, \) and \( C \). The average collision durations \( \tau_{i,j} \) can be computed directly from MD simulations.

### 3. COMPUTATIONAL METHODS

Molecular dynamics (MD) simulations of the NVT ensemble were performed using the 2016 version of LAMMPS\textsuperscript{11,12} with the ReaxFF\textsuperscript{7-10} reactive force field. Each simulation was for 100 molecules with the number of \( \text{H}_2 \) and \( \text{O}_2 \) varied to give composition ratio \([\text{H}_2]/[\text{O}_2] \) values over the range 0.25 to 2.00 in a cubic cell with periodic boundary conditions at 3,000 K, 4,000 K, and 5,000 K. The simulation box size was varied based on the composition ratio to give specific
densities in the range 120.2 kg m$^{-3}$ to 332.7 kg m$^{-3}$. A total of 3,600 trajectories were run for 60 sets of initial conditions (60 per set).

Initial configurations for each simulation were generated using the PACKMOL$^{13}$ code. The molecules were placed in the simulation cell such that all non-bonded atom-atom distances were at least 2.2 Å. The equilibrium bond lengths of 0.759 Å and 1.188 Å, respectively, for H$_2$ and O$_2$ reactants were used to generate the initial configurations, which were adjusted by using the Polak-Ribiere$^{14}$ version of the conjugate gradient algorithm for energy minimization until either the energy, $1 \times 10^{-4}$, or the force criterion, $1 \times 10^{-6}$ kcal mol$^{-1}$ Å$^{-1}$, was satisfied. Atomic velocities were selected from Gaussian distributions corresponding to 3,000 K, 4,000 K, or 5,000 K. The temperature was controlled by a Nosé-Hoover thermostat$^{15,16}$ with damping constant 500 fs. The trajectories were propagated using the velocity Verlet integrator$^{17,18}$ with time step 0.1 fs.

The system was thermally equilibrated, which was ascertained by calculating the average kinetic rotational, vibrational, and translational temperatures, and the system was considered thermally equilibrated when the average kinetic temperatures continuously fluctuated about the target temperature. Depending on the density and temperature complete equilibration required 7-25 ps. The bond distances were monitored during the equilibration simulations, and if a reaction occurred the simulation was abandoned. The configurations of the system at the end of the equilibration simulations were used as the initial states for production runs, which were for 6 ns or until an initial reaction was detected; at 4,000 K and 5,000 K the initial reaction usually occurred within 250 ps. If an initial reaction was detected the simulation was abandoned to preclude the effect of the thermostat on the chemistry.

The molecular bond lengths were checked every 0.5 fs, and when a O$_2$ or H$_2$ bond distance exceeded 2.3 Å and 1.5 Å, respectively, it was assumed that a reaction was underway. These
interatomic cutoff distances were set to give energies within 1 kcal mol\(^{-1}\) of the asymptotic limit of the bond dissociation energy curves predicted by ReaxFF.\(^8\) A reaction cluster was defined for a reaction using critical cutoff values of 3.0 Å, 1.8 Å, and 3.5 Å for the non-bonded atom-atom distances for H···H, H···O, and O···O, respectively. A new bond was considered formed when the atom-atom distance underwent at least four inner turning-points. The internuclear distances were monitored for the reaction cluster beginning 210 fs before the time at which the bond rupture was detected.

The average frequencies of O\(_2\)···H and O\(_2\)···H\(_2\)···O\(_2\) collisions were computed from the MD simulation results for the composition ratio [H\(_2\)]/[O\(_2\)] = 0.25 and density range 120 kg m\(^{-3}\) to 350 kg m\(^{-3}\) at 3,000 K and 4,000 K. For each set of initial conditions ten H\(_2\) molecules were selected from the seven trajectories run for that set of initial conditions. We monitored all the H···O atom-atom distances for each selected molecule over the first 50 ps of the simulation. When an O···H atom-atom distance between a selected H\(_2\) molecule and an O\(_2\) molecule became less than 1.8 Å, it was considered to be a bimolecular H\(_2\)···O\(_2\) collision. When a selected H\(_2\) molecule was found between two O\(_2\) molecules with two O···H distances within 1.8 Å it was considered to be a termolecular O\(_2\)···H\(_2\)···O\(_2\) collision.

We computed the average collision duration \(\tau\) for O\(_2\)···O\(_2\) and O\(_2\)···H\(_2\) pairs for the density range 120-350 kg m\(^{-3}\) and composition ratio [H\(_2\)]/[O\(_2\)] = 0.25 at 3,000 K and 4,000 K to be used in Eq. (23). We began by identifying collision pairs. Ten O\(_2\) molecules were arbitrarily selected from each of the seven trajectories and were tracked for 50 ps. When an O\(_2\) molecule was found to be within an O···O distance of 3.5 Å of another O\(_2\) molecule or within 1.8 Å of a H\(_2\) molecule they were taken to be, respectively, an O\(_2\)···O\(_2\) or O\(_2\)···H\(_2\) collision pair. We excluded collisions in
which the pair collided with a third molecule. The collision duration for a pair was measured as the time during which the molecules of the pair were within the atom-atom cutoff distances.

4. RESULTS AND DISCUSSION

The ReaxFF-based simulations predicted two primary reactions that produce initial radicals: the bimolecular reaction $H_2 + O_2 \rightarrow H + HO_2$, which was usually interrupted by a third molecule ($O_2$ or $H_2$), and the termolecular reaction $2O_2 + H_2 \rightarrow 2HO_2$. These reactions begin with the breaking of a $H_2$ bond. They accounted for 95% and 80% of the initial reactions at 3,000 K and 5,000 K, respectively.

Plots of the probability of initial reactions $H_2 + O_2 \rightarrow H + HO_2$ (circles) and $2O_2 + H_2 \rightarrow 2HO_2$ (squares) for (a) $[H_2]/[O_2] = 0.25$, (b) $[H_2]/[O_2] = 0.5$, (c) $[H_2]/[O_2] = 1$ and (d) $[H_2]/[O_2] = 2$ at 3,000 K (blue), 4,000 K (green), and 5,000 K (red) over the density range 120-350 kg m$^{-3}$ are shown in Figure 1; the straight lines are drawn to guide the eye. The probability of $H_2 + O_2 \rightarrow H + HO_2$ as the initial reaction is larger than that of $2O_2 + H_2 \rightarrow 2HO_2$ over the density range 120-350 kg m$^{-3}$ at 4,000 K and 5,000 K, with the exception that at 4,000 K and 350 kg m$^{-3}$ the green circle and the green square essentially overlap (see Frame b); in fact, the probabilities for most of the other cases cross at high density. The $H_2 + O_2 \rightarrow H + HO_2$ reaction is the most prevalent over the density range 120 kg m$^{-3}$ to 350 kg m$^{-3}$ at 4,000 K and 5,000 K; however, at 3,000 K and $\rho > 250.0$ kg m$^{-3}$ $2O_2 + H_2 \rightarrow 2HO_2$ becomes dominant (see blue points in Figure (1)). At 4,000 K and 5,000 K the probability of $2O_2 + H_2 \rightarrow 2HO_2$ increases with density, although not to the extent that it does at 3,000 K. The probability of the bimolecular reaction decreases with increasing of density.
The frequencies and durations of the O$_2$⋯H$_2$⋯O$_2$ and O$_2$⋯H$_2$ collisions are needed to determine their relative importance, which we have computed using MD with ReaxFF. We cannot directly validate the force field for O$_2$ + H$_2$ collisions, however, we can check the predictions of the equations of state for H$_2$ and O$_2$ since there are results in the literature for comparisons.

We computed the pressure-density isotherms ($P$-$\rho$,$T$) for H$_2$ and O$_2$ by simulating NVT ensembles with 2,500 molecules in cubic simulation boxes of volumes 55×55×55 Å$^3$, 75×75×75 Å$^3$, 85×85×85 Å$^3$, 95×95×95 Å$^3$ and 105×105×105 Å$^3$ at 300 K, 500 K and 1,000 K. Simulations were done for O$_2$ at three more temperatures, 2,000 K, 3,000 K, and 4,000 K, to cover the range of the results in the literature.

The ReaxFF prediction for the EOS of H$_2$ can be compared to the predictions of the truncated virial equation reported by Lemmon et al.,$^{19}$ which is in a good agreement with the data in the NIST Standard Database.$^{20}$ The comparison is shown in Figure 2(a). The relative errors of the computed results shows that ReaxFF underestimates the $P$-$\rho$ isotherms for H$_2$ over the density range 8 kg m$^{-3}$ to 50 kg m$^{-3}$ and the temperature range 300 K to 1,000 K (see Figure 2(b)); however, ReaxFF acceptably reproduces the thermodynamic data.$^{19,20}$ The best agreement is at lower densities and higher temperatures; for example, at 1,000 K the relative errors are in the range 0.7% to 8.5% for the density range of 8 kg m$^{-3}$ to 50 kg m$^{-3}$.

The ReaxFF predictions for the $P$-$\rho$ isotherms of O$_2$ can be compared with the empirical equation derived by Kuznetsov et al.$^{21}$ based on data reported by Sychev et al.$^{22}$ ReaxFF overestimates the $P$-$\rho$ isotherms for O$_2$ over the density range of 130 kg m$^{-3}$ to 330 kg m$^{-3}$ and temperature range of 300 K to 4,000 K. The comparison is shown in Figure 3(a). The best agreement is at lower densities and higher temperatures (see Figure 3(b)). Based on these results
we expect that at densities below 350 kg m⁻³ and temperatures higher than 3,000 K that ReaxFF will provide acceptable results for the frequencies of O₂⋯O₂, H₂⋯H₂, and O₂⋯H₂ collisions in H₂:O₂ mixtures. The use of a different thermostat could improve the agreement between the ReaxFF predictions for the P-ρ isotherms and the empirical EOS; however, we have not explored other thermostats.

We computed average frequencies of O₂⋯H₂ and O₂⋯H₂⋯O₂ collisions for each set of initial conditions for 3,000 K and 4,000 K, density range 120 kg m⁻³ to 350 kg m⁻³, and [H₂]/[O₂] = 0.25. Seventy H₂ molecules selected from seven NVT simulations were traced during the first 50 ps of the simulations. The O₂⋯H₂ and O₂⋯H₂⋯O₂ collisions were identified using the atom-atom distance criterion of \( r_{O_i H_j} \leq 1.8 \) Å. Figure 4 shows the ReaxFF predictions the average collision frequencies for O₂⋯H₂ (red curves and circles) and O₂⋯H₂⋯O₂ (purple curves and squares) collisions as functions of density. The results for 3,000 K are shown by the dashed curves and blue points and those for 4,000 K by the solid curves and green points. At both temperatures the termolecular collisions are less frequent than the bimolecular collisions by a factor of \( \sim 30 \); however, the latter increases linearly while the termolecular collision frequency increases nonlinearly at a faster rate.

The plots in Figure 5 show the collision frequency ratios for 3,000 K (blue points) and 4,000 K (green points). The ratios were calculated by using the results in Figure 4. ReaxFF predicts that termolecular collisions are favored over bimolecular collisions at the higher densities. The upward trend of \( ZT/ZB \) values is in accord with Eq. (23). The explicit dependence of \( ZT/ZB \) on density at a given temperature requires the average collision duration \( \tau_{i,j} \); see Eq. (23).
We computed average collision durations $\tau_{O_2\cdots O_2}$ and $\tau_{H_2\cdots O_2}$ from the MD simulation results for 3,000 K and 4,000 K, density range 120 kg m$^{-3}$ to 350 kg m$^{-3}$, and $[H_2]/[O_2] = 0.25$. For each set of initial conditions 70 O$_2$ molecules were selected from seven different simulations. The ReaxFF predictions of the density dependences of $\tau_{O_2\cdots O_2}$ and $\tau_{H_2\cdots O_2}$ are shown in Figure 6 (left y-axis and right y-axis, respectively) for 3,000 K (dashed lines and blue points) and 4,000 K (solid lines and green points). ReaxFF predicts that $\tau_{H_2\cdots O_2}$ and $\tau_{O_2\cdots O_2}$ decrease linearly with density. The latter drops by 15 fs over the density range 120 kg m$^{-3}$ to 350 kg m$^{-3}$ while the former decreases only 0.3 fs.

The dependence of the average collision durations on density can be expressed as

$$\tau = -a\rho + b,$$

where $a$ and $b$ depend on temperature, which are given for $\tau_{H_2\cdots O_2}$ and $\tau_{O_2\cdots O_2}$ in Table 2 for 3,000 K and 4,000 K. The linear dependence of $\tau$ on $\rho$ is a result of the high density of the mixture. In a low-density mixture bimolecular clusters remain isolated for relatively long times, while at high density they can be interrupted by another species or cluster before they decompose physically or chemically.

Inserting Eq. (24) for the average collision durations into Eq. (23) and using density instead of concentration gives

$$\frac{ZT}{ZB} = \sqrt{T} \left( -\alpha \rho^2 + \beta \rho \right),$$

where $\alpha$ and $\beta$ are temperature dependent parameters due to the dependence of the collision duration on temperature. Equation (25) is characteristic of van der Waals (vdW) liquids. The second and third virial coefficients depend, respectively, on the pair interactions and non-additive 3-body interactions. Therefore, the ratio of the frequencies of termolecular to
bimolecular collisions is a measure of the ratio of the third to second virial coefficient, which for a vdW liquid has a temperature dependent constant with algebraic dependence on density. The curves in Figure 5 are fits of the points to Eq. (25). They show the ratios of frequencies of termolecular to bimolecular collisions at 3,000 K and 4,000 K as functions of density; the parameter values are given in Table 3.

Despite the fact that \( \text{O}_2 \cdot \cdot \cdot \text{H}_2 \cdot \cdot \cdot \text{O}_2 \) collisions are less frequent than bimolecular collisions by a factor of ~30 (see Figure 6), increasing the density over the range 120 kg m\(^{-3}\) to 350 kg m\(^{-3}\) favors termolecular collisions \( \text{O}_2 \cdot \cdot \cdot \text{H}_2 \cdot \cdot \cdot \text{O}_2 \) over bimolecular collisions \( \text{O}_2 \cdot \cdot \cdot \text{H}_2 \). This accounts for the upward trend in the probability of \( 2\text{O}_2 + \text{H}_2 \rightarrow 2\text{HO}_2 \) (see Figure 1). According to Eq. (25) the ratio of collision frequencies of termolecular to bimolecular reaches a maximum with increasing density and falls to negative values at large density. We did not carry out calculations for density values larger than 350 kg m\(^{-3}\) to illustrate this fall-off.

The ReaxFF predictions of the EOSs are good at high temperatures, thus the validity of the prediction of the density dependence of \( Z_T/Z_B \) depends on the accuracy of the ReaxFF predictions of the reaction barriers. The ReaxFF energy barrier for the bimolecular reaction seems to be acceptable, thus the issue is the termolecular reaction.

According to Eq. (25) \( Z_T/Z_B \) depends on temperature directly by the \( \sqrt{T} \) factor and indirectly through the coefficients \( \alpha \) and \( \beta \). To eliminate the direct dependence on temperature we define the \textit{characteristic collision frequency ratio} \( (Z_T/Z_B)^* \) by dividing both sides of Eq. (25) by \( \sqrt{T} \); that is,

\[
(Z_T/Z_B)^* = (\alpha \rho^2 + \beta \rho),
\]

which is plotted in Figure 7 for 3,000 K and 4,000 K using the results of Figure 5. (We use the term “characteristic collision frequency” to stress that Eq. (26) does not have any explicit
dependence on temperature.) The curves in Figure 7, which were obtained by non-linear least-squares fitting of the points Eq. (26), are essentially indistinguishable. ReaxFF predicts that the dependences of coefficients $\alpha$ and $\beta$ on temperature are negligible for $O_2\cdots H_2\cdots O_2$ and $O_2\cdots H_2$ collisions. This could be because the coefficients depend on the temperature via the dependences of collision durations, $\tau_{H_2-O_2}$ and $\tau_{O_2-O_2}$, which are oppositely affected by temperature according to the results in Figure 6. The ratio of $O_2\cdots H_2\cdots O_2$ to $O_2\cdots H_2$ collisions depends explicitly on temperature only via $\sqrt{T}$. Increasing the temperature from 3,000 K to 5,000 K leads to an increase in the number of $O_2\cdots H_2\cdots O_2$ collisions relative to $O_2\cdots H_2$ collisions by 28% according to Eq. (25); however, the probability of $2O_2 + H_2 \rightarrow 2HO_2$ being the initial reaction at 5,000 K is less than that at 3,000 K; see Figure 1. We explain this difference by comparing the classical energy requirements for the two reactions.

The classical energy requirement for a reaction requires that the sum of the internal and relative translational energies of the colliding reactants exceed the barrier height, $E_b$. For a diatomic molecule in the classical limit the average internal energy is $2RT$ according to the equipartition theorem. Therefore, when all the reactants are diatomic molecules the energy requirement is satisfied only if the relative translational energy of the colliders exceeds the threshold value of $(E_{rel})^*$ given by

$$ (E_{rel})^* = E_b - 2NRT, \quad (27) $$

where $N$ is two and three for bi- and ter-molecular collisions, respectively.

The height of the energy barrier to the termolecular reaction is not known with certainty. We can, for the sake of illustration, use the values that are available with the understanding that realistic predictions await an accurate description of the barriers. The energy barriers predicted by ReaxFF are 16.7 kcal mol$^{-1}$ and 59.4 kcal mol$^{-1}$ for the ter- and bi-molecular reactions,
respectively. Also available are values reported by Monge-Palacios and Rafatijo\textsuperscript{23} at CCSD(T) = FULL/aug-cc-pVTZ//CCSD = FC/cc-pVTZ single point \textit{ab initio} level of theory: 49.8 kcal mol\textsuperscript{-1} and 58.9 kcal mol\textsuperscript{-1} for the ter- and bi-molecular reactions, respectively. To our knowledge there has not been any other result reported for the energy barrier to the termolecular reaction.

The probability of a collision involving \(N\) species with the relative translational energy higher than \((E_{\text{rel}})^*\) was computed using

\[
\text{Prob} = \left(2\pi \mu_N kT \right)^{-\frac{3(N-1)}{2}} \int \cdots \int_{(E_{\text{rel}})^*}^{\infty} \mathbf{P}(E_{\text{rel}}) \exp \left(\frac{-E_{\text{rel}}}{kT}\right) d^{3(N-1)} \mathbf{P},
\]

where \(\mathbf{P}\) is the relative momentum vector and \(\mu_N\) is the reduced mass of the system of \(N\) species.

The ReaxFF results are given in Table 4 for \(2\text{O}_2 + \text{H}_2 \rightarrow 2\text{HO}_2\) and \(\text{H}_2 + \text{O}_2 \rightarrow \text{H} + \text{HO}_2\) at 3,000 K, 4,000 K, and 5,000 K. All of the termolecular collisions and less than 1\% of the bimolecular collisions meet the classical energy requirement at 3,000 K. Increasing the temperature has a greater positive effect on the probability of bimolecular collisions than on that for termolecular collisions, thus the bimolecular reaction \(\text{H}_2 + \text{O}_2 \rightarrow \text{H} + \text{HO}_2\) dominates in the production of initial radicals at elevated temperatures as shown by the results in Figure 1. This is in accord with the conclusion by Karkach and Osherov\textsuperscript{24} that \(2\text{O}_2 + \text{H}_2 \rightarrow 2\text{HO}_2\) is the main source of initial radicals at elevated pressures and low temperatures in the initiation mechanism of hydrogen combustion. However, they did not report information about the energetics or the initial conditions that would make this reaction the dominant initial reaction.
5. CONCLUSIONS

We have presented a method based on Bodenstein’s definition\(^3\) of termolecular collisions and the cluster representation of gases that allows determination of the molecularity of collisions in MD simulations of bulk gases. The generalized Bodenstein’s model assumes that higher molecularity collisions is the result of collisions between two clusters of species. Using Bodenstein’s definition of termolecular collisions, an analytical expression is derived for termolecular collision frequencies based on molecular kinetic theory. While we focused on termolecular collisions, the method can be extended to obtain the frequencies of collisions of various molecularities. We have shown the generalization to tetra- and penta-species collisions.

We have defined the ratio of the frequencies of termolecular to bimolecular collisions (ZT/ZB) to determine the relative importance of termolecular collisions in the macroscopic behavior of a gas-phase system. The ZT/ZB depends on temperature and pressure (density), and explicitly on the collision times. We illustrated applicability of the method by computing collision frequencies and collision times for O\(_2\)···H\(_2\)···O\(_2\) and O\(_2\)···H\(_2\) collisions from MD simulations of [H\(_2\)]/[O\(_2\)] = 0.25 mixtures over the density range of 120.2 kg m\(^{-3}\) to 332.7 kg m\(^{-3}\) for 3,000 K and 4,000 K with the ReaxFF reactive force field. The MD results for ZT/ZB qualitatively agree with those predicted by the theory.
AUTHOR INFORMATION

RafatijoH@missouri.edu; mongepalaciosm@gmail.com; thompsondon@missouri.edu

ACKNOWLEDGMENTS

We thank Ward H. Thompson for some useful comments. This work was supported by the U. S. Army Research Laboratory and the U. S. Army Research Office under grant number W911NF-14-1-0359.
References


(23) Monge-Palacios, M.; Rafatijo, H. On the role of the termolecular reactions $2O_2 + H_2 \rightarrow 2HO_2$ and $2O_2 + H_2 \rightarrow H + HO_2 + O_2$ in formation of the first radicals in hydrogen combustion: *ab initio* predictions of energy barriers. *Phys. Chem. Chem. Phys.* **2017**, *19*, 2175-2185. They did not take into account spin contamination in the termolecular reaction, thus the reported value is likely not accurate.

Table 1: The temperature dependent factors $f_n(T)$ for $n = 3, 4,$ and $5.$

<table>
<thead>
<tr>
<th>$n = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sum_{ij,k} f_{ij,k}(T)K_{i,j}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$n = 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sum_{ij,kl} f_{ij,kl}(T)K_{i,j}K_{k,l} + \sum_{ijk,l} f_{ijk,l}(T) \left( \sum_{xy,z} K_{xy,z}K_{x,y} \right)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$n = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sum_{ijkl,m} f_{ijkl,m}(T) \left( \sum_{xy,z} K_{xy,z}K_{x,y} \right)K_{i,m}$</td>
</tr>
<tr>
<td>$+ \sum_{ijkl,m} f_{ijkl,m}(T) \left( \sum_{xy,z} K_{xy,z}K_{x,y}K_{z,t} + \sum_{xyz,t} K_{xyz,t} \sum_{xy,z} K_{xy,z}K_{x,y} \right)$</td>
</tr>
</tbody>
</table>

Table 2: Values of the parameters $a$ and $b$ in Eq. (24) for $\tau_{H_2-O_2}$ and $\tau_{O_2-O_2}$ at 3,000 K and 4,000 K.

<table>
<thead>
<tr>
<th></th>
<th>$a$ (fs m$^3$/kg)</th>
<th>$b$ (fs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{H_2-O_2}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,000 K</td>
<td>-0.0016</td>
<td>6.33</td>
</tr>
<tr>
<td>4,000 K</td>
<td>-0.0017</td>
<td>6.45</td>
</tr>
<tr>
<td>$\tau_{O_2-O_2}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,000 K</td>
<td>-0.0829</td>
<td>69.46</td>
</tr>
<tr>
<td>4,000 K</td>
<td>-0.0899</td>
<td>68.03</td>
</tr>
</tbody>
</table>
Table 3: Values of parameters $\alpha$ and $\beta$ in Eq. (25) for the ratio of the frequencies of termolecular ($O_2\cdot\cdot\cdotH_2\cdot\cdot\cdotO_2$) to bimolecular ($O_2\cdot\cdot\cdotH_2$) at 3,000 K and 4,000 K.$^a$

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$ (m$^6$/K$^{1/2}$ kg$^3$)</th>
<th>$\beta$ (m$^3$/K$^{1/2}$ kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ZT/ZB</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,000 K</td>
<td>$-8.213\times10^{-12}$</td>
<td>$2.051\times10^{-6}$</td>
</tr>
<tr>
<td>4,000 K</td>
<td>$-3.059\times10^{-12}$</td>
<td>$2.019\times10^{-6}$</td>
</tr>
</tbody>
</table>

$^a$ Obtained by nonlinear least-squares fitting of Eq. (25) using the data in Figure 5.

Table 4: The probabilities of meeting the energy requirements for the ter- and bi-molecular reactions at 3,000 K, 4,000 K, and 5,000 K.$^a$

<table>
<thead>
<tr>
<th></th>
<th><strong>3,000 K</strong></th>
<th><strong>4,000 K</strong></th>
<th><strong>5,000 K</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Termolecular</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy available</td>
<td>35.76</td>
<td>47.68</td>
<td>59.61</td>
</tr>
<tr>
<td>$(E_{rel})^b$</td>
<td>0.0 (14.04)</td>
<td>0.0 (2.12)</td>
<td>0.0 (0.0)</td>
</tr>
<tr>
<td>Probability$^b$</td>
<td>1 (0.581)</td>
<td>1 (0.997)</td>
<td>1 (1)</td>
</tr>
<tr>
<td><strong>Bimolecular</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy available</td>
<td>23.84</td>
<td>31.79</td>
<td>39.74</td>
</tr>
<tr>
<td>$(E_{rel})^b$</td>
<td>35.56 (35.06)</td>
<td>27.61 (27.11)</td>
<td>19.66 (19.16)</td>
</tr>
<tr>
<td>Probability$^b$</td>
<td>0.008 (0.008)</td>
<td>0.074 (0.078)</td>
<td>0.266 (0.277)</td>
</tr>
</tbody>
</table>

$^a$ The average internal energy was calculated in accordance with the equipartition theorem. All energy values are in kcal/mol.

$^b$ The values in ( ) were computed with the barrier heights reported by Monge-Palacios and Rafatijo.$^{23}$ The values that are not in ( ) were computed with ReaxFF barrier heights.
In the cluster representation a gas is considered to be a mixture of clusters of different molecularities at any instant of time.