

A Neural Network Parametrized Coagulation Rate Model for < 3 nm Titanium Dioxide Nanoclusters

Tomoya Tamadate, Suo Yang, Christopher J. Hogan Jr.*

Department of Mechanical Engineering, University of Minnesota, Minneapolis, MN USA

Submitted to:

The Journal of Chemical Physics

*To whom correspondence should be addressed: hogan108@umn.edu

ABSTRACT

Coagulation is a key factor governing the size distribution of nanoclusters during the high temperature synthesis of metal oxide nanomaterials. Population balance models are strongly influenced by the coagulation rate coefficient utilized. Although simplified coagulation models are often invoked, the coagulation process, particularly for nanometer-scale particles, is complex, affected by the coagulating nanocluster sizes, the surrounding temperature, and potential interactions. Towards developing improved models of nanocluster and nanoparticle growth, we have developed a neural network (NN) model to describe titanium dioxide (TiO_2) nanocluster coagulation rate coefficients, trained with molecular dynamics (MD) trajectory calculations. Specifically, we first calculated TiO_2 nanocluster coagulation probabilities via MD-trajectory calculations varying the nanocluster diameters from 0.6 nm to 3.0 nm, initial relative velocity from 20 to 700 m s^{-1} , and impact parameter from 0.0 to 8.0 nm. Calculations consider dipole-dipole interactions, dispersion interactions, and short-range repulsive interactions. We trained a NN model to predict whether a given set of nanocluster diameters, impact parameter, and initial velocity would lead to the outcome of coagulation. The accuracy between the predicted outcomes from the NN model and the MD trajectory calculation results are $> 95\%$. We subsequently utilized both the NN model and MD trajectory calculations to examine coagulation rate coefficients at 300 K and 1000 K. The NN model predictions are largely within the range 0.65-1.54 of MD predictions, and importantly NN predictions capture the local minimum coagulation rate coefficients observed in MD-trajectory calculations. The NN model can be directly implemented in population balances of TiO_2 formation.

I. INTRODUCTION

The synthesis of oxide nanomaterials in flames typically involves the volatilization of an organometallic precursor, oxidation of the precursor vapor, and subsequently the formation of metal oxide nanoclusters, which are solid species with nanometer scale dimensions¹⁻⁴. Upon the formation of nanoclusters⁵⁻⁷, further growth into nanoparticles proceeds via surface growth by nanocluster collisions with partially oxidized precursor molecules,^{8, 9} as well as by nanocluster-nanocluster collisions (coagulation)^{10, 11}. While the relative importance of these two processes is dependent upon the temperature-time history in the reactor and the properties of the organometallic precursor vapor^{12, 13}, nanocluster-nanocluster coagulative growth is likely a key contributor to nanoparticle growth, as the precursor rapidly reacts and becomes depleted in concentration in flame reactors.

For this reason, prediction of metal oxide nanoparticle growth rates in flames hinges upon accurate calculation of nanocluster-nanocluster coagulation rate coefficients; these rate coefficients are essential inputs to population balance models for nanocluster and nanoparticle size distribution evolution predictions¹⁴. Coagulation rates are dependent not only on the nanocluster sizes and system temperature, but also the potential interactions between nanoclusters closely approaching one another^{10, 15-18}. Unfortunately, the conventional approach to coagulation rate prediction is rather coarse. It is commonplace in modeling growth dynamics in combustion systems to utilize the free molecular hard-sphere collision model for nanocluster-nanocluster coagulation, in which the coagulation rate coefficients are the products of the mean thermal speed of the reduced nanocluster masses and the nanocluster-nanocluster combined projected areas¹⁹. Attractive potential interactions are not explicitly considered in the hard-sphere collision model. In instances where the effects of attractive potential interactions are considered, it is commonplace

to invoke a temperature-independent, constant “enhancement factor” for the coagulation rate coefficient, i.e. a constant factor multiplied by the hard-sphere-predicted coagulation rate coefficient^{20, 21}. However, there is ample evidence, both computational and experimental, that enhancement factors for metal oxide clusters are large (>2 for many materials), strongly temperature-dependent, and also dependent on the coagulating nanocluster sizes. For example, Goudeli et al²², theoretically examining SiO₂ nanocluster coagulation through molecular dynamics trajectory calculations, found enhancement factors in the 3–9 for nanometer scale clusters, in the 300-1500 K temperature range, with enhancement factors decreasing with increasing temperature. Yan et al²³ and Zhang et al¹⁷ utilized molecular dynamics simulations to argue that for TiO₂ nanoclusters ~ 2 nm in diameter, the coagulation rate coefficient at room temperature will be more than 8 times higher than the hard-sphere predicted model, and this is largely due to dipole-dipole interactions between closely approached nanoclusters. Experimentally, Sharma et al¹¹ estimated enhancement factors in the 3-7 range for 5-6 nm TiO₂ particles, also decreasing with increasing temperature in the 673-1073 K range.

While prior research consistently yields large, temperature-dependent enhancement factors for metal oxide nanocluster coagulation rates, studies to-date utilizing atomistic simulations to infer nanocluster-nanocluster potential energy surfaces and coagulation rates have been limited to specific circumstances^{17, 22-25}. To incorporate improved coagulation rate models into nanoparticle growth predictions, nanocluster-nanocluster collision models need to be used to provide coagulation rate coefficients not only for explicitly simulated circumstances, but also over wider nanocluster size and system temperature ranges. Towards this end, using TiO₂ nanoclusters in the 0.6-3.0 nm diameter range as a model system, we utilize molecular dynamics (MD) trajectory calculations, similar to prior studies, to precisely examine nanocluster-nanocluster coagulation

rates. Subsequently, we utilize a neural network model to fit results for the coagulation probability of nanoclusters for a given set of nanocluster diameters, impact parameter, and initial relative velocity. While the neural network is ultimately, a “black-box” approach, because neural network training and testing is readily facilitated by modern computational tools, the combination of trajectory calculations with detailed potential interactions between nanoclusters (atomistic simulations) and neural network training at present appears to be a more tractable approach coagulation rate coefficient estimation than does the use of simpler potential function approximations between nanoclusters and analytical modeling. We show that the fit coagulation probability map can be used to predict coagulation rate coefficients for TiO_2 nanoclusters under widely variable temperatures and nanocluster diameters, i.e. the resulting modeling can be directly used in population balance models for nanoparticles.

II. COMPUTATIONAL METHODS

We utilize MD-trajectory calculations^{22, 24-26} to examine TiO_2 nanocluster collisions, defining the coagulation rate coefficient and enhancement factor in rate with respect to hard-sphere collisions, and then subsequently developing a neural network (NN) model to fit trajectory outcomes, and to develop a function for the coagulation rate coefficient. In section A, we define the coagulation rate coefficient and enhancement factor. In section B, the methods employed to determine probability maps for the outcomes of trajectories are briefly described, with reference to prior work adopting similar methods. In section C, we provide a description of the neural network (NN) fit to coagulation probability maps, including the structure of the NN, used equations in NN, and training process using MD simulation results obtained in B.

A. Theories for coagulation rate coefficient and enhancement factor calculations

The coagulation rate coefficient for two nanoclusters, β_{ij} , where the subscripts “ i ” and “ j ” denote the nanocluster “type” (typically diameter or mass), can be calculated as²²:

$$\beta_{ij} = \int_0^\infty \int_0^\infty L_{ij} db dv_0 \quad (1a)$$

$$L_{ij} = 2\pi b \cdot P(b, v_0) \cdot v_0 f_{\text{MB}}(v_0) \quad (1b)$$

$$f_{\text{MB}}(v_0) = \left(\frac{m_{ij}}{2\pi k_b T} \right)^{1.5} 4\pi v_0^2 \exp\left(-\frac{m_{ij} v_0^2}{2k_b T}\right) \quad (1c)$$

L_{ij} has been termed binding length²²; it is a function of $f_{\text{MB}}(v_0)$, the Maxwell-Boltzmann distribution, v_0 , the relative speed approaching nanoclusters, b , the impact parameter, m_{ij} , the reduced mass, k_b , Boltzmann’s constant, T , the background system temperature, and $P(b, v_0)$, the probability that coagulation occurs, for a given b and v_0 . With the hard sphere assumption (subscript “HS”) in the free molecular regime:

$$P_{\text{HS}}(b, v_0) = \begin{cases} 1 & (b \leq a_{ij}) \\ 0 & (b > a_{ij}) \end{cases} \quad (2)$$

where a_{ij} is the collision distance of the nanoclusters, defined as $a_{ij} = \frac{D_{p,i} + D_{p,j}}{2}$, when $D_{p,i}$ and $D_{p,j}$ are the nanocluster i and j diameters. Substituting equation 2 into equation 1 and integrating yields the hard sphere coagulation rate coefficient²⁷:

$$\beta_{\text{HS}} = \pi a_{ij}^2 \sqrt{\frac{8k_b T}{\pi m_{ij}}} \quad (3)$$

Although the coagulation rate coefficient is predictable from this equation, as shown in previous studies, the calculated values are often underestimations (especially on the nanocluster). The ratio of the true coagulation rate coefficient (or at least improved approximations to it) to the hard sphere coefficient is defined as the enhancement factor, η_{ij} ^{15, 16}:

$$\eta_{ij} = \frac{\beta_{ij}}{\beta_{\text{HS}}} \quad (4)$$

MD-trajectory calculations are performed with the intention of evaluating $P(b, v_0)$ considering potential interactions between nanoclusters.

B. Probability mapping, $P(b, v_0)$ prediction via MD simulation

While simple potential models have been found accurate for atomic scale systems with stronger Coulombic interactions (i.e. those dominated by ion-dipole interactions²⁶ or ion-ion interactions²⁸), simple models have been less successful for electrically neutral nanoclusters wherein dipole-dipole interactions may be significant. We simulated the coagulation of two TiO_2 nanoclusters in the size range of 0.6 to 3.0 nm on the open-source MD software LAMMPS²⁹ customized with a coagulation (collision) event detection sub-routine. A schematic diagram depicting the coagulation MD simulation is shown in Figure 1. Prior to these simulations, we prepared the different diameter TiO_2 test nanoclusters. Approximately spherical nanoclusters were cut from a TiO_2 bulk crystal (10 nm $\times$ 10 nm $\times$ 10 nm, with the anatase structure) with target sizes as the diameters while ensuring one titanium atom to two oxygen atoms. While simulations of aerosol processes need to be performed unthermostated when the particles or nanoclusters are smaller than the mean free path of the gas (i.e. thermal equilibration does not occur on the timescale of the process simulated^{24, 28, 30, 31}), initialization of the simulated species does require an equilibration step. The nanoclusters were thermally relaxed individually at 300 K in canonical ensemble (constant number, volume, and temperature: NVT) MD simulations with a Nose-Hoover thermostat³² for 2 ns. This thermal relaxation process was performed at only 300 K; strictly, the internal (vibrational and rotational) energies of atoms with nanoclusters will vary with system

temperature. However, prior work²⁴, and preliminary simulations show that for nanocluster coagulation the nanocluster internal temperature does not significantly influence on the coagulation rate, although the processes after collision and binding is certainly affected (i.e. the rate of coalescence^{33, 34}, not examined in the present study). Two sizes of thermally-relaxed nanoclusters were chosen as nanocluster i and nanocluster j (Figure 1) and they were positioned in a simulation domain at center of masses $r_i = (-x_0, b, 0)$ and $r_j = (0, 0, 0)$ with random orientations. x_0 was fixed at 10 nm while the impact parameter, b , was varied in the range from 0 to 8 nm, significantly larger than nanocluster radii. In such simulations, the nanocluster-nanocluster initial distance ($\sqrt{x_0^2 + b^2}$) needs to be long enough such that potential interactions are negligible initially in comparison to the translational kinetic energy used to initialize motion. We found that the minimum initial distance, $x_0 = 10$ nm was sufficient in this study as long-range Coulombic interactions between two “neutral” nanoclusters diminished at this distance. At the same time, we remark that for net charged nanoclusters, particularly with attractive interactions, it would be necessary to increase this distance, and likely to consider the influence of neutral gas on nanocluster trajectories as in Tamadate *et al.*^{28, 30, 35}. An initial center-of-mass translational velocity $v_{ij} = (v_0, 0, 0)$ in the range of 20 to 700 m/s was given to nanocluster i , while the nanocluster j was initialized with a zero center-of-mass velocity, where the range of v_0 is determined to sufficiently cover the nanocluster Maxwell-Boltzmann thermal speed distribution at standard to high temperature (300 K to 1000 K). The velocities of individual atoms were the sum of the center-of-mass velocity and their thermal velocities resulting from thermal relaxation simulations. MD-trajectory simulations were performed in the absence of neutral gas, in the micro-canonical ensemble (constant number, volume, and energy: NVE) with varying the conditions ($D_{p,i}$, $D_{p,j}$, b , v_0). The combined conditions of simulations performed are summarized

in table S1 of supporting information and also mapped visually in figures S1 to S4 of the supporting information. As shown in these table and figures, some of the impact parameter calculation ranges are limited to reduce the total MD simulations since MD-trajectory calculation outcomes are predictable when impact parameter is far away from the critical impact parameter (See Figure 2). The distance between two nanocluster center-of-masses, r_{ij} was continuously tracked during each MD simulation to determine the simulation outcome: (1) collision, when the distance is smaller than the collision distance ($r_{ij} < a_{ij}$) for 200 ps and (2) non-collision, when the square of the distance exceeds 200 nm^2 ($r_{ij}^2 > 200 \text{ nm}^2$). Once one of the events (1) or (2) was detected, the calculation in question was ceased and the outcome recorded. We note that all of simulations performed in this study were classified as one of the events (collision or no-collision) in the calculation total simulation time of 50 ns and collision induced dissociation was never observed in the tested velocity range ($< 700 \text{ m/s}$). Each set of trajectory initial conditions ($D_{p,i}$, $D_{p,j}$, b , v_0) was repeated 10 times. While this is a coarser approximation for $P(b, v_0)$ than in prior simulations, because $P(b, v_0)$ is typically either near “1” or “0” for all input conditions, 10 trajectory calculations appear to be reasonable for L_{ij} estimation.

MD-trajectory calculations are of course strongly dependent on the potential interaction model selected and the extent to which this model accurately describes the chemical properties of the nanoclusters under examination. For TiO_2 potential interactions, we utilized the Matsui-Akaogi potential³⁶; this model has been utilized previously in TiO_2 nanocluster collision studies^{17, 23, 37}. The potential between atoms (including Ti-Ti, Ti-O, and O-O) are given by the expression

$$U_{kl} = \frac{C_k C_l}{r_{kl}^6} + f_0 (B_k + B_l) \exp\left(\frac{A_k + A_l - r_{kl}}{B_k + B_l}\right) + \frac{q_k q_l}{4\pi\epsilon_0 r_{kl}} \quad (5)$$

where, subscription k and l are the indexes of the atoms, ϵ_0 is permittivity of vacuum, A , B , C , and f_0 are the parameters for Matsui-Akaogi potential whose values are provided in the supporting

information, adjusted for the atom type. The Matsui-Akaogi potential is the summation of the Buckingham and the Coulomb potentials, and hence considers short-range attractive (1st term) and repulsive interactions (2nd term), as well as Coulomb interactions (3rd term) resulting from partial charges. The partial charges were assigned electron equivalent values of $q = +2.196$ for titanium and $q = -1.098$ for oxygen atoms, in accordance with the Matsui-Akaogi potential. The time step was fixed to 0.2 fs in the thermal relaxation process and 1.0 fs in the coagulation simulations. A smaller time step was required in the structural relaxation process, particularly for the calculation of the initial nanocluster architect, where high initial velocities after obtaining structures from bulk crystals necessitated smaller time steps. Time steps of 1.0 to 1.6 fs have been widely applied in MD simulations with Matsui-Akaogi force field including the original study³⁶ as well as in prior TiO₂ nanocluster calculations^{17, 23, 37, 38}.

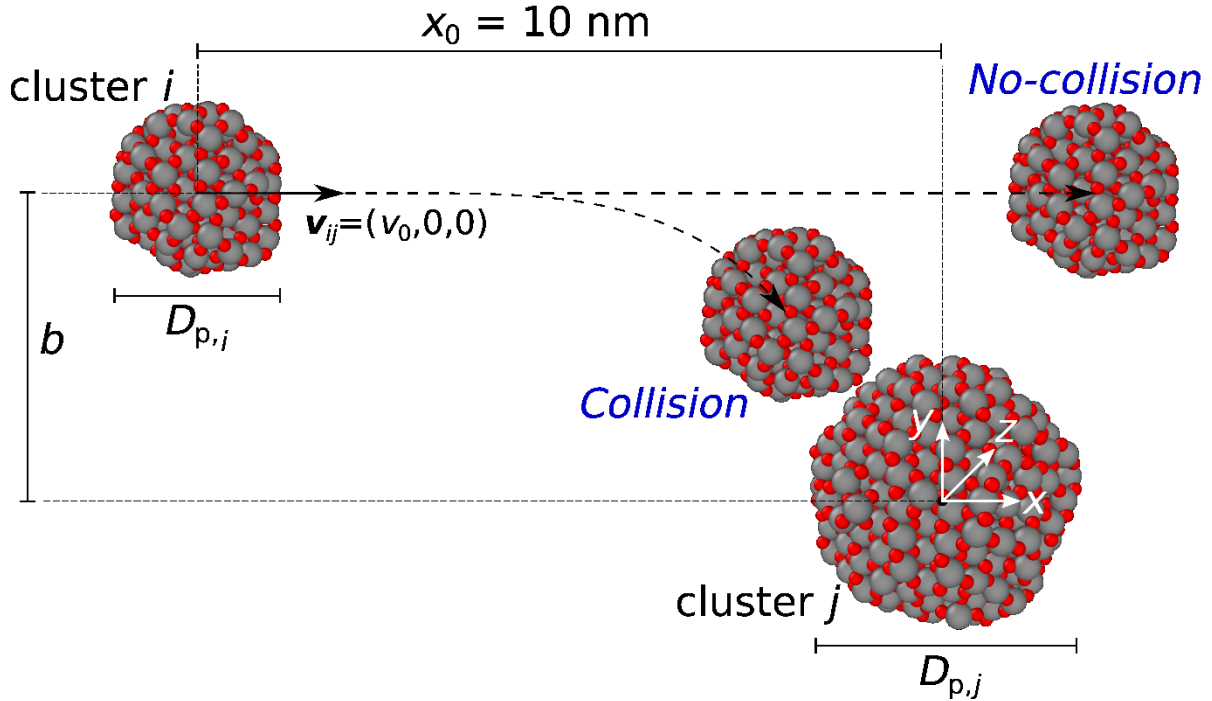


Figure 1. A schematic diagram depicting the MD trajectory calculation setup.

C. Neural network (NN) setup

After obtaining $P(b, v_0)$ maps through trajectory calculations, for prescribed $D_{p,i}$, $D_{p,j}$ combinations, we parametrized a NN coded in Python (version 3.8.10) using the Pytorch module³⁹ of v1.12.1 with a CPU. We assumed that the coagulation probability is expressed as a simple classification problem and the NN was created to generate a one-dimension output label, i.e. collision ($P = 1$) or no-collision ($P = 0$) from a four-dimensional input parameter set ($D_{p,i}$, $D_{p,j}$, b , v_0). Based on the obtained outcome coagulation probability maps both in this work and elsewhere, we find that the binary output for the coagulation probability is reasonable for most (b, v_0) input pairs. The NN output layer was composed of two parameters, the probabilities of (1) coagulation and (2) no-collision; the higher probability label is employed as the output label. Three hidden layers were set in between the input and output layers with 10, 5, and 5 neurons, respectively. Cross-entropy loss was applied as a loss function and the Adam optimizer was adopted for updating weights and biases at each layer⁴⁰. These NN hyperparameters were selected after examining 1-3 layers consisting of 5-10 neurons per layer and selecting the network deemed most accurate after training. The training data sets were prepared from the MD simulation results as the input parameters ($D_{p,i}$, $D_{p,j}$, b , v_0) and outcome labels created by rounding collision probabilities from MD results. We trained the NN model under the assumption that nanocluster i is larger than nanocluster j ($D_{p,i} \geq D_{p,j}$). Although we performed MD simulations with ($D_{p,i} < D_{p,j}$), such data sets were introduced as training data sets reversing two nanocluster diameters under assumption that the reversed particle size pairs, ($D_{p,i}$, $D_{p,j}$) and ($D_{p,j}$, $D_{p,i}$), yield an identical result. The training data sets ($N = 41,600$) were randomly split 8:2 for the training and the test process in each epoch. The number of epochs was 150 to reach converged losses and accuracy while avoiding overtraining. The model accuracy was defined as a ratio of the correct labels to the total labels, as is

commonplace in NN training for classification problem. After constructing the NN model, we evaluated the model through comparing the MD-predicted and the NN-predicted critical impact parameters as a function of velocity for coagulation to occur. From MD trajectory calculation results, the critical impact parameter was calculated by linearly fitting the collision-to-no-collision transition on the probability map, using probability ($P = 0.5$) to define the critical impact parameter, $b_{c,MD}$. For the NN model, the probability map was also created and NN critical impact parameter, $b_{c,NN}$ was determined directly from the boundaries between coagulation and no-collision outcomes. We also calculated the NN coagulation rate coefficient, $\beta_{ij,NN}$ from the NN-predicted collision probability map through equation (1) and $\beta_{ij,NN}$ was compared with MD coagulation rate coefficient, $\beta_{ij,MD}$.

In addition to the NN, we also attempted to develop a logistic regression model. In the training process for the logistic regression model, the 41,600 data sets were randomly split into 8:2 for training and testing, as they were for the NN. The label accuracy, critical impact parameter, $b_{c,NN}$, and coagulation rate coefficient, $\beta_{ij,NN}$, were also compared with MD predicted values.

III. RESULTS AND DISCUSSION

A. MD-Trajectory Calculations

Coagulation probability heat maps are provided in Figure 2, where the diameter of the center nanocluster was fixed at 2.0 nm and the incoming nanocluster diameter was varied from 0.6 to 3.0 nm. We note that the original heat map point densities are 35×20 (numbers of impact parameters $\times$ initial velocities tested) but we linearly interpolated the original grid to display 50x higher resolution heat maps. Heat maps for the other size combination results are available in the supporting information (Figure S5-7). Overall, features of the heat maps are similar for different

combinations, and in general: (1) all heat maps show a larger critical impact parameter, b_c (where boundary of coagulation to no-collision occurs, i.e. red to purple color transition in Figure 2) than the hard sphere critical impact parameter, $b_c = a_{ij}$, which is drawn by black dashed line in the Figure over entire v_0 range; (2) the critical impact parameters, b_c , are close to the hard sphere collision distance at higher initial velocities, v_0 . The first feature demonstrates that the coagulation rate coefficient predicted by MD trajectories calculations will be larger than the hard sphere approximation (equation 3), hence the enhancement factor is also more than unity. This increase is caused by the attractive van der Waals and dipole-dipole potential interactions between two nanoclusters. From (2), we infer that the coagulation rate may be better described by the hard sphere approximation at high velocity, i.e., high temperature. We also note larger critical impact parameter differences between MD and hard sphere theory with smaller nanoclusters, specifically $D_{p,j} = 0.6$ and 0.8 nm in Figure 2. The smaller nanoclusters experience proportionally larger attractive acceleration due to their small masses, an observation consistent with earlier MD-trajectory calculations^{22, 24}. In this study, we extensionally confirmed this finding while systematically varying the nanocluster size combinations in wide nanocluster size range ($D_p = 0.6$ to 3.0 nm) with fairly higher size resolution in comparison to prior studies.

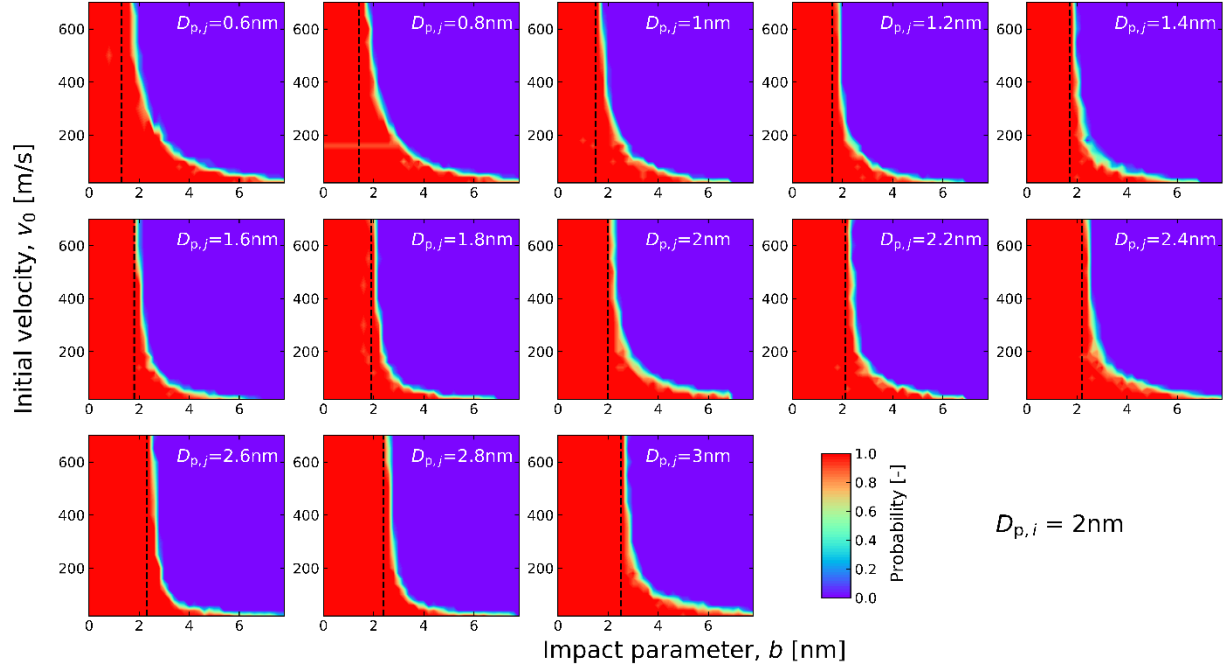


Figure 2. Coagulation probability heat maps $P(b, v_0)$ as functions of the impact parameter and the initial velocity when the center particle diameter, $D_{p,i}$ is fixed to 2.0 nm and the second particle diameter, $D_{p,j}$ is varied from 0.6 to 3.0 nm. The vertical dashed line denotes the critical collision parameter of the hard sphere model ($b = a_{ij}$).

Figure 3 shows the binding length maps of four selected nanocluster combinations at 300, 600, and 900 K. These binding length maps are obtained directly from coagulation probability maps; they provide insight into the coagulation process as they demonstrate the most probable impact parameter and velocity combinations at which coagulation occurs for a given set of nanocluster diameters and a system temperature. Dashed vertical lines in Figure 3 denote the hard-sphere collision radius, while dashed horizontal lines denote the mean thermal speed of the reduced mass of the nanoclusters. The hard sphere model assumes that coagulation occurs for all velocities, but only in instances where the impact parameter is less than or equal to the hard sphere collision radius. However, similar to the findings of Goudeli et al²², the binding length maps obtained in this study demonstrate that the most probable encounters leading to coagulation are for impact parameters greater than the hard sphere collision radius and at speeds which can exceed the mean thermal speed. Importantly, binding length heat maps also show that “head-on”, zero impact

parameter coagulation events are extremely rare, and instead coagulation occurs with nanocluster oriented at “glancing” angles with respect to one another. The newly formed nanocluster will thus have an elevated rotational energy about its center of mass. Results suggest that models of particle collision processes should not utilize the zero-impact parameter, head-on collision framework, though this is commonplace⁴¹⁻⁴³. With increasing temperature, the region of (b, v_0) space where coagulation is most probable shifts to the upper left quadrant in all plots. The velocity shift is clearly caused by the raising initial velocity at higher temperature, while the shift in impact parameter range is due to diminished influence of attractive potentials at high initial kinetic energies. We note that the high probability region of the binding length heat map at 1200 K for two 0.6 nm nanoclusters coagulation is still bounded by, but is close to the initial velocity upper limit, $v_0 = 700$ m/s, suggesting 1200 K is the highest temperature which can be examined using the calculations reported here.

Figure 4 displays plots of the coagulation rate coefficients calculated for selected nanocluster sizes, while Table 1 displays coagulation rate coefficients for all examined nanocluster pairs at selected three temperatures (300 K, 500 K, and 1,000 K). Beginning first by examining each subfigure individually, we find that in all cases lower coagulation rate coefficients are obtained when one of the nanocluster diameters is either 1.2 nm or 1.8 nm, suggesting the coagulation rate coefficient curve with size typically has a local minimum in the intermediate nanocluster size range. More specifically, nanocluster pairs corresponding to minima in curves for fixed size of nanocluster i are shaded in Table 1 which is $D_{p,j} = 1.0$ to 1.2 nm for $D_{p,i} = 0.6$ nm, $D_{p,j} = 1.0$ nm for $D_{p,i} = 1.0$ nm, and $D_{p,j} = 1.2$ to 1.6 nm for $D_{p,i} = 2.0$ nm. Such a minimum must arise from a combination of factors; the coagulation rate increases with the physical cross-section of the colliding nanoclusters, but also decreases with increasing the mass of the smaller nanocluster

(which more strongly affects the thermal speed), and potential interactions are affected by nanocluster size and encountering nanocluster size disparity in a complex manner. Without considering the potential interactions, such as in equation (3), the nanocluster size combination exhibiting the local minimum is simplified; the size ratio of the two nanoclusters $\frac{D_{p,j}}{D_{p,i}} = 0.856$ leads to the local minimum in rate coefficient for nanocluster i (see section 4 in Supporting information). However, MD-trajectory calculations yielded particle size-dependent local minimum nanocluster size ratios, from 1.67 - 2.0 for 0.6 nm, 1.0 for 1.0 nm, and 0.6 - 0.8 for 2.0 nm nanoclusters. This suggests the potential interaction has a significant influence on the coagulation rate coefficient for nanoclusters, and beyond observation of a minimum coagulation rate coefficient for the intermediate size range examined (1 nm to 1.6 nm), it is difficult to make generalizations about the size-dependency of the coagulation rate coefficient from MD trajectory calculations (we remark this is a common finding in such studies²²). Furthermore, coagulation rate coefficients generally have a weak-dependency on temperature, with rates increasing in some instances with temperature, and decreasing in others, and yet still other size combinations with a local minimum in the coagulation rate coefficient versus temperature curves. The hard sphere collision rate coefficient, and correspondingly hard sphere rates multiplied by a temperature dependent enhancement factor, have $\beta_{ij} \propto \sqrt{T}$ (c.f. equation 3). Weak and fluctuating temperature dependency is due to the interplay between the influence of translational kinetic energy and potential energy on the coagulation. Increased translational energy of course increases the frequency of nanocluster encounters due to increased velocities, but reduces the region of impact parameter space in length where successful coagulation events occur. Combined, variations in coagulation rate coefficients with particle diameters and with temperature shown rather non-

systematic variations which at present appear difficult to parameterize via simple models. For this reason, we turn to a neural network approach in the subsequent section.

The coagulation rate coefficients in Figure 4 were also converted to enhancement factors, η_{ij} , through Equation (4), and are shown in Figure 5. Figure 5 is also smoothed to have higher resolution, akin to Figure 4. Because enhancement factors are normalized by hard-sphere coefficients, they monotonically decrease with increasing temperature. For most of the diameters and temperatures examined, the enhancement factor is less than 10, in agreement with prior studies^{11, 17}. The much higher enhancement factor exceptionally obtained with 0.6 nm – 0.6 nm coagulation is caused by the large critical impact parameters, relative to the hard sphere collision radius, leading to coagulation for these nanoclusters; this is shown in the coagulation probability of map Figure S5. We note that the number of atoms in 0.6 nm nanocluster is 9 (Ti_3O_6) and it is the out of the nanocluster size ranges (below) in previous studies. Yang et al²⁴ calculated smaller enhancement factors (< 2.0) for gold nanocluster-monomer atom collisions, but such encounters are subject to weaker potential interactions as a dipole-dipole potential is not present (as it is for TiO_2 nanoclusters).

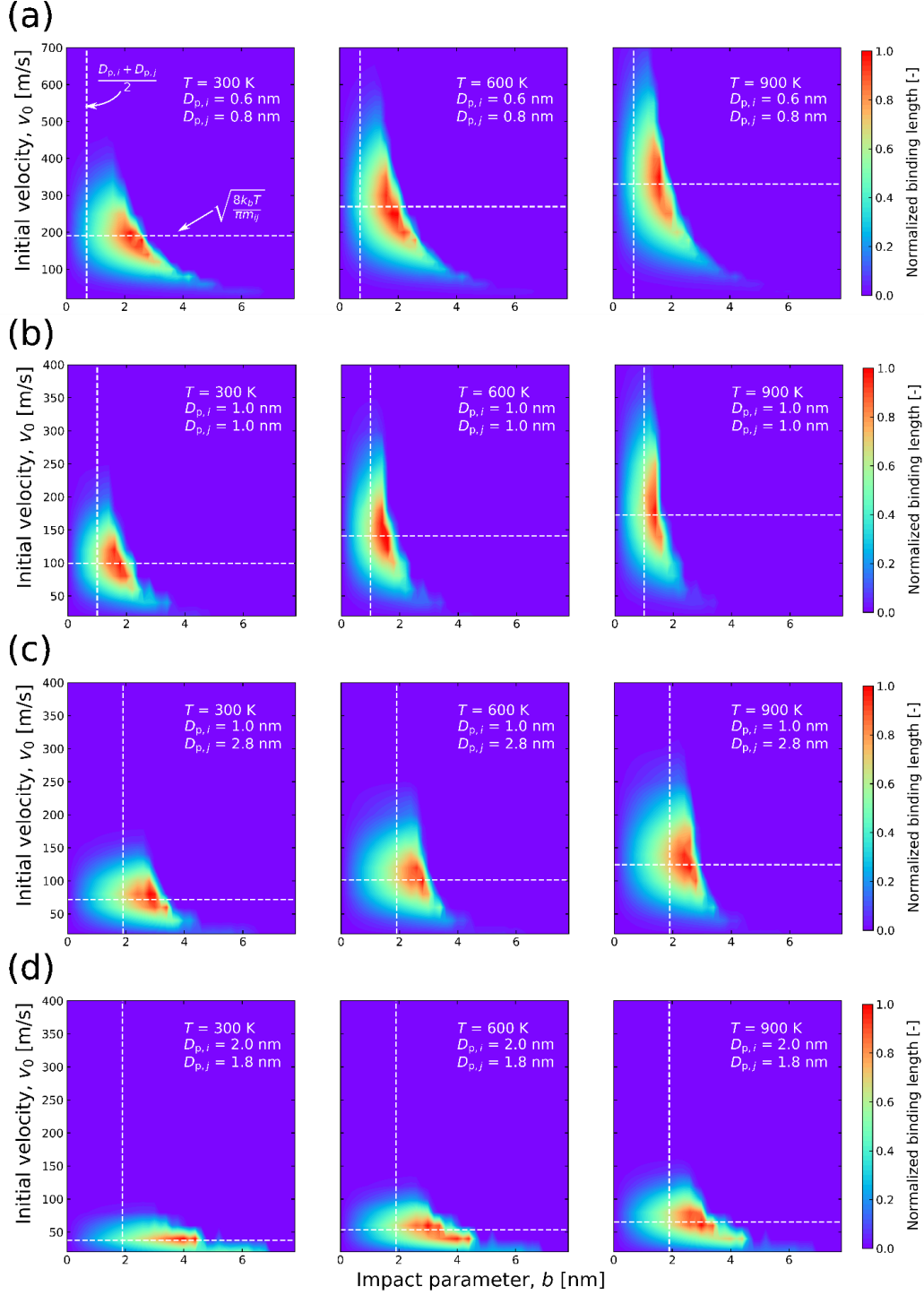


Figure 3. Integrand heat maps of the 4 different particle size combinations $(D_{p,i}, D_{p,j}) = (0.6 \text{ nm}, 0.8 \text{ nm})$ (a), $(D_{p,i}, D_{p,j}) = (1.0 \text{ nm}, 1.0 \text{ nm})$ (b), $(D_{p,i}, D_{p,j}) = (1.0 \text{ nm}, 2.8 \text{ nm})$ (c), $(D_{p,i}, D_{p,j}) = (2.0 \text{ nm}, 1.8 \text{ nm})$ (d) at 3 different temperatures ($T = 300, 600, 900 \text{ K}$). The horizontal lines express the mean thermal speeds and vertical lines are hard sphere collision distances. The binding length is normalized by the maximum value in each temperature and size combination.

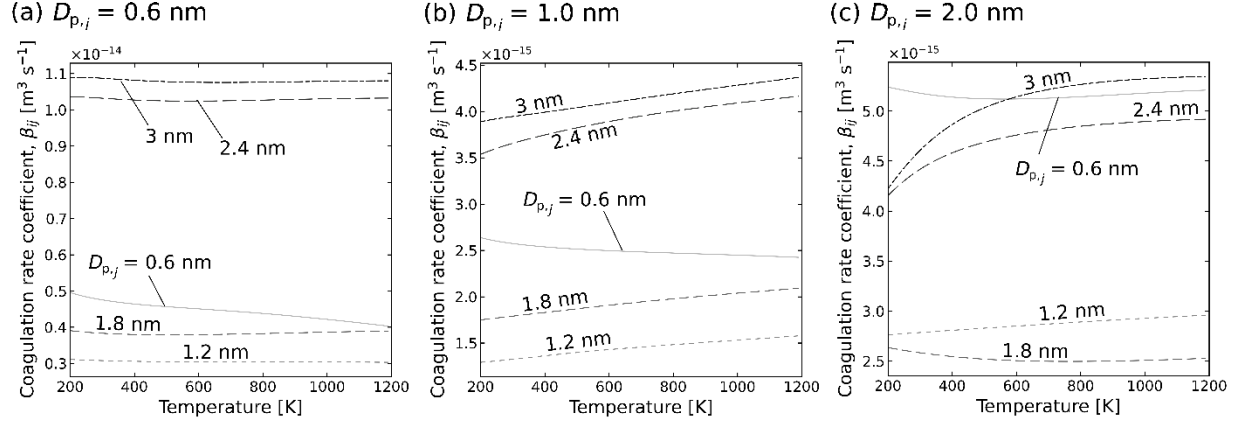


Figure 4. Coagulation rate coefficients, β_{ij} as functions of temperature for selected diameter combinations. The diameter of the nanocluster i is fixed in each panel at $D_{p,i} = 0.6$ nm (a), 1.0 nm (b), and 2.0 nm (c), while size of the nanocluster j is varied $D_{p,j} = 0.6$ nm, 1.2 nm, 1.8 nm, 2.4 nm, and 3.0 nm.

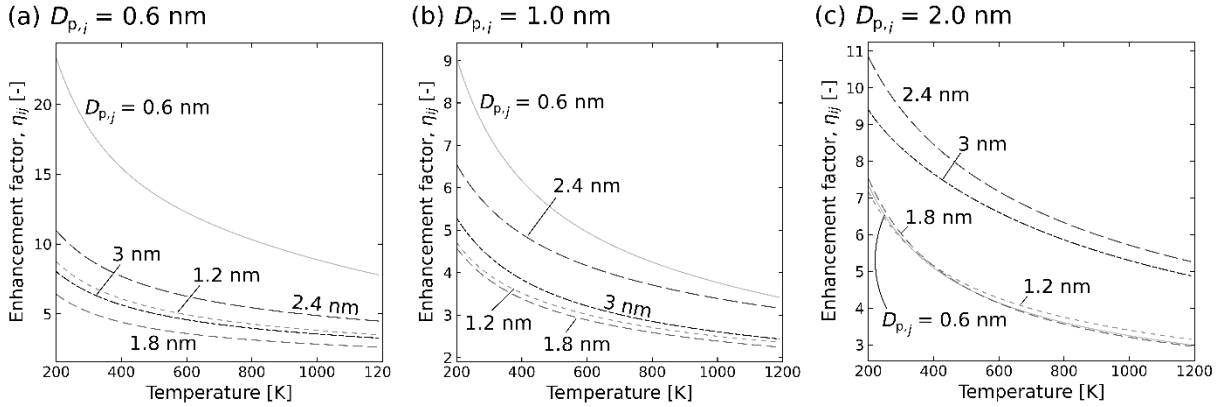


Figure 5. Enhancement factors, η_{ij} as functions of temperature for selected diameter combinations. The diameter of nanocluster i is fixed in each panel at $D_{p,i} = 0.6$ nm (a), 1.0 nm (b), and 2.0 nm (c), while size of the nanocluster j is varied $D_{p,j} = 0.6$ nm, 1.2 nm, 1.8 nm, 2.4 nm, and 3.0 nm.

B. Neural network prediction

We calculated collision probabilities, P from 41,600 different input conditions ($D_{p,i}$, $D_{p,j}$, v_0 , b) in MD-trajectory calculations, utilizing a NN to predict the binary coagulation probability based on these results. The loss functions and accuracy of the labels during the NN training process are shown in the supporting information (Figure S8). Both of loss function and accuracy

converged with increasing epochs, where we confirmed that the accuracy reached a sufficiently reliable value ($> 95\%$) within 150 epochs in both of the training and testing samples. We note that this training process includes the obvious collision and no-collision initial conditions, e.g. those of high initial velocity and large impact parameter, hence the accuracy at the boundary of the outcomes is predicted at lower than 95% accuracy. Figure 6a displays comparison of the critical impact parameter obtained by MD simulation (a function of velocity, and determined as the point where the coagulation probability is 0.5) and the prediction via trained NN model. The coefficient of determination (R^2) was 0.9544, suggesting the NN is a sufficiently reliable model for coagulation probability prediction. We also note that most of the points plotted in Figure 6a fall within a shaded gray region, which is $0.8 - 1.25 (= 1/0.8)$ times from the 1:1 line; more than 95 % of data fall in this region. The NN is particularly accurate at higher velocities, likely because higher velocities lead to better convergence with the well-behaved hard sphere model. The larger disagreement observed for larger nanoclusters is presumably due to their increased degrees of freedom, leading to greater uncertainty in the dipole moment orientation throughout the coagulation process and complicating its influence on trajectories; this in turn complicates NN training. We further compare MD-trajectory and NN predictions by directly plotting the calculated coagulation rate coefficients, i.e., $\beta_{ij,MD}$ vs. $\beta_{ij,NN}$ where “MD” and “NN” denote molecular dynamics and the neural network, respectively, in Figures 6b and 6c (with the ratios in Table 1). The coefficients of determination at 300 K (6b) and 1000 K (6c) are 0.8865 and 0.9204, respectively, suggesting again that higher velocities yield improved NN predictions. There is a higher degree of variability between the coagulation rate coefficients from NN calculations and MD-trajectory calculations than observed for the critical impact parameter, as the coagulation rate coefficients derive from the squares of critical impact parameters, magnifying deviations.

Nonetheless, more than one standard deviation of the distribution ($> 68.3\%$; 71.7% and 74.4% at 300 K and 1000 K) of NN-predicted rate coefficients fall within the shaded $0.8 - 1.25$ ($1/0.8$) region and two-standard deviations of the distribution ($> 95.0\%$; 97.4% and 100% at 300 K and 1,000 K) are within $0.65 - 1.54$ ($= 1/0.65$), suggesting the NN approach will be much more appropriate to describe TiO_2 nanocluster growth than use of a constant enhancement factor in conjunction with the hard sphere model.

Also for reference with a simpler classification model, logistic regression parameterization is compared with MD predictions in Figure S9 of the supporting information. While outcome label accuracy was 0.91, the coefficient of determination for critical impact parameter was 0.7321, and for coagulation rate coefficient at 300 K it was 0.6724, which is unsatisfactory for implementation. This further demonstrates the present need for non-linear classification algorithms, such as the NN approach, though future studies may find alternative approaches more accurate for coagulation rate coefficients. Towards implementation, those interested in using the NN results may download it at https://github.com/tamadate/TiO2_collision_rate_NN. Because the NN model is trained with the MD trajectory calculations, the input parameter range is limited to $0 \text{ nm} \leq b \leq 8.0 \text{ nm}$, $20 \text{ m/s} \leq v_0 \leq 700 \text{ m/s}$, $D_p \leq 3.0 \text{ nm}$, with an approximate upper temperature limit of 1200 K. The extension of coagulation rate coefficient estimation to larger nanoclusters also remains the subject of future work. We do expect that the larger nanoclusters will yield closer agreement with hard sphere theories or with models considering simpler, angularly averaged (spherically symmetric) potential interactions since (1) the length scale of the potential interaction becomes short-distance relatively to the nanocluster diameters and (2) the nanocluster acceleration is inversely-proportional to nanocluster mass. Future larger scale

simulations will tell us the boundary of the size of nanoclusters where the NN remains accurate and where simpler theories can be applied.

We also use NN results to predict the coagulation rate coefficients and enhancement factors at 300 K and 1000 K, respectively, with a variety of nanocluster size combinations with results shown in Figure 7. Enhancement factors are also displayed in Table 1. The predicted heat maps are smooth and the range of both the rate coefficients and the enhancement factors fall well-within the range expected based on MD-trajectory simulations (2 to 20). In the coagulation rate coefficient heat maps shown in Figure 7a and 7b, local minimum values are evident in the nanocluster diameter range of 1.0 – 1.5 nm, consistent with MD simulations. In addition, as shown in Figure 7c, significantly higher enhancement factors are obtained for small nanoclusters at low temperature.

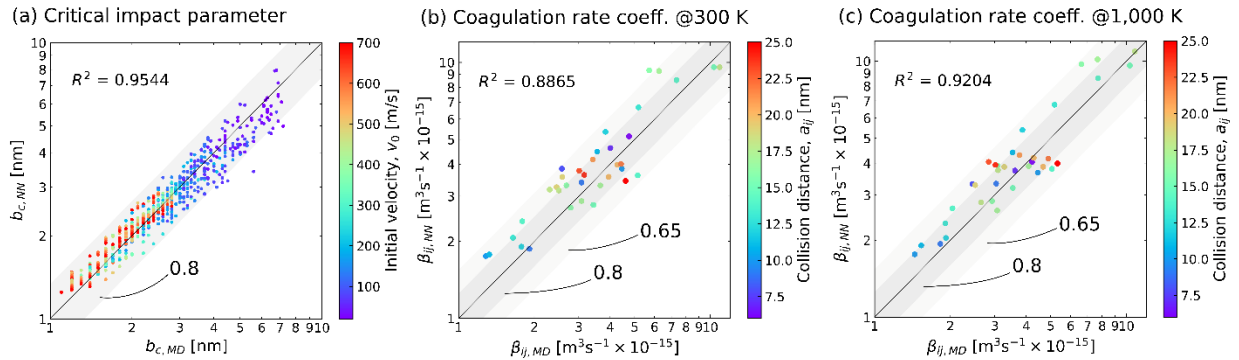


Figure 6. NN- predicted versus MD-predicted critical impact parameters, b_c (a), coagulation rate coefficient, β_{ij} at 300 K (b), and coagulation rate coefficient, β_{ij} at 1,000 K (c). The plot color denotes the initial velocity, v_0 in (a) and collision distance a_{ij} in (b) and (c). The diagonal line is the 1:1 line ($b_{c,NN} = b_{c,MD}$ or $\beta_{ij,NN} = \beta_{ij,MD}$), dark gray shadings are 0.8 to 1.25 (1/0.8) times from the 1:1 line, and light gray shadings are 0.65 to 1.54 ($= 1/0.65$). R^2 is the coefficient of determination.

Table 1. Comparison of MD-determined coagulation rate coefficients $\beta_{ij,\text{MD}}$ and enhancement factors $\eta_{ij,\text{MD}}$ along with the ratio of MD and NN -determined coagulation rate coefficients $\chi_{ij} = \beta_{ij,\text{NN}}/\beta_{ij,\text{MD}}$ at 300 K, 500 K, and 1,000 K. The unit of D_p is [nm] and $\beta_{ij,\text{MD}}$ is [$\text{m}^3 \text{s}^{-1} \times 10^{-15}$]. $\eta_{ij,\text{MD}}$ and χ_{ij} are dimensionless. The shaded numbers denote the local minimums for three base particle diameters (0.6, 1.0, 2.0 nanometers) and three temperatures (300, 500, and 1,000 Kelvins).

$D_{p,i}$	$D_{p,j}$	$\beta_{ij,\text{MD}}$ 300 K	$\eta_{ij,\text{MD}}$ 300 K	χ_{ij} 300 K	$\beta_{ij,\text{MD}}$ 500 K	$\eta_{ij,\text{MD}}$ 500 K	χ_{ij} 500 K	$\beta_{ij,\text{MD}}$ 1,000 K	$\eta_{ij,\text{MD}}$ 1,000 K	χ_{ij} 1,000 K
0.6	0.6	4.75	19.60	1.08	4.65	14.59	1.02	4.25	9.52	1.01
0.6	0.8	4.01	14.41	1.15	4.24	10.62	1.11	3.94	7.11	1.09
0.6	1.0	2.57	7.66	1.49	3.61	5.80	1.44	3.49	3.99	1.42
0.6	1.2	3.08	7.52	1.10	3.32	5.77	1.09	3.50	4.07	1.15
0.6	1.4	4.45	8.99	0.86	3.77	6.87	0.86	3.86	4.92	0.87
0.6	1.6	3.55	6.01	1.26	4.44	4.58	1.27	4.64	3.27	1.31
0.6	1.8	3.83	5.49	1.40	5.39	4.20	1.42	5.67	3.03	1.47
0.6	2.0	5.18	6.35	1.29	6.76	4.87	1.32	7.11	3.48	1.37
0.6	2.2	7.33	7.78	1.16	8.70	6.12	1.17	9.03	4.48	1.17
0.6	2.4	10.30	9.56	0.93	9.76	7.35	0.96	10.10	5.23	0.98
0.6	2.6	5.71	4.66	1.63	9.58	3.81	1.59	10.30	2.96	1.56
0.6	2.8	6.25	4.52	1.48	9.62	3.80	1.42	10.70	3.04	1.40
0.6	3.0	10.90	7.02	0.88	9.97	5.39	0.92	11.50	3.81	1.06
1.0	0.6	2.57	7.66	1.49	3.61	5.80	1.44	3.49	3.99	1.42
1.0	0.8	1.91	6.21	0.98	1.88	4.65	1.02	2.03	3.24	1.12
1.0	1.0	1.28	4.10	1.37	1.76	3.27	1.33	1.85	2.53	1.28
1.0	1.2	1.32	3.94	1.35	1.84	3.22	1.31	2.01	2.49	1.31
1.0	1.4	1.78	4.77	1.07	2.01	3.77	1.11	2.15	2.82	1.12
1.0	1.6	1.65	3.94	1.25	2.21	3.23	1.27	2.45	2.51	1.28
1.0	1.8	1.79	3.81	1.34	2.57	3.08	1.37	2.80	2.38	1.37
1.0	2.0	3.43	6.49	0.81	3.02	5.06	0.88	3.37	3.66	0.95
1.0	2.2	2.68	4.52	1.24	3.63	3.68	1.29	4.11	2.85	1.33
1.0	2.4	3.66	5.52	0.90	3.60	4.47	0.94	4.09	3.38	1.00
1.0	2.6	2.30	3.12	1.38	3.45	2.72	1.34	3.95	2.30	1.28
1.0	2.8	2.45	3.00	1.33	3.47	2.58	1.28	4.07	2.19	1.24
1.0	3.0	3.95	4.38	0.90	3.81	3.47	0.94	4.41	2.60	1.03
2.0	0.6	5.18	6.35	1.29	6.76	4.87	1.32	7.11	3.48	1.37
2.0	0.8	5.14	8.23	0.70	3.73	6.32	0.73	4.00	4.43	0.79
2.0	1.0	3.44	6.51	0.81	3.02	5.07	0.87	3.37	3.67	0.95
2.0	1.2	2.79	5.86	0.97	2.76	4.61	0.98	2.99	3.37	1.02
2.0	1.4	3.08	6.88	0.93	2.79	5.31	0.91	2.75	3.75	0.90
2.0	1.6	2.50	5.74	1.26	3.09	4.53	1.21	2.95	3.32	1.12
2.0	1.8	2.58	5.94	1.38	3.53	4.49	1.40	3.45	3.16	1.37
2.0	2.0	4.26	9.63	0.93	4.03	7.97	0.89	4.07	5.77	0.87
2.0	2.2	3.41	7.47	1.23	4.40	6.12	1.22	4.53	4.51	1.20
2.0	2.4	4.42	9.29	0.91	4.24	7.63	0.90	4.38	5.63	0.90
2.0	2.6	3.01	6.04	1.25	4.03	4.51	1.39	4.25	3.11	1.50
2.0	2.8	3.15	5.99	1.15	3.91	4.49	1.28	4.14	3.14	1.37

2.0	3.0	4.61	8.27	0.75	3.77	7.00	0.75	4.20	5.24	0.79
-----	-----	------	------	------	------	------	------	------	------	------

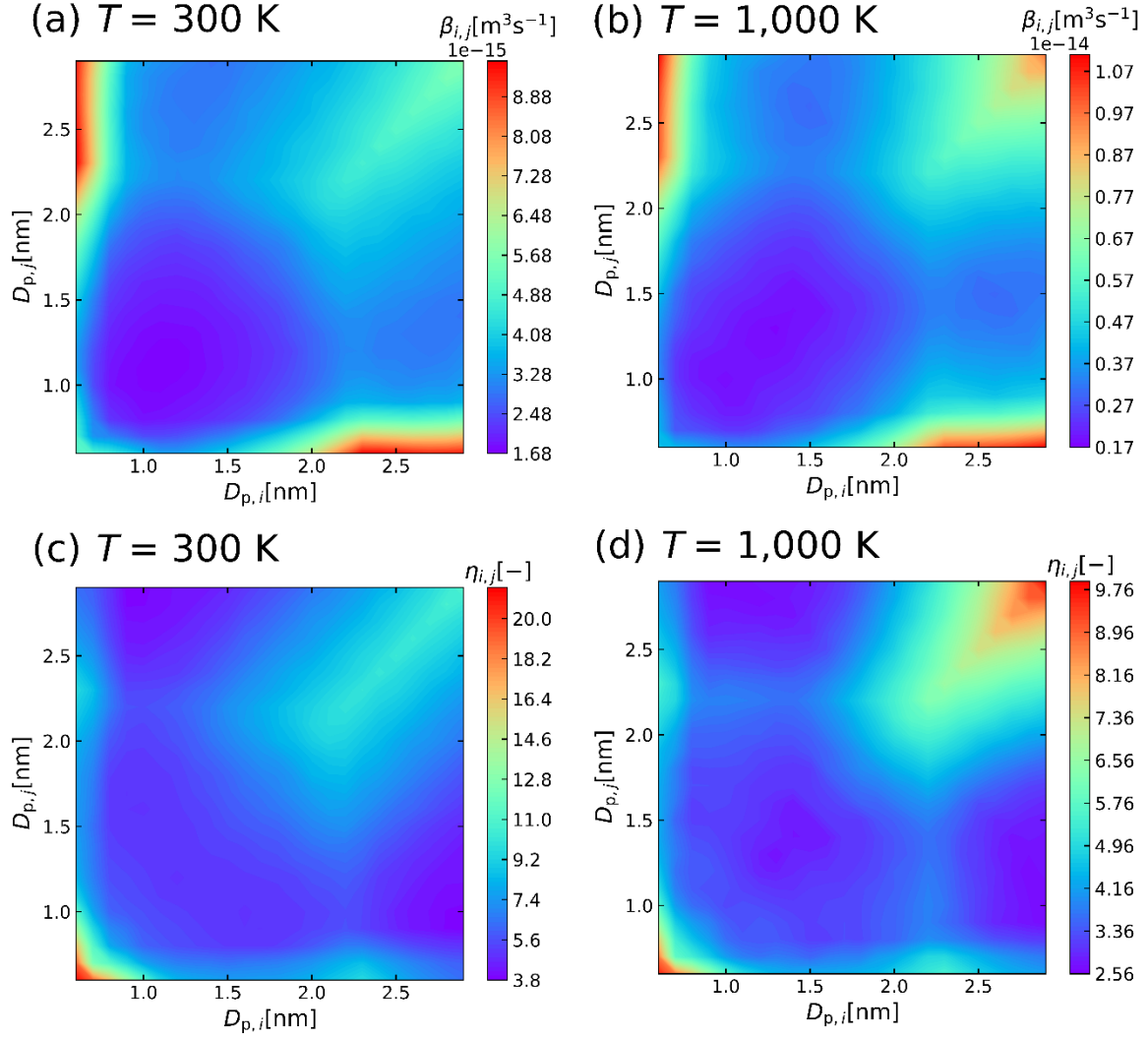


Figure 7. NN-predicted coagulation rate coefficients, β_{ij} (a-b) and enhancement factors, η_{ij} (c-d) maps at two temperatures, 300 K (a & d) and 1000 K (c & d) with different particle diameter combinations ($D_{p,i}$, $D_{p,j}$).

IV. CONCLUSIONS

We computed the trajectories of TiO_2 nanoclusters *via* MD simulations, using them to compute coagulation rate coefficients for equal and unequal sized nanoclusters with diameters in the 0.6 – 3.0 nm size range. Similar to prior efforts utilizing MD-trajectory calculations to study coagulation, we find that the TiO_2 nanocluster coagulation rate is significantly greater than the hard sphere coagulation rate, with enhancement factors in the 20 to 2 range as the temperature increases from 300 K to 1000 K. Simulations suggest that coagulation rate coefficients vary non-monotonically with temperature nanocluster diameter, and nanocluster diameter disparity (relative size with respect to one another), with a unique minimum in the rate observed when one of the nanoclusters falls in the 1.0-1.8 nm diameter range (though this finding cannot be extrapolated outside the tested diameter range).

We trained a NN to predict the critical impact parameter for coagulation to occur as a function of the nanocluster diameters, and approach velocity. The trained NN shows reasonable agreement with MD-trajectory calculations for both the critical impact parameter, and the correspondingly calculated coagulation rate coefficient; most NN-predicted coagulation rate coefficients are within a factor of 0.8-1.25 from the MD-trajectory calculated rate coefficient. In addition to providing a NN which can be used to model TiO_2 nanocluster growth in high temperature gas phase systems, the approach we apply here, i.e. MD-trajectory calculations followed by NN training, is sufficiently general that it can be applied to a wide variety of systems as a means to develop robust coagulation rate coefficient models as inputs to population balance models. Improvements in both MD-modeling (more accurate potentials, consideration of variable nanocluster internal energy) and further advanced machine learning approaches may enable even more accurate and robust coagulation rate coefficient models in future work. In addition, while

the analysis presented in this study focuses on metal oxide nanoclusters growth, as would be expected in oxygen-rich synthesis systems, the approach can be adapted for metal and semimetal nanoclusters or other ceramics, synthesized in low oxygen content gas phase reactors⁴⁴⁻⁴⁶ or plasma synthesis systems^{47, 48}.

ACKNOWLEDGEMENTS

This work was supported by US National Science Foundation (NSF) Award Number 2038173. Computations were performed using computing resources of the Minnesota Supercomputing Institute (MSI) at the University of Minnesota.

SUPPORTING INFORMATION

A list of MD simulation conditions, figures noting the ranges of impact parameters and velocities used in MD-trajectory calculations, a table noting molecular dynamics potential parameters, additional coagulation probability maps from MD-trajectory calculations, loss & accuracy plots for NN-training, and a comparison of logistic regression model training to MD-trajectory calculations are available online.

REFERENCES

- ¹M. T. Swihart, *Current Opinion in Colloid & Interface Science* 8, 127 (2003).
- ²S. Li, Y. Ren, P. Biswas, and S. D. Tse, *Progress in Energy and Combustion Science* 55, 1 (2016).
- ³F. Meierhofer and U. Fritsching, *Energy & Fuels* 35, 5495 (2021).
- ⁴H. Wang, *Proceedings of the Combustion Institute* 33, 41 (2011).
- ⁵Y. Qiao, L. Li, J. Chen, S. Yang, and C. J. Hogan, *Journal of Aerosol Science* 163, 105981 (2022).
- ⁶Y. Wang, J. Kangasluoma, M. Attoui, J. Fang, H. Junninen, M. Kulmala, T. Petäjä, and P. Biswas, *Proceedings of the Combustion Institute* 36, 745 (2017).
- ⁷M. Domaschke, C. Lübbert, and W. Peukert, *Journal of Aerosol Science* 137, 105438 (2019).
- ⁸P. Buerger, D. Nurkowski, J. Akroyd, and M. Kraft, *Proceedings of the Combustion Institute* 36, 1019 (2017).
- ⁹P. Buerger, D. Nurkowski, J. Akroyd, S. Mosbach, and M. Kraft, *The Journal of Physical Chemistry A* 119, 8376 (2015).
- ¹⁰D. Hou, D. Zong, C. S. Lindberg, M. Kraft, and X. You, *Journal of Aerosol Science* 140, 105478 (2020).
- ¹¹G. Sharma, S. Dhawan, N. Reed, R. Chakrabarty, and P. Biswas, *Journal of Aerosol Science* 127, 27 (2019).
- ¹²S. Tsantilis and S. E. Pratsinis, *Journal of Aerosol Science* 35, 405 (2004).
- ¹³N. Morgan, C. Wells, M. Kraft, and W. Wagner, *Combustion Theory and Modelling* 9, 449 (2005).
- ¹⁴A. Boje and M. Kraft, *Journal of Aerosol Science* 159, 105895 (2022).
- ¹⁵N. A. Fuchs and A. G. Sutugin, *Journal of Colloid Science* 20, 492 (1965).
- ¹⁶W. H. Marlow, *J Chem Phys* 73, 6284 (1980).
- ¹⁷Y. Zhang, S. Li, W. Yan, Q. Yao, and D. Tse, *The Journal of Chemical Physics* 134, 084501 (2011).
- ¹⁸H. Ouyang, R. Gopalakrishnan, and C. J. Hogan, *The Journal of Chemical Physics* 137, 064316 (2012).
- ¹⁹S. N. Rogak and R. C. Flagan, *J Colloid Interf Sci* 151, 203 (1992).
- ²⁰M. E. Mueller, G. Blanquart, and H. Pitsch, *Proceedings of the Combustion Institute* 32, 785 (2009).
- ²¹M. E. Mueller, G. Blanquart, and H. Pitsch, *Combustion and Flame* 156, 1143 (2009).
- ²²E. Goudeli, J. Lee, and C. J. Hogan, *Journal of Aerosol Science* 146, 105558 (2020).
- ²³W. Yan, S. Li, Y. Zhang, Q. Yao, and S. D. Tse, *The Journal of Physical Chemistry C* 114, 10755 (2010).
- ²⁴H. Yang, E. Goudeli, and C. J. Hogan, *The Journal of Chemical Physics* 148, 164304 (2018).
- ²⁵R. Halonen, E. Zapadinsky, T. Kurtén, H. Vehkamäki, and B. Reischl, *Atmos. Chem. Phys.* 19, 13355 (2019).
- ²⁶I. Neefjes, R. Halonen, H. Vehkamäki, and B. Reischl, *Atmos. Chem. Phys.* 22, 11155 (2022).
- ²⁷W. G. Vincenti and C. H. Kruger, *Introduction to Physical Gas Dynamics* (Krieger, Huntington, NY, 1975).
- ²⁸T. Tamadate and C. J. Hogan, *Journal of Aerosol Science* 163, 105994 (2022).
- ²⁹S. Plimpton, *Journal of Computational Physics* 117, 1 (1995).

- ³⁰T. Tamadate, H. Higashi, T. Seto, and C. J. Hogan, *The Journal of Chemical Physics* 152, 094306 (2020).
- ³¹H. Yang, G. Song, and C. J. Hogan, *Journal of Aerosol Science* 159, 105891 (2022).
- ³²S. Nosé, *Progress of Theoretical Physics Supplement* 103, 1 (1990).
- ³³M. R. Zachariah and M. J. Carrier, *Journal of Aerosol Science* 30, 1139 (1999).
- ³⁴E. Goudeli and S. E. Pratsinis, *AIChE Journal* 62, 589 (2016).
- ³⁵T. Tamadate, H. Higashi, C. J. Hogan, and T. Seto, *Physical Chemistry Chemical Physics* 22, 25215 (2020).
- ³⁶M. Matsui and M. Akaogi, *Molecular Simulation* 6, 239 (1991).
- ³⁷B. Buesser, A. J. Gröhn, and S. E. Pratsinis, *The Journal of Physical Chemistry C* 115, 11030 (2011).
- ³⁸V. V. Hoang, *Nanotechnology* 19, 105706 (2008).
- ³⁹A. Paszke *et al.*, *NIPS'19: Proceedings of the 33rd International Conference on Neural Information Processing Systems*, 8026 (2019).
- ⁴⁰D. P. Kingma and J. Ba, *arXiv preprint arXiv:1412.6980* 2014).
- ⁴¹B. D. Doan, A. R. Dove, and P. K. Schelling, *Journal of Aerosol Science* 155, 105742 (2021).
- ⁴²T. Hawa and M. R. Zachariah, *Physical Review B* 69, 035417 (2004).
- ⁴³E. Guevara-Chapa and S. Mejía-Rosales, *Journal of Nanoparticle Research* 16, 2757 (2014).
- ⁴⁴L. Wergen, M. Domaschke, P. Herre, M. Wu, E. Spiecker, and W. Peukert, *Journal of Aerosol Science* 128, 50 (2019).
- ⁴⁵T. Rosenberger, J. Neises, D. Kiesler, and F. E. Kruis, *Journal of Aerosol Science* 144, 105531 (2020).
- ⁴⁶M. Snellman, N. Eom, M. Ek, M. E. Messing, and K. Deppert, *Nanoscale Advances* 3, 3041 (2021).
- ⁴⁷U. R. Kortshagen, R. M. Sankaran, R. N. Pereira, S. L. Girshick, J. J. Wu, and E. S. Aydil, *Chemical Reviews* 116, 11061 (2016).
- ⁴⁸A. Kumar, S. Kang, C. Larriba-Andaluz, H. Ouyang, C. J. Hogan, and R. M. Sankaran, *Nanotechnology* 25, 385601, 385601 (2014).