SCATTERING DYNAMICS OF RAREFIED GASES

A HYBRID ATOMISTIC/MACHINE LEARNING APPROACH

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Eindhoven,
op gezag van de rector magnificus prof. dr. ir. F.P.T. Baaijens, voor een
commissie aangewezen door het College voor
Promoties, in het openbaar te verdedigen
op 6 april 2023 om 13:30 uur

door

Shahin Mohammad Nejad

geboren te Rasht, Iran.

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommisie is als volgt:

voorzitter: prof. dr. ir. P.D. Anderson
promotor: prof. dr. ir. D.M.J. Smeulders
copromotor: dr. ir. A.J.H. Frijns
copromotor: dr. S.V. Gaastra-Nedea

leden:	prof. dr. ir. E.H. van Brummelen
	prof. dr. S. Calero
	prof. dr. P.A.J. Hilbers
	prof. A. Frezzotti (Politecnico di Milano)

Het onderzoek dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

Printed by: Proefschrift All In One

Cover design: Anastasija Mass

This work is part of the research program RareTrans with Project No. HTSM-15376, which is (partly) financed by the Netherlands Organization for Scientific Research (NWO)

A catalogue record is available from the Eindhoven University of Technology Library.
ISBN 978-90-386-5699-1

CONTENTS

Summary **ix**

1 Introduction **1**
1.1 Background and motivation . . . 2
1.2 Simulation methods for rarefied gases . . . 6
1.2.1 Molecular Dynamics . . . 7
1.2.2 Direct Simulation Monte Carlo . . . 10
1.3 Thesis outline . . . 13

2 Impact of intermolecular interactions on the accommodation coefficients **15**
2.1 Introduction . . . 16
2.2 Materials and Methods . . . 18
2.2.1 Molecular Dynamics (MD) Simulations . . . 18
2.2.2 Computing Accommodation Coefficients . . . 20
2.3 Results and Discussion . . . 21
2.4 Conclusions . . . 26

3 Modeling rarefied monoatomic gas-solid surface interactions using Gaussian Mixture Approach **27**
3.1 Introduction . . . 28
3.2 Methodologies . . . 30
3.2.1 MD simulation . . . 30
3.2.2 CLL scattering kernel . . . 31
3.2.3 Gaussian mixture model . . . 33
3.3 Results and Discussion . . . 34
3.3.1 Different approaches for feeding the training data into Gaussian mixture model . . . 34
3.3.2 Performance evaluation of GM model . . . 38
3.3.2.1 Ar-Au system . . . 39
3.3.2.2 He-Au system . . . 44
3.4 Conclusions . . . 48

4 Development of a scattering model for diatomic gas- solid surface interactions based on Gaussian Mixture approach **49**
4.1 Introduction . . . 50
4.2 Methodologies . . . 52
4.2.1 MD simulation model . . . 52
4.2.2 Scattering kernel . . . 54
4.2.3 Gaussian Mixture model . . . 55

4.3 Results and Discussion . . . 57
4.3.1 MD scheme validation . . . 57
4.3.2 GM model Assessment . . . 58
4.3.2.1 H_2-Ni system . . . 59
4.3.2.2 N_2-Ni system . . . 69
4.3.3 Impact of the adsorption layer on behavior of reflected gas molecules 77
4.4 Conclusions . . . 80

5 Generalized Gaussian Mixture wall model for rarefied gases 83
5.1 Effect of gas pressure on the post-collisional behavior of monoatomic gas molecules . . . 84
5.2 Effect of surface temperature on the post-collisional behavior of monoatomic gas molecules . . . 87
5.3 Effect of gas pressure on the post-collisional behavior of diatomic gas molecules 89
5.4 Effect of surface temperature on the post-collisional behavior of diatomic gas molecules . . . 92
5.5 Constructing a generalized GM scattering model for Ar-Au system . . . 95
5.6 Constructing a generalized GM scattering model for H_2-Ni system . . . 102
5.7 Assessment of the sampling approaches used to generate GM data . . . 107

6 A hybrid GM-DSMC method to study the Fourier thermal problem 115
6.1 Introduction . . . 116
6.2 Methodologies . . . 118
6.2.1 MD simulation . . . 119
6.2.2 Gas-surface interaction models . . . 120
6.2.2.1 CLL scattering model . . . 121
6.2.2.2 GM scattering model . . . 121
6.2.3 DSMC simulation . . . 124
6.3 Results and discussion . . . 125
6.3.1 Isothermal Ar-Au system . . . 125
6.3.2 Non-isothermal Ar-Au system . . . 128
6.3.3 Isothermal H_2-Ni system . . . 129
6.3.4 Non-isothermal H_2-Ni system . . . 132
6.4 Conclusions . . . 137

7 Conclusions and outlook 139
7.1 Conclusions . . . 140
7.2 outlook . . . 141

A Appendix - Comparison between velocity correlations, PDFs, and ACs obtained from training, testing, and the GM model for the Ar-Au system 143

B appendix-Velocity correlations and PDFs for the Ar-Au system that are not discussed in the main text 145

C Appendix-Velocity correlations and PDFs for the He-Au system that are not discussed in the main text **149**

D Appendix- Sensitivity analysis to decide the number of Gaussians used in the GM model for diatomic gas molecules **153**

E Appendix-Correlation graphs and PDFs for H_2-Ni case studies that are not discussed in the main text **155**

F Appendix-Correlation graphs and PDFs for N_2-Ni case studies that are not discussed in the main text **161**

G Appendix-Correlation graphs and PDFs of the generalized GM model for the Ar-Au system **169**

H Appendix-Correlation graphs and PDFs of the generalized GM model for the H_2-Ni system **173**

I Appendix-Sensitivity analysis to decide the number of Gaussians used in the GM model for mono/diatomic gas molecules **177**

J Appendix-Integration of the GM scattering kernel into the DSMC solver **179**

Bibliography **183**

Acknowledgements **195**

Curriculum Vitæ **197**

List of Publications **199**

SUMMARY

The topic of rarefied gas dynamics is an active research topic in numerous cutting edge engineering applications such as vacuum technology, aerospace engineering, aerosol industry, and micro/nano systems. In these applications, above a certain degree of rarefaction (Kn>0.1) important transport properties such as shear stress, pressure drop, and heat flux cannot be anymore predicted using continuum flow and heat transfer models. Under such circumstances, Boltzmann transport equations are typically applied to describe the flow field properties, which can be solved using different simulation techniques, such as Direct Simulation Monte Carlo (DSMC), Method of Moment (MoM), and Lattice Boltzmann Methods (LBM).

The reliability and accuracy of the results obtained based on all the aforementioned simulation approaches highly depend on the applied boundary conditions that describe the energy and momentum transfer between gas and the adjacent solid surfaces. The reason lies in the fact that in the rarefied gas flow systems, as Knudsen number increases, the gas-solid surface interactions in comparison with gas-gas interactions become more dominant. Therefore, in these systems a fundamental understanding of the physics of the gas-solid surface interactions is of crucial importance to model flow field properties. Over the last century, considerable efforts have been dedicated towards developing rigorous and realistic gas-surface interaction (GSI) models. However, owing to the complexity of the microscopic interaction between gas molecules and adjoining solid surfaces, the physics of GSI phenomenon are still not well understood, and reliable and generalized GSI models suitable to describe complex surface interactions and highly nonequilibrium flow in modern engineering applications are still lacking.

Molecular Dynamics (MD) simulation is the most promising approach that can be used to study gas-surface interactions. In the most common manner, an MD simulation is carried out to derive energy and momentum accommodation coefficients (ACs), which are important parameters in a GSI model, quantifying the degree at which a gas thermally accommodates to an adjacent surface. In an MD simulation, the type of pair potential employed to describe the interaction between gas and solid molecules is the most prominent factor that can affect the MD results. To investigate this issue, an MD simulation setup consisting of two parallel walls (Au atoms), and an inert gases (Ar or He) confined between them, was used as a case study. A pairwise Lennard-Jones (LJ) potential was employed to describe gas-surface interactions. The LJ potential parameters were computed based on different approximation approaches such as the Lorentz–Berthelot (LB) and the Fender–Halsey (FH) mixing rules, as well as taking from an existing quantum mechanic type computation called ab-initio technique. It has been shown that an interaction potential based on ab-initio computations is the most reliable one to compute different ACs. Besides, in the case of the Ar-Au system the LB mixing rule highly overestimates the potential well depth. This results in a fully saturated solid surface, which is not physically correct.

The obtained ACs from MD simulations can be fed into the phenomenological GSI models such as Maxwell or Cercignani–Lampis-Lord (CLL) that commonly are used as the GSI models for the higher scale simulation approaches such as DSMC. The performance of these GSI models is highly dependent on the value of the ACs. Nevertheless, it was observed that the aforementioned models, which are based on just a few limited number of parameters, in the case of complex flow systems or in the presence of highly non-equilibrium situation are incapable to describe the flow field characteristics accurately.
To fulfill this shortcoming in this work the Gaussian Mixture (GM) model, which is an unsupervised machine learning technique, was employed to construct a stochastic GSI model from MD simulation results. Here, the entire collisional data obtained from MD simulations is used directly to train the GM model, and no intermediate parameter such as the AC is needed. Considering the MD simulations results as the reference solutions, the performance of the GSI model obtained from the GM approach was assessed against the CLL model in different benchmark systems, such as the Fourier thermal problem, the Couette flow problem, and the Fourier+Couette flow problem for monoatomic and diatomic gases (Ar-Au, He-Au, H_2-Ni, N_2-Ni). The assessment has been carried out in terms of the distribution of the different velocity components and energy modes, as well as various accommodations coefficients (ACs) predicted by the GM and CLL boundary models. It was observed that the results from the GM model are in excellent agreement with MD results outperforming the CLL model.
In the next step, the capability of the GM approach to construct a generalized GSI model was examined. Two parallel walls system including monoatomic and diatomic gases was used as a benchmark system here. Considering the surface temperature as one of the most important parameters affecting the behavior of gas molecules after reflecting from the surface, initially, MD simulations at different wall temperatures were carried out. The resulting collisional data obtained from different simulations were gathered together and used for the training of the GM model as a whole. The comparison of the predicted results for different temperatures by the GM model with the original MD results and the predicted results by the CLL model confirmed the superiority of the GSI model derived from the GM approach.
Finally, an efficient hybrid algorithm was proposed by coupling the GM derived GSI model with DSMC method (GM+DSMC). As an application, the proposed model was used for modeling heat transfer in 1D Fourier thermal problem in the case of Ar-Au system. The comparison between the obtained results from the proposed model and the pure DSMC simulation with the CLL boundary model showed a better performance of GM+DSMC model.

1

INTRODUCTION

1

1.1. BACKGROUND AND MOTIVATION

Rarefied gas dynamics is considered one of the most important research subjects of this century. This fact stems from the practical application of this topic in numerous industrial and environmental areas such as semiconductor industry, micro/nanofluidics, aerospace engineering, nuclear reactor safety, heterogeneous catalysis, and seawater desalination [1–3]. To achieve an optimal design a fundamental understanding of the physics of gas flow in the aforementioned systems is crucial. Usually, at design and fabrication stages of these devices, various numerical approaches are employed to predict gas flow transport properties. These numerical tools are chosen based on the degree of rarefaction in the system and the computational power that can be afforded.

As depicted in Figure 1.1, the degree of rarefaction in a gas is characterized by the Knudsen number (Kn), which is the ratio of the molecular mean free path (λ) to a flow characteristic length scale (L). The mean free path is defined as the average distance a molecule travels between consecutive collisions. For Kn < 0.001, gas is considered as a continuum medium, and Navier-Stokes-Fourier (NSF) equations with no-slip boundary conditions are used to determine flow field properties [4]. By going to a higher rarefaction regime (0.001 < Kn < 0.1), the non-continuum effects, such as velocity slip or temperature jump, appear in a thin gas layer adjacent to the solid surface, known as the Knudsen layer. In this regime, NSF equations are still valid. However, to include the experienced non-continuum effects, it is necessary to introduce some modifications to the boundary conditions. For 0.1 < Kn < 10 gas is in the transitional regime. In this regime, due to stronger non-continuum phenomena, NSF equations are no longer applicable to describe flow properties, and the Boltzmann Equation (BE) or methods based on the kinetic theory should be applied [5]. The BE describes rarefied gas behavior, assuming that gas molecules move and collide with each other constantly and randomly at the microscopic level. As Kn increases, the collisions of the gas molecules with the adjacent surfaces become more dominant in determining flow properties than intermolecular collisions. That is the main reason we need an approach to account for the behavior of individual molecules with their surrounding molecules. When Kn > 10, gas is in the free molecular regime. In this case, the collisions between gas molecules are very rare. As a result, it is very common to completely neglect the gas-gas collisions in this flow regime [4].

Based on the applied governing equations in the system, different numerical schemes are used to solve the gas flow problem [6]. For instance, in the case of a continuum system based on the NSF equations, initially, the partial differential equations are discretized into a set of algebraic equations using numerical techniques such as Finite Ele-

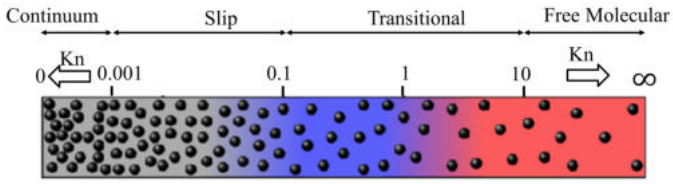

Figure 1.1: Gas flow regimes characterized by Knudsen number

ment Method [7]. The driven set of equations can then be solved using numerical algorithms such as SIMPLE [8].
Substituting the NSF equations with BE to describe gas flow properties induces the demand for other numerical schemes. Method of Moments (MoM) [9] is a commonly used numerical approach to solve the integro-differential form of BE. In this approach simplified models are applied to characterize the collisions between particles. Besides, the relations between particle velocity distributions and macroscopic properties of the system, such as mass density or flow velocity, are used to investigate gas transport properties. Solving the resulting partial differential equations becomes more complex by going toward a higher degree of rarefaction. Therefore, these approaches are usually used in the medium rarefied regime ($0.01 < Kn < 1$).
Another class of numerical techniques that are based on BE are particle-based simulation techniques, such as Lattice Boltzmann Method [10] (LBM), Direct Simulation Monte Carlo [11] (DSMC), and Molecular Dynamics (MD) simulation [12]. In LBM and DSMC, computational particles are typically employed to represent a group of real molecules. While particles streaming and collisions are treated in separate steps in both methods, different strategies are applied to deal with these processes [13]. For example, during the streaming step in LBM, particles at each lattice site can only move along a finite number of directions with certain prescribed velocities. On the other hand, there is no such constraint applied in DSMC. Such fundamental differences make DSMC a more adoptable simulation tool that can be utilized in a broad range of rarefaction regime ($Kn > 0.1$). However, LBM is more efficient when dealing with near continuum flows ($Kn < 0.1$). MD is another particle-based simulation technique that makes it possible to explicitly study the interaction between individual particles in a system. Since the exact movement of all the particles is modeled in the MD simulations, these simulations are relatively very expensive. Thus, employing pure MD simulations for macroscopic engineering problems is not realistic. Nevertheless, MD can be used to achieve a detailed understanding of transport phenomena happening at the microscale level in a rarefied gas system.
While the aforementioned numerical methods differ in various technical aspects, the prescription of a rigorous boundary model is a key parameter to achieving reliable results in all of them. In fact, as the Knudsen number increases in a system, gas-surface interactions become relatively more influential than gas-gas interactions in determining flow field properties. Due to the very complex nature of gas-surface interactions, it is exceedingly difficult to develop an analytically tractable and general model to describe such interactions. To circumvent such complexity, many researchers combine their theoretical understanding of gas-surface interactions with either experimental measurements or MD simulations and propose boundary models with different levels of sophistication that can be used in various rarefied gas flow systems [14–35].
Following the probabilistic nature of the BE, all the gas-surface interaction (GSI) models have been presented in the probabilistic form, indicating the relation between the incidence state and the reflection state of gas molecules on a solid surface. The fully diffuse and specular reflection models are the first and most fundamental descriptions of gas-surface interactions established by Maxwell in 1859 [14]. The fully diffuse GSI model assumes that the surface condition fully determines the state of reflected gas molecules, and the re-emitted gas molecules totally forget the information related to

the incident state. On the other hand, the specular reflection model assumes that the incident gas molecules experience a fully elastic collision with the surface, i.e., their tangential momentum maintains, and the normal momentum only changes the direction. Later, experimental studies revealed that pure diffuse or specular reflections could not properly describe the real gas scattering process. Maxwell proposed a new model combining these two GSI models [15]. According to this model, a fixed fraction, α, of the gas molecules is reflected in a completely diffusive fashion, and the remainder, 1-α, is reflected in a specular fashion. Here, α refers to the accommodation coefficient (AC), which generally is a phenomenological parameter used to quantify either momentum or energy exchange between a gas and its neighboring surface. In the case of the Maxwell model, the tangential momentum AC was applied in the model. Despite the successful implementation of the Maxwell model in various rarefied gas systems, many works demonstrated that this model gives inaccurate results in some engineering applications such as Fourier thermal problem, thermal transpiration, and evaluation of Knudsen force [36, 37].

To overcome the shortcomings of the Maxwell model, Cercignani and Lampis proposed a more elaborated GSI model, in which, in addition to the tangential momentum AC the normal energy AC was also incorporated into the model [16]. This consideration makes the Cercignani-Lampis (CL) scattering model able to describe the momentum and energy transport simultaneously, while the Maxwell model does not offer this possibility. Later on, the CL scattering model was extended by Lord to describe also the diatomic gas molecules scattering process [17]. The Maxwell and Cercignani-Lampis-Lord (CLL) scattering models, as the most commonly used boundary models in the technological rarefied gas systems, are the backbones of an important category of scattering models known as the classical or phenomenological scattering models [18–25]. Generally, in these models several ACs are applied in the model to describe the scattering process in a specific application. Measuring ACs is the major bottleneck affecting the performance of the phenomenological GSI models.

The researchers have devoted a significant amount of work to evaluating various ACs utilizing several different techniques [38–40]. The problem is that the superposition of numerous parameters can affect the gas-surface interactions. For instance, gas properties (i.e., gas temperature, purity, and molecular weight) and surface condition (i.e., surface roughness, cleanness, temperature, and chemistry) are the parameters that need to be considered while measuring ACs. However, controlling all these parameters simultaneously with high precision at the microscale level is quite challenging. Therefore, sometimes major discrepancies in the values of the reported ACs for the same gas-surface pair can be found between different experimental studies [38]. This observation can be caused by the differences in the applied experimental approach, sample preparation, and the theoretical model for deriving the ACs.

Although the experimental results elucidate general trends of ACs with various system properties, the MD simulation can deliver far more insight into the fundamental physics of gas-surface interactions. MD can be used to measure ACs in certain rarefied regimes that are very complicated to reconstruct experimentally. The accuracy of the MD results highly depends on the supplied intermolecular force fields. Quantum Mechanics (QM) computations are usually combined with some experimental measurements to derive

the required force fields [12]. However, this task for some gas-solid pairs or under certain extreme physical conditions (e.g., very high or low temperature) can be significantly challenging. Regardless of the all addressed difficulties in computing ACs, it has been shown that the empirical GSI models are generally not adequate to fully capture the underlying physics in systems encountered with highly nonequilibrium and complex gas flow conditions. Under such circumstances, significant non-continuum phenomena are likely to happen that can not be fully described by a model based on a limited number of constant parameters that need to be determined a priori [41].

Another category of GSI models is the physical-based GSI models. The hard-cube model [26] is the first model in this category proposed by Logan and Stickney. In this model, the solid surface is demonstrated by an ensemble of cubes. An impulsive force of repulsion is employed to describe the interaction of a gas molecule with a surface atom. The gas-surface interactions are simplified as one-dimensional binary collisions between the gas molecules and effective surface cubes. The soft-cube model [27] is another physical-based scattering model. The fundamental difference between these two cube models is that in the second one, a stationary square-well attractive potential and an exponential repulsive potential are applied to describe gas-surface interactions. Therefore, the soft-cube model can describe the trapping-desorption phenomenon. The washboard [28] model, as the extension of previously mentioned physical models, is the most realistic one, which by incorporating corrugation parameters, can be used to study the gas flow passing a rough surface. While these models are qualitatively consistent with the scattering phenomena observed in molecular beam experiments, they do not satisfy the reciprocity principle [4], which is the basis of constructing a physically realistic GSI model. To overcome this problem, Liang et al. [29] coupled the original washboard model with the CLL kernel and proposed a new model based on three input parameters: the surface corrugation strength, the attractive potential well depth, and local collision AC. However, their model can not describe the interaction of polyatomic gas molecules, including rotational or vibrational energy modes, with solid surfaces.

Another category of GSI models is nonparametric models [31, 32, 42]. Unlike the aforementioned scattering models, the nonparametric models are not based on intermediate parameters, such as ACs, that are required to be computed in advance. Besides, they do not need any predefined analytical form. These models are constructed directly based on the MD collisional data. M. Liao et al. [31], using the molecular beam MD setup, investigated the interactions between CO_2 and CH_4 gases with Graphite surface. A conditional probability distribution function was introduced based on the relation between the incoming and outgoing velocity components presented in the MD collisional data. This function can generate postcollisional velocities, given the precollisional velocities. A similar procedure was also followed in the works by N. Andric et al. [32] and Liu et al. [42] while they studied different gas-solid pairs. The nonparametric models are very flexible and considered a promising tool for constructing a generalized boundary model for highly nonequilibrium conditions. Nevertheless, in these scattering models, solid surface is considered perfect and clean, and the gas-gas interactions that can affect the reflected gas molecules properties in the early transition regime (0.1<Kn<1) are ignored. Besides, the usual non-parametric approximations of high-dimensional multivariate data can be a complex task requiring advanced learning methods [43]. To avoid

this problem, before the main fitting step, usually various techniques are employed to reduce the dimensionality of the dataset. [32].

Machine learning is another approach that has been recently employed to construct a GSI model based on MD data [33, 35, 44]. Gaussian Mixture (GM) [45] approach was used by Liao et al. [33] to construct a scattering kernel for monoatomic gases. The GM approach is a robust unsupervised machine learning technique that has been utilized in a broad range of scientific fields, such as pattern recognition, language identification, and object tracking of multiple objects. The GM scattering model is a parametric model, but unlike the common parametric models (e.g., the CLL model), its performance does not depend on a finite number of parameters. On the contrary, due to the presence of numerous fitting parameters in the model, it retains the flexibility of a nonparametric model. The main advantage of the GM approach over nonparametric schemes for deriving scattering kernel is that based on this approach, it is more straightforward to deal with high-dimensional multivariate data. In addition, the whole collision data obtained from the MD simulation is used to train the GM model. Therefore, this model can capture all the underlying physical phenomena happening at the gas-surface interface. Investigating both smooth and rough atomic surfaces, Liao et al. [33] reported good performance of the GM-driven scattering model in predicting the physical and statistical properties of the gas-wall interfaces. However, since their model is also based on the molecular beam MD setup, it cannot deal with absorption-related problems. Since in their MD simulations a clean and perfect solid surface is assumed.

Despite devoting a considerable affort to investigate the physics behind the GSI phenomenon, there has still been a lack of a generalized model that can be used to describe such a complex phenomenon in different rarefaction regimes. The main objective of this thesis is to construct a generalized scattering model based on the MD simulation data that can be used in higher-scale simulation approaches, such as DSMC, to study the flow field properties of rarefied mono/diatomic gases in highly non-equilibrium situations. To achieve this goal, the GM approach is employed in this work. While in the existing GSI models, only gas-wall interactions are included in the scattering model, in this work, both gas-gas and gas-wall interactions are employed to construct the scattering model. This issue guarantees the capability of the model to deal with adsorption-related problems in the case of systems in which a considerable number of gas molecules are adsorbed on the solid surface. As an example, the surface coverage for N_2-Ni and H_2-Ni systems are 3.8 and 0.35 $1/nm^3$, respectively. As will be discussed in more detail in Chapter 4, such a higher value of surface coverage in the case of N_2-Ni system considerably affects the scattering behavior of gas molecules.

1.2. SIMULATION METHODS FOR RAREFIED GASES

In this work, MD simulation is used to derive a fundamental understanding of physics behind gas-surface interactions and generate the initial data required to construct a gas scattering model. At the end, the developed scattering models are implemented in a DSMC solver to study their compatibility and performance in a coarse-grained simulation technique. Some basics of MD and DSMC simulations are presented in the following subsections.

1.2.1. MOLECULAR DYNAMICS

MD is a computational tool originally developed in the late 1950s and has been extensively used in chemical physics, materials science, and biophysics. The main idea is to mimic what atoms do in real-life systems, presuming a given potential energy function. In such a way, MD helps us to interpret experiments, provide alternative interpretations, as well as to assess theories of intra- and intermolecular interaction. In the conventional implementation of MD, given the initial coordinates of the atoms in a system, one can compute the force exerted on each atom by all the other atoms in the system. For the atom *i*, this is done by differentiation of the global potential energy with respect to the position of atom *i*:

$$F_i = -\nabla_{r_i} V \tag{1.1}$$

where V denotes potential energy and contains information regarding interatomic interactions as:

$$V = \sum_i V_1(\boldsymbol{r}_i) + \sum_i \sum_{j>i} V_2(\boldsymbol{r}_i, \boldsymbol{r}_j) + \sum_i \sum_{j>i} \sum_{k>j} V_3(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k) + \dots \tag{1.2}$$

where V_1 refers to the effect of an external field (e.g., the boundaries or the gravitational field). The remaining terms represent particles interactions in the system. V_2 and V_3 represents two-body and three-body potentials, respectively. For gases and liquids, the interaction potential is mainly dominated by pair potential term V_2. Usually, in MD simulations, the pair potential is replaced by an effective pair potential V_2^{eff} that contains also the three-body effects [12]:

$$V = \sum_i V_1(\boldsymbol{r}_i) + \sum_i \sum_{j>i} V_2^{eff}(\boldsymbol{r}_i, \boldsymbol{r}_j) \tag{1.3}$$

The set of parameters required to describe a potential energy function is derived from quantum mechanics calculations and, typically, from certain experimental measurements.

In classic MD, for a system of N atoms interacting via a potential field V as in Equation 1.2, the spatial position of each atom as a function of time is calculated by numerically solving Newton's equations of motion:

$$m_i \frac{\partial^2 r_i}{dt^2} = F_i \tag{1.4}$$

where F_i is the force experienced by atom *i*, m_i is the atomic mass of atom *i*, and r_i is its position. Velocity Verlet's algorithm is the most common and stable time integrator for solving the differential equation of motion in an MD simulation. In this approach, the position and velocity of an atom at consecutive time steps can be computed as [12]:

$$r_i(t+\Delta t) = r_i(t) + \Delta t\, v_i(t) + \frac{1}{2}\Delta t^2 \frac{F_i(t)}{m_i} \tag{1.5}$$

$$v_i(t+\Delta t) = v_i(t) + \frac{1}{2}\Delta t \left(\frac{F_i(t)}{m_i} + \frac{F_i(t+\Delta t)}{m_i} \right) \tag{1.6}$$

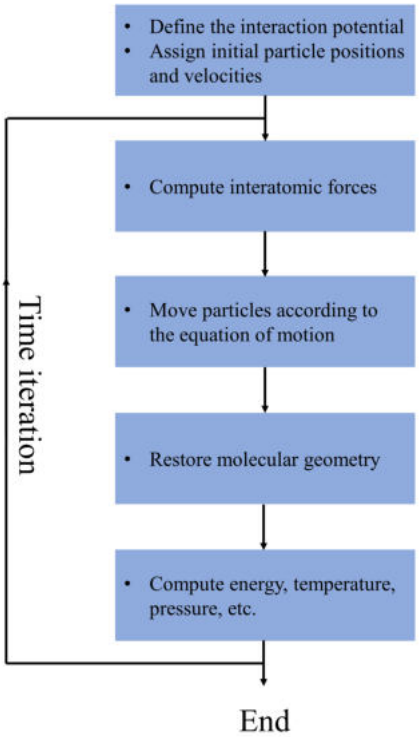

Figure 1.2: Flowchart of an MD simulation

where Δt is the time step used in the simulation, typically in the order of femtosecond. The flowchart of an MD simulation is shown in Figure 1.2.

An MD simulation generates for every time step positions and velocities of all particles in the system at the microscopic scale. A collection of such information is called a microstate. On the other hand, a collection of all microstates representing the same macrostate or subjected to the same macroscopic constrains is called an ensemble [46]. Different ensembles can be considered in an MD simulation. The most common one is the microcanonical (NVE) ensemble. In this ensemble the number of particles, N, the volume of the system, V, and the total energy of the system, E, are constant. Translation of the microscopic level information into the desired macroscopic properties (e.g., energy, temperature, pressure, etc.) is carried out by statistical mechanics. As an example, assuming a system of N monoatomic particles, the overall temperature, T, of the system can be computed as:

$$KE = \frac{3}{2} N k_B T \tag{1.7}$$

where KE is the total kinetic energy of the system and k_B is the Boltzmann constant. KE is directly related to the velocity of atoms as follows:

$$KE = \frac{\sum_{i=1}^{N} m_i v_i^2}{2} \tag{1.8}$$

To study the behavior of a system at a specific temperature, an MD simulation is carried out based on the canonical (NVT) ensemble. In this ensemble the number of particles, N, the volume of the system, V, and the temperature of the system, T, are constant. In an NVT ensemble the total energy is not specified. However, the temperature is controlled by coupling the system to a so called thermostat. A thermostat, acting as an external heat bath, allows energy to enter and leave the simulated system to maintain a constant temperature. Practically, using thermostats the overall temperature in the system is controlled by modifying velocities of subsets of particles. The most commonly used thermostats in MD simulations are the Berendsen [47] and Nose-Hoover [48] thermostats. As an example, using the Berendsen thermostat, to achieve the desired temperature, T_0, the velocities are corrected at each time step, such that the rate of change of temperature is proportional to the difference in temperature:

$$\frac{dT(t)}{dt} = \frac{1}{\tau}(T_0 - T(t)) \tag{1.9}$$

where $T(t)$ is the system temperature at time t, and τ is the damping factor, which determines how tightly the bath and the system are coupled together.
The MD method is a valuable computational tool for several reasons. First of all, it makes it rather straightforward to capture the exact trajectory of atoms in the system, which is a very challenging task with any empirical technique. Secondly, the simulation condition is exactly known and can be controlled with very high precision. While very accurate, MD simulations are computationally very expensive. The reason is that a substantial amount of calculations must be performed in each time step. Based on the model characteristics, such as the level of complexity of the potential energy function, 100 to 1000,000's atoms can be studied in an MD simulation for the overall simulation time of a maximum 1 μs. Despite this, in the last decade, the availability of parallel computing and using graphics processing units (GPUs) for running MD simulations have been a great help for researchers to speed up MD computations.
As it has been addressed in Section 1.1, computing different ACs is one the main approaches that MD simulations can help to investigate the rarefied gases scattering process. Based on MD simulation results, there are two methods for computing ACs. In the first method, referred to as the classical (Cla) approach in this work, different ACs are computed as:

$$\alpha_q = \frac{\langle Q_I \rangle - \langle Q_R \rangle}{\langle Q_I \rangle - \langle Q_w \rangle} \tag{1.10}$$

where q can be any kinematic quantity, such as gas translational velocity in a certain direction, as well as its total translational or rotational kinetic energy content. $\langle Q_I \rangle$ is the impinging flux of quantity Q, and $\langle Q_R \rangle$ denotes the outgoing flux of quantity Q. $\langle Q_w \rangle$ stands for the outgoing flux associated with the fully accommodated condition.
Another method for computing ACs based on the MD results is the Correlation (Cor) method, developed by P. Spijker et al. [49]. In this method different ACs can be calculated

based on the least-squares approximation on the MD results containing impinging Q_I and outgoing Q_R quantities as follows:

$$\alpha_q = 1 - \frac{\sum_i (Q_I^i - \langle Q_I \rangle)(Q_R^i - \langle Q_R \rangle)}{\sum_i (Q_I^i - \langle Q_I \rangle)^2}. \tag{1.11}$$

The main difference between these two methods is that in the case of the Classic method, as the difference between the impinging gas and surface temperatures goes toward zero, numerical instability arises. On the other hand, this limitation is not encountered in the Correlation method, and basically, this method can be used even for the situation, where gas and surface have the same temperature.

1.2.2. DIRECT SIMULATION MONTE CARLO

DSMC is the most popular particle-based simulation technique in rarefied gas flow simulation that has been utilized in a broad range of engineering applications. The main reason behind the popularity of DSMC is that despite the substantially lower computational cost of DSMC compared to MD, it provides acceptable precision in measuring rarefied gas transport properties in relatively larger domains (in the order of μm to m). In fact, in DSMC the total computation needed is proportional to the number of molecules simulated N, in contrast to N^2 for MD simulation [50]. Unless a truncated potential and a nearest neighbor list is used. However, common DSMC solvers cannot model a phase transition in the fluid, while MD can model this phenomenon. Besides, going towards denser systems (e.g., slip flow regime), the accuracy of DSMC reduces and the LBM approach will be a better option in this case. DSMC, introduced by G.A. Bird for the first time, can be considered as a Monte Carlo approximation of the time-dependent nonlinear Boltzmann equation [51]. The Boltzmann equation illustrates the evolution of a rarefied gas at the level of the single particle distribution function as:

$$\frac{\partial f}{\partial t} + v. \bigtriangledown f = \Omega(f, f) \tag{1.12}$$

where f represents the probability density function of particles and v is the particle velocity. $\Omega(f, f)$ is the collision integral, which describes the collisions between particles within a cell.

In DSMC, many independent simulating particles are employed to model gaseous flows, where each DSMC particle is representative of a large number of real gas molecules ($10^4 - 10^8$ real molecules). The concept of representative particles is the key feature that allows us to rescale length and time to model larger systems. In the first step, the simulation domain is randomly populated with particles. The initial velocities of the particles are generated based on the Boltzmann distribution at the desired temperature. In the case of a polyatomic gas molecule, its initial internal energy (E_{int}), which can be either rotational or vibrational energy, is assigned based on the equipartition theorem at the desired temperature.

The DSMC particles are allowed to move and collide within the simulation domain; however, the main approximation of DSMC is the uncoupling of the particle motions and the intermolecular collisions over small time intervals that as smaller than the mean collision time. While the particles' motion is treated deterministically, collisions between

them are considered stochastically. In order to implement DSMC, the simulation domain is decomposed into cells with dimensions less than the gas mean free path. In each time step (τ), the DSMC particles are initially moved as if they do not interact with each other. Thus, different from MD, DSMC particles can move through each other. Every particle's position is reset as:

$$\boldsymbol{r}_i(t+\tau) = \boldsymbol{r}_i(t) + \boldsymbol{v}_i(t)\tau \tag{1.13}$$

Particles that reach an interface are treated according to the boundary model assigned to that specific interface. After moving all the particles, they are sorted into spatial cells, and only particles within the same computational cell can collide. In each cell, a set of representative collisions is processed in each time step. The No-Time-Counter (NTC) approach is the most common collision-sampling approach implemented in most DSMC solvers. In this approach, the number of candidate collisions pairs (N_{cand}) from each cell is determined as:

$$N_{cand} = \frac{N_c^2 F_N (\sigma_T c_r)_{max}}{2V_c} \tag{1.14}$$

where N_c is the number of simulated particles in the cell, F_N is the number of the real molecules represented by the simulated one, $(\sigma_T c_r)_{max}$ is the maximum value of the product of total collision cross-section and relative velocities, and V_c is the volume of the cell. Different molecular models are developed to describe the collision cross-section. The main idea behind these models is to reproduce the required temperature dependence of the transport properties (e.g., viscosity and self diffusion) in a real gas. Variable hard sphere (VHS) is currently the most widely used molecular collision model in DSMC because of its simplicity and compelling approximation of real intermolecular potential [52]. While collisions between monoatomic particles are considered elastic in a DSMC simulation, the collisions can be inelastic for polyatomic particles. Typically, for polyatomic gas molecules phenomenological models, such as the Borgnakke-Larsen (BL) model [53], are employed to describe the energy exchange between the internal and translational energy modes.

The NTC approach applies an acceptance-rejection technique to choose the collision partners. Here, a randomly selected pair of particles will collide if:

$$R < \frac{\sigma_T c_r}{(\sigma_T c_r)_{max}} \tag{1.15}$$

where R is a uniformly distributed random number. If the pair is accepted, the postcollisional velocity of particles is determined based on the conservation of linear momentum and energy laws. This algorithm is continued until N_{cand} pairs are investigated. After resetting the velocity of all colliding particles, the process is repeated for the next time step. The macroscopic flow properties can then be evaluated by time averaging the cell-based values. The flowchart of a DSMC simulation is shown in Figure 1.3.

1

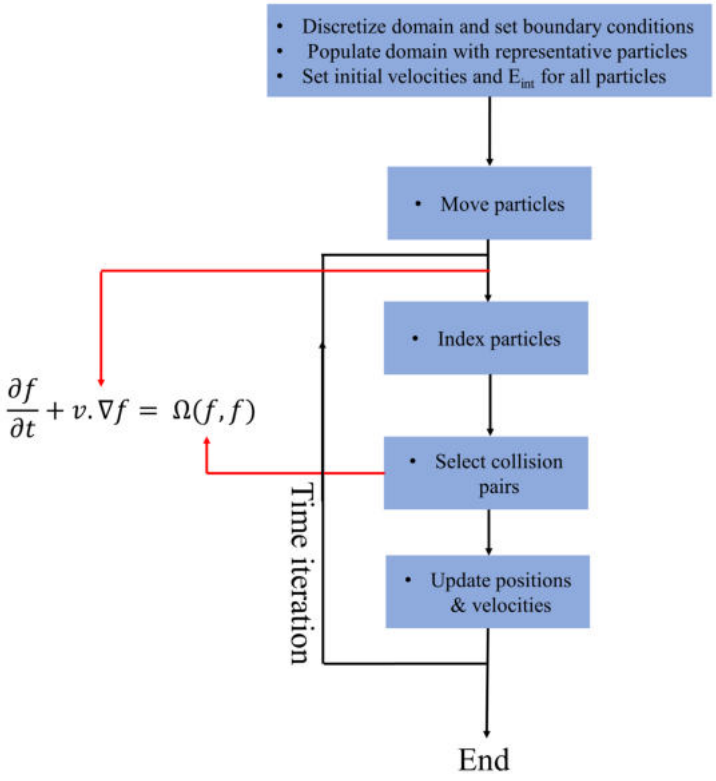

Figure 1.3: Flowchart of an DSMC simulation

1.3. THESIS OUTLINE

This thesis is organized as follows. Chapter 2 mainly focuses on how applying different interaction potentials between gas and solid can affect the MD results for ACs. The MD simulation setup is a two parallel walls system construed from Au atoms and interacting with either Ar or He atoms.In Chapter 3, the GM model is employed to construct a scattering model for monoatomic gases (Ar, He) interacting with a Au surface. Herein, considering that both gas-gas and gas-wall interactions can affect the gas flow solution in the vicinity of a solid wall, a two-parallel wall system was used as the MD setup. Benchmarked by MD results, the performance of the GM scattering model is examined against the CLL model in different physical problems, such as the Fourier thermal problem, the Couette flow problem, and a combined Fourier-Couette flow problem.In Chapter 4, the GM approach is used to construct a stochastic GSI model for diatomic gas molecules (H_2, N_2) interacting with the Ni surface. A two parallel walls setup is used in MD simulation to gather collisional data. Here, the entire translational and rotational velocity components of the gas molecules before and after colliding with the surface are utilized for training the GM model. This creates the possibility of accurately capturing the energy exchange between the different molecular modes that the classical scattering kernels cannot capture. Similar to Chapter 3, the performance of the GM scattering model is evaluated against the CLL model in different benchmarking systems.
In Chapter 5, the capability of the GM approach to construct a generalized GSI model for monoatomic and diatomic gas molecules is investigated. The thing is, in Chapters 3 and 4, for each case study, the corresponding GM model was trained separately. In this chapter, initially, it is shown that the surface temperature is a crucial parameter affecting the scattering process in the two parallel walls Ar-Au and H_2-Ni systems. In the next step, MD simulations at different wall temperatures are carried out for both case studies. For each gas-surface pair, the collisional data at different wall temperatures are gathered in one data set and used to train the GM model at once. Using the original MD results as the reference solutions, it is shown that the GM scattering model can predict the post-collisional behavior of gas molecules at different wall temperatures with good precision. In chapter 6, the GM scattering kernels driven for the Ar-Au and H_2-Ni systems for Fourier thermal problem are coupled to a commercial DSMC solver.
Finally, in Chapter 7, the main finding of this thesis are summarized, and recommendations for further research are addressed.

2

IMPACT OF INTERMOLECULAR INTERACTIONS ON THE ACCOMMODATION COEFFICIENTS

Molecular dynamics (MD) simulations are conducted to determine energy and momentum accommodation coefficients at the interface between rarefied gas and solid walls. The MD simulation setup consists of two parallel walls, and of inert gas confined between them. Different mixing rules, as well as existing ab-initio computations combined with interatomic Lennard-Jones potentials were employed in MD simulations to investigate the corresponding effects of gas-surface interaction strength on accommodation coefficients for Argon and Helium gases on a gold surface. Comparing the obtained MD results for accommodation coefficients with empirical and numerical values in the literature revealed that the interaction potential based on ab-initio calculations is the most reliable one for computing accommodation coefficients. Finally, it is shown that gas–gas interactions in the two parallel walls approach led to an enhancement in computed accommodation coefficients compared to the molecular beam approach. The values for the two parallel walls approach are also closer to the experimental values.

This chapter is published as a peer-reviewed article. Mohammad Nejad, S., Nedea, S., Frijns, A. and Smeulders, D. The influence of gas–wall and gas–gas interactions on the accommodation coefficients for rarefied gases: A molecular dynamics study. *Micromachines*, 11(3), p.319 (2020).

2.1. INTRODUCTION

Rarefied gas condition is encountered in a broad range of modern engineering applications; for example, in low pressure devices such as semiconductor manufacturing and spacecraft flying at high altitudes, as well as small-scale structures such as microelectronic devices and micro/nanoelectromechanical systems (M/NEMS) [1]. In all these applications to achieve the effective thermal management, a fundamental understanding of gas–surface interactions (GSI) is of paramount importance. The degree of rarefaction of a gas is quantified by Knudsen number ($Kn = \frac{\lambda}{L}$), where λ is the mean free path of the gas molecule and L is the characteristic length scale. It is known that a gas is regarded as rarefied if Kn > 0.01 . Over the years, a wide range of experimental [38, 54, 55] and numerical studies [56–58] have been carried out, investigating rarefied gas-solid surface interactions. Due to the complexities involved in the instrumentation, an empirical study of GSI is a very challenging and time-consuming task. Regarding computational techniques, due to the noncontinuum gas behavior adjacent to the solid surface, common continuum approaches (Navier–Stokes equations) are not applicable to describe energy and momentum exchange at GSI. Herein, particle-based simulations methods such as MD simulations [12] are considered a promising candidate to study GSI.

MD simulations can provide an atomistic-level understanding of the scattering dynamics of the gas molecules interacting with solid surfaces. Energy and momentum accommodation coefficients (E/MACs), which are the most relevant parameters involved in GSI models, describe the degree at which a gas attains its thermal or kinematic equilibrium with a surface while interacting with it. MD simulation is a very promising tool to determine different accommodation coefficients. These coefficients can be fed into semi-empirical GSI models such as Maxwell's model [59] or Cercignani–Lampis–Lord (CLL) model [60] that can be employed as boundary conditions for higher-scale simulation techniques such as Direct Simulation Monte Carlo (DSMC) [11], Lattice Boltzmann method (LBM) [61], and method of moments (MoM) [62] to describe heat and mass flow at macroscopic level under rarefied condition. In literature there are various numerical studies in which MD simulations are employed to determine accommodation coefficients for different gas-solid surface combinations [63–70]. The general objective of all such investigations is to find the correlation between the energy and momentum accommodation coefficients and input parameters such as the gas temperature or purity, gas molecular weight (MW), surface condition (i.e., surface roughness, cleanness, temperature and chemistry), as well as the gas–surface interaction strength. Briefly summarized, results in the literature reveal that E/MACs decrease by increasing the temperature and implicitly the kinetic energy of the molecules. Moreover, increasing the surface roughness, gas molecular mass, and gas–solid interaction strength lead to an increase in E/MACs. Gas molecules approaching a surface are sometimes trapped by the potential well and stager on the surface for a while as they are physically adsorbed. The gas molecules may escape the potential well after some residence time through which they lose sufficient amount of their thermal or kinetic energy such that they accommodate to the surface temperature at a higher degree (i.e., resulting in a higher accommodation coefficient). This phenomenon called trapping-desorption is more likely to happen at lower temperature, higher surface roughness and, higher gas MW, as well as stronger gas-solid interaction which at the end causes higher E/MACs at aforementioned condi-

tions.

Due to the superposition of many factors affecting the phenomenon of gas-surface interaction, sometimes in literature for the same pair of gas–solid surface, a considerable discrepancy in the values of accommodation coefficients, obtained by different MD simulation approaches is found. As an example, for the Platinum–Argon combination the obtained values for the tangential momentum accommodation coefficient by Yamamoto et al. [71] and Hyakutake et al. [72] were 0.19 and 0.89, respectively. Accurate values of accommodation coefficients are essential for the better assessment of the overall transport properties of rarefied gases. Therefore, we will compare most common MD approaches and study the effect on the resulting thermal and momentum accommodation coefficients, and compare the results with experimental values.

In most previous MD simulations [67, 69, 70] to compute E/MACs, the molecular beam approach was employed, in which the gas molecule adjacent to the surface is assumed to interact only with the wall atoms, and its initial velocity is sampled from an equilibrium distribution at certain temperature. Such assumptions are valid for a highly-rarefied gases (i.e., Kn > 10). However, the condition is very different in most M/NEMS applications, where less degree of rarefaction is encountered (Kn < 1) [1]. In such systems, in the case of a temperature difference between gas and surface, gas molecules will experience non-equilibrium processes. Furthermore, in this rarefaction regime, both of gas–wall and gas–gas interactions can equally impact the ACs values [73, 74].

In the present work, two parallel plates MDapproach is applied to calculate E/MACs of noble gases (Ar and He) interacting with Gold (Au) surface. Firstly, the dependence of E/MACs on the gas pressure in the system was investigated, which to our best knowledge has not been studied by MD simulations, previously. The impact of gas-solid interatomic potential on E/MACs was also characterized. To do so, a pairwise Lennard-Jones (LJ) potential was considered at the solid-gas interface. The LJ potential parameters were computed using different approximation methods such as Lorentz–Berthelot (LB) and Fender–Halsey (FH) mixing rules, as well as taking form existing ab-initio calculations. It has been observed that an interaction potential based on quantum calculations such as ab-initio computations is the most reliable one for computing different ACs. Such a behavior has been reported by Mane et al. [69] and Daun et al. [75], where using molecular beam approach, they studied the interaction between monoatomic gases with aluminum and iron surfaces, respectively. At the end, to unravel the importance of including gas–gas interactions on the MD obtained accommodations coefficients, we compare our two parallel plates results with those obtained by the molecular beam approach.

2.2. MATERIALS AND METHODS

2.2.1. MOLECULAR DYNAMICS (MD) SIMULATIONS

The MD simulation setup considered in this work is a three-dimensional system, in which a monatomic gas is confined between two parallel walls (see Figure 2.1).

The cross-section area of the walls is 10 by 10 nm, and each of them consists of 8750

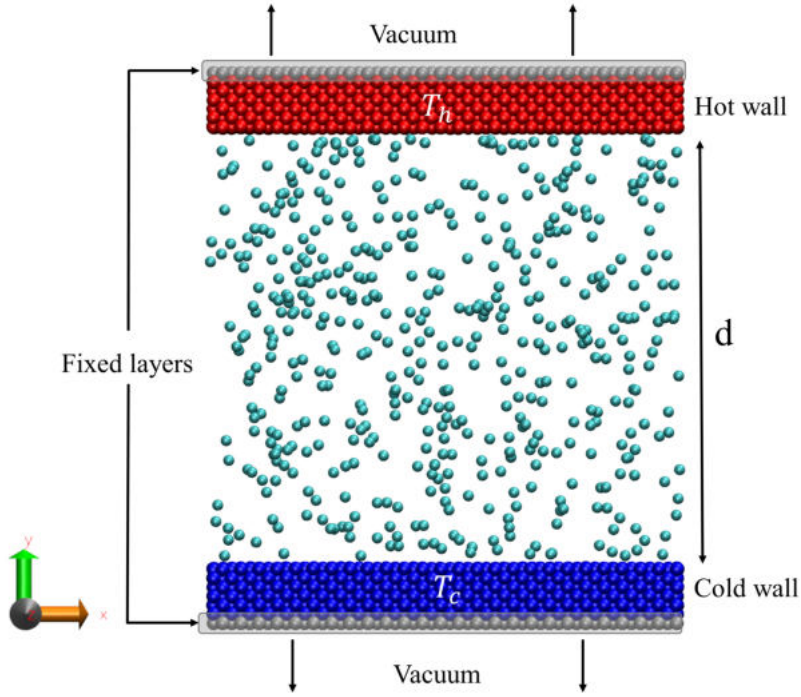

Figure 2.1: Schematic representation of the simulation setup; two walls kept at a distance d apart; they are thermostated at a low temperature T_c (bottom wall) and a high temperature T_h (top wall).

gold atoms arranged in a FCC structure. The walls are separated from each other in the *y* direction. The temperatures of the bottom and top walls are maintained at T_c= 300 K and T_h = 350 K, respectively, using Berendsen thermostats. In each wall, the outermost layer is fixed to prevent the wall from any translational or rotational motion. Periodic boundary conditions are applied in all three directions. Since the simulation box is also periodic in *y* direction, the walls are not placed directly at the periodic boundary. Vacuum has been considered between the walls at the periodic boundary in order to prevent direct contact and heat conduction between the hot and cold walls. At the beginning of the simulation, the temperature of the gas molecules is set at the 300 K, which corresponds to the root mean square velocity (V_{rms}) of 423 and 1367 m/s, for Ar and He gases, respectively. Herein, the gas molecules are not coupled to an external heat bath, and their temperature change is caused only via collisions with other particles (gas and solid particles). As it will be discussed in detail in the proceeding section, in order to obtain

reliable results for E/MACs, a large number of collisions between gas molecules and the wall surface are required: 100,000 collisions have been considered here, as it has been recommended in previous studies [64, 67] for a similar simulation setup. The distance between the two solid walls is fixed at d = 11 nm and d = 102 nm for Ar and He gases, respectively. For the Ar-Au system the Knudsen number and the reduced density (η) are Kn = 0.23 and η = 0.02, respectively. The reduced density is defined as ($\eta = \pi n a^3/6$, where n is the number density and a is the particle diameter [76]. On the other hand, for the Au-He system Kn = 0.56 and η = 0.00043. The interaction between Au atoms are described by the embedded atom model(EAM) potential [77], whereas the gas–gas and gas–solid interactions are modeled by the Lennard-Jones (LJ) potential. Herein, noble gas-Au pair potential coefficients have been calculated by commonly used (LB) mixing rule:

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2}, \epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \tag{2.1}$$

and (FH) mixing rule:

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2}, \epsilon_{ij} = \frac{2\epsilon_{ii}\epsilon_{jj}}{\epsilon_{ii} + \epsilon_{jj}} \tag{2.2}$$

as well as existing ab-initio pair potentials [70,78], derived from computational quantum mechanics.

The detailed gas-gas and gases-Au interatomic potential parameters used for mixing rules are presented in Table 2.1. In addition, the pair potential coefficients based on existing ab-initio calculations are listed in Table 2.2 The LJ cut-off distance (r_c) for gas-gas interactions is set at 2.5 times the LJ length parameter (σ_{ii}). In addition, for both of Au-Ar and Au-He pairs we used r_c = 12 Å which is similar to the cut-off radii that were used in the ab-initio simulations [70]. For implementing MD simulations, LAMMPS

Table 2.1: Gases-gold interaction potential parameters used by mixing rules and molecular weights

Atom Type	ϵ_{ii} (meV)	σ_{ii} (Å)	MW (a.m.u)
Au [79]	229.4	2.63	196.96
Ar [80]	12.2	3.35	39.94
He [80]	0.94	2.64	4.00

Table 2.2: Gases-gold interaction potential parameters based on ab-initio computation [70]

Parameter	Value
ϵ_{Au-Ar}	11.36 (meV)
σ_{Au-Ar}	3.819 (Å)
ϵ_{Au-He}	0.787 (meV)
σ_{Au-He}	4.342 (Å)

molecular dynamics package was used [81]. All simulation setups were initially equilibrated for 1 ns (time step = 1 fs) for Ar and 3 ns for He. Afterwards, the production run

was started and proceeded for the next 20 and 60 ns for Ar and He cases, respectively.

2.2.2. COMPUTING ACCOMMODATION COEFFICIENTS

To compute E/MACs using MD simulations, the trajectory of each gas molecule in the simulation box is monitored during the specified simulation time. Herein, the collision is defined in such a way that if a gas molecule approaching the surface crosses a virtual border placed in certain distance from the surface, its velocity components are recorded. For the same particle after re-emitting from the surface, when it reaches back to the virtual border, its velocity components are recorded again. Both scattered and trapped particles that are desorbed from the wall are considered here. As it is depicted in Figure 2.2, in this study the virtual border is located at one gas-wall interaction cut off distance (r_c = 12 Å) away from the wall surface in order to guarantee that the gas molecule is not affected by the wall potential.

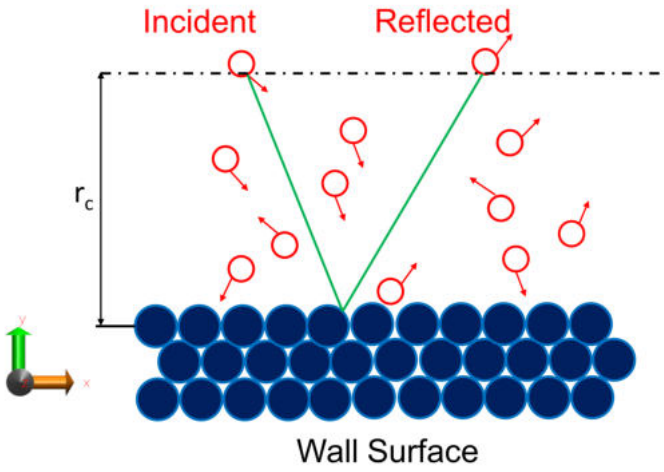

Figure 2.2: Schematic of gas-surface interaction used to compute accommodation coefficients in molecular dynamics (MD) simulations. Here the green line is a simplified example of gas molecule trajectory. The virtual border for the computation of the accommodation coefficients is placed at distance r_c = 12 Å from the solid surface.

The accommodation coefficients were calculated by the least-square method proposed by Spijker et al. [66]. In this approach, which is based on the correlation between input (Impinging) and output (Reflected) data obtained from MD simulations, the AC (α) can be computed as follows:

$$\alpha_q = 1 - \frac{\sum_i (Q_I^i - \langle Q_I \rangle)(Q_R^i - \langle Q_R \rangle)}{\sum_i (Q_I^i - \langle Q_I \rangle)^2} \tag{2.3}$$

where subscript q can be any kinematic quantity, such as the gas molecule velocity in a certain direction (for momentum accommodation) or its total kinetic energy (for thermal energy). Q_I and Q_R referred to the considered quantity for the impinging and reflected particles, respectively. To be more specific, when a gas atom approaching to the surface ($v_y < 0$) passes through the virtual border (dotted line in Figure 2.2) it will classified as an impinging particle. From the other side, when a gas atom going outwards with respect to the surface ($v_y > 0$) passes through the virtual border it will be considered as reflected particle. The bracket notations denote that the average value for these quantities is calculated.

2.3. RESULTS AND DISCUSSION

In order to compute E/MACs, after gathering collision data from MD simulation, the correlation between relevant impinging and outgoing velocities was studies. These correlations can be illustrated as a two-dimensional probability distribution profile, which for a particular impinging velocity gives the distribution of reflected velocities (see Figure 2.3). Herein, a line is fitted to the collision data based on least-square approximation (red line in Figure 2.3). Comparing the slope of this line, which is actually the fractional part of Equation 2.3, with dashed horizontal (fully diffusive) and diagonal (fully specular) lines gives us the accommodation coefficient.

First of all, the dependency of computed E/MACs on gas pressure between two walls was investigated. Herein, ab-initio pair potential was employed to describe gas-wall interactions. By decreasing the number of gas atoms in the simulation box, the pressure was reduced from 2.75 MPa to 0.42 MPa, and from 0.21 to 0.04 MPa in the case of Ar and He, respectively. As it is depicted in Table 2.3, for both gases, decreasing pressure does not have a significant impact on obtained accommodation coefficients, but considerably increases the MD simulation time required to record the same number of collisions. Similar pressure dependency has been also reported in an experimental study by Thomas and Brown [82]. Therefore, in the remaining part of this chapter we perform our simulations at the pressures of 2.75 and 0.21 MPa for Ar and He gases, respectively.

In the next step, the comparison between gas-wall interaction potentials obtained using different methods is shown in Figure 2.4. This figure shows that the FH mixing rule is relatively softer than the LB mixing, and that the ab-initio potential is softer than both mixing rules. Furthermore, it is depicted that the LB mixing rule highly overestimates the potential well depth. Such an overprediction in the case of a heavy gas like Ar causes that during MD run all Argon atoms are adsorbed on the solid surfaces (see Figure 2.5a),

Table 2.3: Variation of energy and momentum accommodation coefficients and MD simulations running time with the pressure in the simulation box for Au–Ar and Au–He pairs

System	Pressure (MPa)	Number Density ($1/nm^3$)	MFP(nm)	EAC	MAC	MD Simulations time (ns)*
Au-Ar	2.75	0.59	2.63	0.874	0.883	20
	1.27	0.27	5.71	0.832	0.846	50
	0.84	0.18	8.57	0.816	0.822	70
	0.42	0.09	17.14	0.783	0.791	100
Au-He	0.21	0.048	58.73	0.048	0.059	60
	0.13	0.029	97.89	0.046	0.057	90
	0.08	0.019	146.84	0.043	0.052	150
	0.04	0.009	293.70	0.042	0.049	250

* The time required to record 100,000 collisions

and that they do not leave the surface anymore. Therefore, for none of the impinging gas particles an outgoing velocity can be recorded, and E/MAC values are numerically unobtainable. In addition, the normalized number density distributions in the case of aforementioned system using varied interaction potentials are depicted in Figure 2.5b. Herein, initially it can be understood that the gas density adjacent to the wall surfaces is higher than the bulk density (n_0) in the central part of the system: stronger is the gas-wall interaction the higher is the density profile peak near to the wall. This is in agreement with the behavior reported in [83]. Furthermore, in the case of the LB potential, except in the vicinity of the walls, the gas density goes to zero.

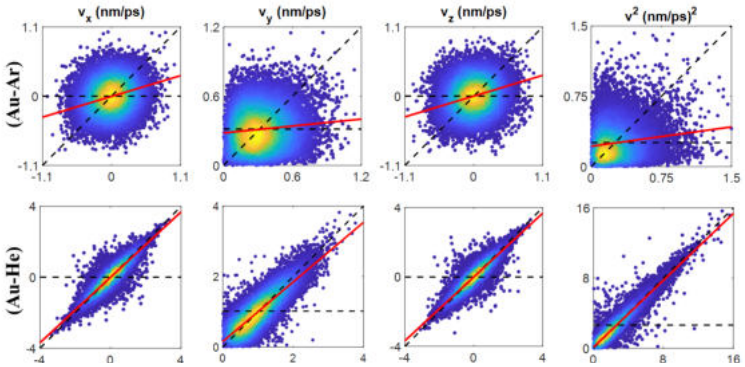

Figure 2.3: Velocity correlations of impinging (x-axis) and reflected (y-axis) velocity components of Ar and He on Au surface at 300 K using ab-initio potential. The dashes horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. The red line refers to the linear fit of the collision data obtained by MD simulations.

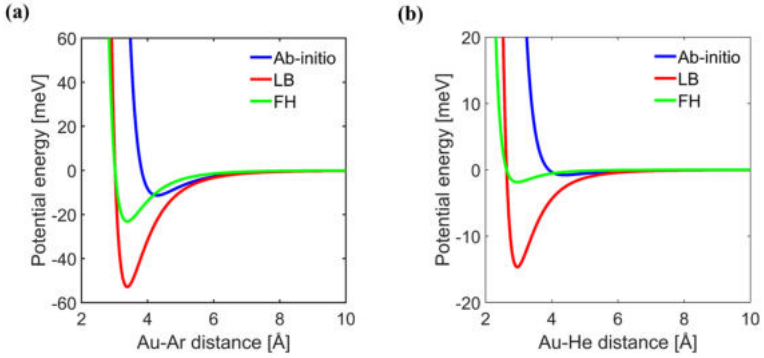

Figure 2.4: Pair potential energy plots of noble gases interaction with Au surface: (a) Au–Ar; (b) Au–He.

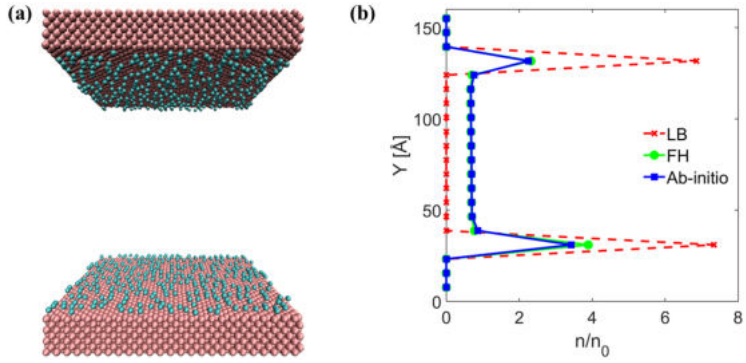

Figure 2.5: (a) Adsorption of Ar molecules on Au surface based on pair potential obtained from LB mixing rule; (b) normalized number density for different Au–Ar interaction potentials.

The obtained E/MACs for the aforementioned systems at the bottom wall (T_w= 300 K) are reported in Table 2.4. The reason why only E/MACs on bottom wall are reported here is for a further comparison with experimental data which are in the same temperature range. As it is shown in [63, 66], the temperature gradient between two walls causes only a minor reduction in obtained values for E/MACs on the bottom wall. Therefore, this can be neglected. However, in order to resemble the experimental two parallel plates, in which the presence of a small temperature gradient ($T_h - T_c \ll Tc$) is a must [84], a temperature difference was imposed between the two plates in our MD simulations. For both noble gas cases, E/MACs increase with increasing the potential well depth. This is in agreement with the behavior as reported in other numerical studies [64, 69]. The reason lies in the fact that higher gas–wall interaction strength increases the likelihood of trapping-desorption phenomenon, which at the end causes higher energy and momentum exchange between the solid surface and the neighboring gas. In the case of Au–Ar,

the reported empirical value for EAC (α_E) using two parallel plates approach at 296 K is 0.85 [84], which is in excellent agreement with obtained value for EAC using the potential based on ab-initio computations for our parallel walls assembly. Furthermore, Agrawal and Prabha [85] have reported that the tangential-MAC (TMAC) for Ar on commonly employed surface materials is 0.893, which this value also is consistent with obtained results for TMAC in different directions (α_x and α_z) in our case based on ab-initio pair potential. The EAC using FH mixing rule is 0.913, which is slightly higher than the value obtained by ab-initio potential.

Table 2.4: Variation of energy and momentum accommodation coefficients and MD simulations running time with the pressure in the simulation box for Au–Ar and Au–He pairs

System	Pair potential	α_x	α_y	α_z	α_E
Au-Ar	Ab_initio (Parallel walls)	0.824	0.913	0.832	0.874
	Ab-initio (Molecular beam) [70]	0.40	0.77	0.40	0.56
	Fender Halsey	0.915	0.934	0.913	0.913
	Experimental results: α_E = 0.85 [84] ; TMAC(α_x, α_z) = 0.893 [85]*				
Au-He	Ab_initio (Parallel walls)	0.036	0.0.113	0.038	0.048
	Ab-initio (Molecular beam) [70]	0.013	0.046	0.014	0.017
	Fender Halsey	0.245	0.347	0.221	0.069
	Lorentz-Berthelot	0.642	0.748	0.653	0.187
	Experimental results: α_E = 0.31 [84]				

* No temperature is reported in reference [85]

For Au–He, Trott et al. [84] have measured EAC = 0.31 at 296 K, which is higher than all values reported in Table 2.4 for the same combination. To elucidate the possible reason behind the observed mismatch, it is noteworthy to mention that the surface roughness and the type of premeasurement treatment employed to clean the surface under investigation have significant impacts on obtained experimental results for E/MACs. For instance, for He–Pt pair at room temperature Mann [86] reported EAC = 0.03. For the same combination, in another experimental study [87] the measured EAC at 303 K was 0.238. The most important difference between the two aforementioned studies is that in the first study the metal surface is perfectly clean, but in the latter case the metal surface is somehow contaminated and it is also partially covered by testing gas. In addition, in another experiment [88] using Tungsten as the substrate, it was reported that EAC for He can vary from 0.017 (clean surface) to 0.23 (untreated surface). In Reference [84] it is also mentioned that recontamination of the surface is highly possible after the surface treatment technique that they used. Considering that in the performed MD simulation the surface is assumed to be atomistically smooth and clean, it can be inferred that MD results for E/MACs at the first place should be compared with experimental results on very clean and pure solid surface. Herein, assuming that clean Au surface has a similar behavior as clean Platinum and Tungsten surfaces, it is seen that the obtained results for

EAC of Au–He using pair potential based on both FH mixing rule and ab-initio computations are similar to the values reported in [86] and [88] for He on clean Pt and Tungsten surface, respectively.

The presence of contamination on a real surface causes the accumulating of gas molecules on the surface which at the end leads to measuring higher E/MACs. To support this, the trajectories of Ar and He molecules interacting with Au surface using ab-initio computations pair potential during our MD simulation are depicted in Figure 2.6. In this figure the red crosses indicate the y-position of each gas molecules as function of simulation time. When an accumulation of red crosses is shown in certain area, it means that gas atom remains for a longer time in that area. It is shown that in the case of Ar molecule, which has higher molecular weight ($MW_{Ar}/MW_{He} \approx 10$) and stronger interaction with Au surface, higher number of multiple collisions with Au surface has occurred (accumulation of red dots in vicinity of Au surface for Au–Ar pair). This means that in MD simulation of Au–Ar pair, there is higher chance that we see a layer of gas molecules adjacent to the solid surface. This layer of gas molecules can resemble the contamination in the case of experimental study, in which achieving a real clean surface is very challenging. Due to the presence of this layer on the surface in the case of Au–Ar combination, gas–gas interactions near the wall are more frequent. Since the gas–gas interaction strength is typically higher than gas-wall interaction strength (see Tables 2.1 and 2.2), it can be deduced that forming a gas layer at vicinity of surface leads to deriving higher values for E/MACs by MD simulations. Therefore, it is more likely that E/MACs obtained by MD simulations for a heavy argon gas match the experimental results.

Liao et al. [70] computed also E/MACs for Ar–Au and He–Au pairs using a molecular

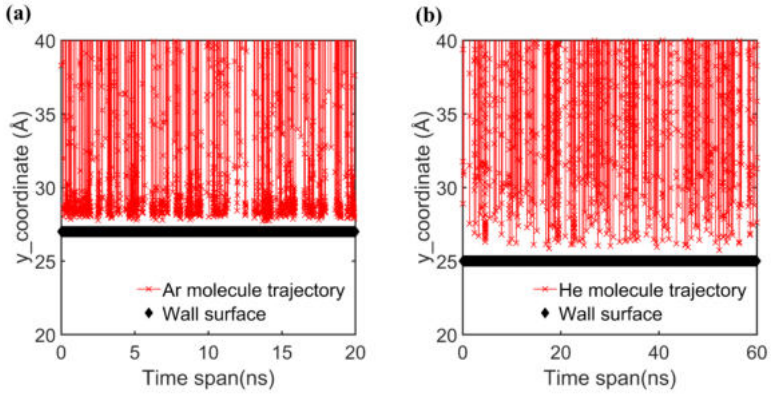

Figure 2.6: Trajectories of noble gases in vicinity of Au surface. (a) Ar–Au pair; (b) He–Au pair. Red crosses depict the y-coordinate of gas atoms as a function of simulation time.

beam approach based on the same ab-initio potential that was employed here. As it is shown in Table 2.4, their results for E/MACs are lower than those obtained in this study. In the case of molecular beam approach gas molecules adjacent to the surface are considered to interact only with wall atoms, and gas-gas interactions are ignored. While in the case of the parallel plates approach due to presence of other gas molecules in the

system, gas particles will be partly reflected back by other particles to the surface resulting in relatively more gas–wall interactions per molecule and therefore resulting in an increase in thermal and momentum accommodation coefficients.

2.4. CONCLUSIONS

The energy and momentum accommodation coefficients of monoatomic gases (Ar and He) on gold surface were determined through MD simulations. Initially, the impact of gas pressure on computed accommodation coefficients was investigated. Reducing gas pressure has a minor influence on accommodation coefficient values, whereas it can considerably increase the MD simulations time.

The effect of gas-wall interaction strength on accommodation coefficients was studied for different mixing rules, as well as for ab-initio pair potential. It was concluded that larger energy well depths in potential energy function leads to higher accommodation coefficients. For Ar–Au and He–Au gas-wall interactions, the energy well depth is overestimated for both Lorentz–Berthelot and for Fender–Halsey mixing rules resulting in an overestimation of the E/MACs. In the case of Ar–Au, the Lorentz–Berthelot mixing rule highly overestimates the potential well depth. This issue causes a fully saturated solid surface in which computing the accommodation coefficients from numerical point of view is not possible. It has been found out that the accommodation coefficients obtained by pair potential based on ab-initio computations are always in a reasonable agreement with experimental and numerical results in the relevant literature.

Comparing the obtained results for accommodation coefficients in this work with another study in which a molecular beam MD approach was used to compute accommodation coefficients reveals that gas–gas interaction is an important aspect that needs to be taken into account in the transient Knudsen regime since it leads to an enhancement in obtained accommodation coefficients.

3

MODELING RAREFIED MONOATOMIC GAS-SOLID SURFACE INTERACTIONS USING GAUSSIAN MIXTURE APPROACH

In rarefied gas flows, discontinuity phenomena such as velocity slip and temperature jump commonly appear in the gas layer adjacent to a solid boundary. Due to the physical complexity of the interactions at the gas-solid interface, particularly in the case of systems with local nonequilibrium state, boundary models with limited number of parameters cannot completely describe the reflection of gas molecules at the boundary. In this work, the Gaussian mixture (GM) model, which is an unsupervised machine learning technique, is employed to construct a statistical gas-solid surface scattering model based on the collisional data obtained from molecular dynamics (MD) simulations. The GM model is applied to study Couette flow for different inert gases (Ar and He) confined between two parallel infinite gold walls at different temperatures. A direct comparison between the results obtained from the GM model and the Cercignani-Lampis-Lord (CLL) scattering kernel against the MD collisional data in terms of the distribution of the predicted postcollisional velocities, and accommodation coefficients has shown that the results from the GM model are an excellent match with the MD results outperforming the CLL scattering kernel. As an example, for He gas, while the predicted energy accommodation coefficient by the CLL model is more than two times higher than the MD predictions, the value computed by the GM model is in excellent agreement with the MD results. This superior performance of the GM model confirms its high potential to derive a generalized boundary condition in systems encountered with highly nonequilibrium and complex gas flow conditions.

This chapter is published as a peer-reviewed article. Mohammad Nejad, S., Iype, E., Nedea, S., Frijns, A. and Smeulders, D. Modeling rarefied gas-solid surface interactions for Couette flow with different wall temperatures using an unsupervised machine learning technique. *PHYSICAL REVIEW E*, 104(1), p.015309 (2021).

3.1. INTRODUCTION

Heat and momentum transfer of fluid flows in contact with surfaces in noncontinuum or rarefied conditions has been an active research topic in numerous modern engineering applications such as vacuum technology, micro or nanoelectromechanical systems (M/NEMS), astronautics, and particle sizing techniques used in the aerosol industry [1, 89–91]. In these systems jump conditions for temperature and velocity, which manifest themselves mainly at the gas-solid interfaces, are of crucial importance in computing the drag force and heat transfer experienced at the solid surfaces. To simulate the fluid flow in such systems under the moderately rarefied condition, that means a Knudsen number (Kn) less than 0.1, the Navier Stokes (NS) equations may be used but slip boundary conditions should be considered. However, for a higher degree of rarefaction (Kn > 0.1) the NS equations break down and must be substituted by more sophisticated equations such as Boltzmann equations. Computational-particle-based simulation techniques such as direct simulation Monte Carlo (DSMC) [11], Lattice Boltzmann Method (LBM) [61], and method of moments (MoM) [62] are commonly employed to find approximate solutions to the different forms of the Boltzmann equation. Nevertheless, in all these simulation techniques, rigorous prescription of boundary conditions at solid surfaces is the crucial parameter to guarantee the reliability of the simulation results. Although the gas-surface interaction (GSI) has been long studied [59, 92, 93], and a wide range of numerical models have been developed to describe it, a generalized formalism applicable in varied applications is still missing. The most well-known available GSI models are phenomenological models expressed by scattering kernels, which are probability density functions (PDF) correlating the velocity distributions of the gas molecules before and after colliding with the solid surface.

The Maxwell model [59] is the oldest empirical GSI model. Maxwell postulated that a fraction of the incident gas molecules was reflected in a diffuse manner, while the remaining part underwent a specular reflection. Despite its simplicity and acceptable accuracy, the Maxwell model was incapable of reproducing the lobular patterns observed in the molecular beam experiment [94]. Later on, to fulfill this shortcoming, Cercignani and Lampis [92] developed a more elaborate scattering kernel, which was extended further by Lord [93], and showed better performance compared to the Maxwell model [95]. The Maxwell and Cercignani-Lampis-Lord (CLL) models are controlled by one or several constant parameters known as accommodation coefficients (ACs), which quantify the energy and momentum exchange at the gas-solid interface and must be known a *priori*. In fact, evaluating ACs is the major bottleneck in the performance of the aforementioned empirical scattering models. This is due to the superposition of many factors affecting GSI at the microscopic level, which causes a notable discrepancy among the reported values for ACs in experimental and numerical studies [38, 84, 85, 96]. Several investigations [95, 97–99] assessed the performance of the existing empirical GSI models. Considering hard assumptions like, for instance, the usage of constant ACs for all impinging gas molecules, assuming fixed functional relationships between the momentum and energy ACs, as well as neglecting the interplay between different components of gas molecule velocities, these models do not fully capture the physics behind the gas-wall scattering process. Moreover, these models are not capable of capturing the flow properties in a highly nonequilibrium state, frequently present in M/NEMS. This involves, for

instance, impinging or reflecting molecular gas fluxes for complex flow conditions and for complex interfacial structures and conditions [97, 100]. To construct more reliable scattering kernels some researchers extended the classical GSI models. For instance, Struchtrup [22] incorporated a velocity dependent AC with classical Maxwell scattering kernel. To describe gas scattering adjacent to an anisotropic surface, Dadzie and Meolans [19], as well as To et al. [20] proposed the usage of different ACs in each spatial direction. However, these boundary models, due to higher number of tunable parameters, are very complex to implement, and therefore have not been employed much in practice. To study Couette flow using DSMC simulations Yamamoto et al. [21], adapted the Maxwell model by using different directional ACs to generate each velocity component. Nevertheless, it was shown that the velocity correlations obtained from the Yamamoto model can still be very different from the explicit wall simulations [66].

A promising approach to achieve a detailed understanding about GSI at the atomistic level is molecular dynamics (MD) simulations, which has been successfully interplayed with DSMC to study rarefied gas flow [101–103]. However, it is well known that MD simulations are computationally expensive. A new class of scattering kernels, known as nonparametric scattering kernels, were proposed in the literature [104–106]. The main idea behind developing such models is to overcome some intrinsic limits of the existing empirical GSI models, which are mainly caused by employing a limited number of fitting parameters (i.e., accommodation coefficients). Generally speaking, in these models collision data obtained by MD simulations are directly used to construct the scattering kernel and no additional parameters such as ACs are needed to be computed in between. Assessment of the postcollisional velocities predicted by nonparametric GSI models against the results obtained from the classical parametric GSI models, MD simulations results, as well as experimental studies confirm the superior performance of the nonparametric GSI models.

Recently a new type of GSI model using the Gaussian mixture (GM) model, which is a robust unsupervised machine learning technique, was proposed by Liao et al. [33]. In the aforementioned study, for training the GM model the collisional data obtained from the molecular beam MD setup were utilized, in which only gas-wall interactions were taken into account, and gas-gas interactions were ignored. The main advantage of using the GM model over the classical scattering kernels is that since it does not require the use of any ACs as input parameters, it maintains all the important physics included in the collisional data. Investigating both smooth and rough atomic surfaces, they reported a good performance of the GM model in predicting physical and statistical properties of the gas-wall interfaces. However, it was shown that in many microgas flow systems the adsorption of gas molecules on solid surfaces impacts the gas scattering from the surfaces [73]. In the presence of the gas adsorbed layer, both the gas-gas and gas-wall interactions are not negligible. To our knowledge, no GM-based model exists that includes gas flows, gas-gas interactions, and highly nonequilibrium effects.

The main aim of the present work is to investigate the capability of the GM model in the study of the rarefied gas-solid surface interactions in the case of systems encountered with highly nonequilibrium and flow conditions. Herein, considering that both gas-gas and gas-wall interactions can influence the gas flow solution in the vicinity of a solid wall, a nanochannel system consisting of two infinite parallel plates with Ar and

He gases confined between them is chosen as a case study. A Couette flow was used to study the influence of the wall velocity on the residence time of the gas molecules in the interface region and the molecular scattering effects. Two types of systems, with walls at the same temperature and with walls at different temperatures, are studied using MD simulations. In each case, the performance of the GM model is evaluated in comparison with the CLL model, as well as the original MD-obtained collision data. The evaluation is carried out in terms of correlations between gas molecules's velocities before and after colliding with the surface, the PDF of the postcollisional gas velocities, and predicted ACs.

3.2. METHODOLOGIES

3.2.1. MD SIMULATION

The three-dimensional schematic diagram of the MD simulation setup is shown in Figure 3.1. The system model consists of two parallel walls including of 8750 Au atoms each with either 800 Ar or 400 He atoms confined between them. In each wall Au atoms are arranged in a FCC lattice structure with a cross section area of 10 nm by 10 nm. The normal distance between the walls d is fixed at 12 nm for Ar and 102 nm for He gas. The number of gas atoms confined between the walls, and the distance between the walls were chosen in a way that the Knudsen number remains in the early transition regime ($0.1 < \text{Kn} < 1$), and the pressure in the gas domain does not overpass the critical pressure (P_{cr}) value in the system (for Ar: $P_{cr} = 4.89$ MPa, and for He: $P_{cr} = 0.23$ MPa [107]). Therefore, for the Ar-Au system the Knudsen number and the reduced density (η) are Kn = 0.2 and $\eta = 0.02$, respectively. In addition, for the He-Au system Kn = 0.71 and $\eta = 0.00045$, respectively.

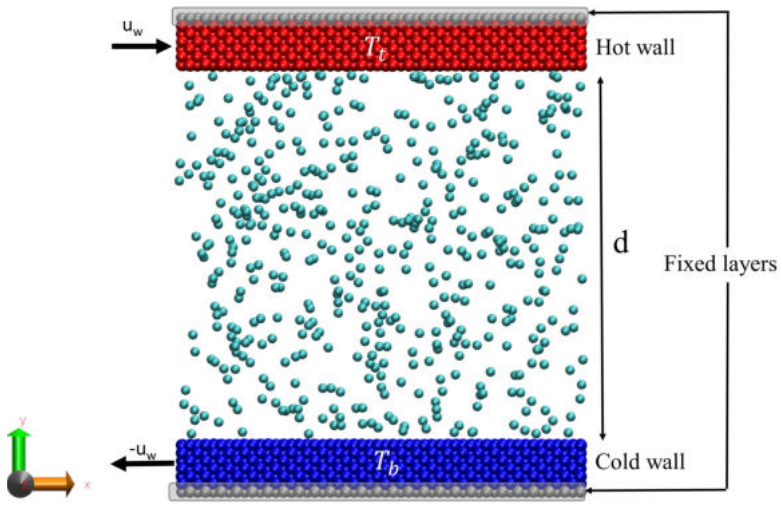

Figure 3.1: Schematic representation of the system under consideration for the MD simulations; d: distance between the two walls; T_b: Temperature of the bottom wall;T_t: Temperature of the top wall; u_w: Imposed velocity on the walls.

A Berendsen thermostat with a damping constant of 100 fs is employed to maintain walls temperature at the desired values. Regarding the gas atoms, initially their temperature is set at the 300 K. Afterwards, gas atoms are in a microcanonical ensemble (NVE) and their temperature can only evolve through collisions with other atoms in the simulation box.
In this system, periodic boundary conditions are applied in the directions parallel to the wall (x,z) such that only a section of an infinite wall has to be modeled. The interaction between Au atoms are modeled by the embedded atom model (EAM) potential [77]. The gas-gas and gas-solid interactions are modeled by the Lennard-Jones (LJ) potential and the potential parameters are listed in Table 3.1. It is noteworthy to mention that in the atomistic study of the GSI, the pair potential parameters employed to describe the interaction strength at the gas-solid interface are of prominent importance [108, 109]. It has been shown that the interaction potentials based on mixing rules cannot accurately describe GSI, and pair potentials based on the quantum calculations are the most reliable ones to study the GSI at the molecular scale [69, 75, 108]. Therefore, in this work the gas-wall interaction potential parameters are calculated based on existing quantum *ab-initio* computations [70]. To model the Couette flow condition, walls have been moving with the relative velocity difference of $2\,u_w$ (see Figure 3.1). The speed ratio S_w is defined as the ratio between the wall velocity u_w and the most probable speed of the gas at the temperature T_b: $S_w = \frac{u_w}{\sqrt{2k_B T_b/m_g}}$, where k_B is the Boltzmann constant, T_b is the temperature of the bottom wall, and m_g is the mass of the gas atom.
In all the simulations, after performing the initial surface energy minimization, the MD setup has been equilibrated at the desired temperature for 3 ns using the time step of $\Delta t = 1$ fs and 0.5 fs for Ar-Au and He-Au systems, respectively. Once the system is fully equilibrated the production run is started. To guarantee reliable statistics it is proceeded for the next 25 ns and 100 ns for Ar-Au and He-Au systems, respectively. More details regarding the gathering collisional data based on the discussed MD simulation setup can be found in our previous work [108]. All the MD simulations were performed using the LAMMPS [81] molecular dynamics package.

Table 3.1: Lennard–Jones potential parameters

Atom pair	ϵ (meV)	σ (Å)
Ar-Au	11.36	3.819
He-Au	0.787	4.342
Ar-Ar	12.2	3.35
He-He	0.94	2.64

3.2.2. CLL SCATTERING KERNEL

When considering the flow of a dilute gas adjacent to a solid surface, prescribing the proper boundary condition is the key factor for obtaining a reliable solution of the Boltzmann equation. In other words, it is of crucial importance to know how gas molecules impinging with certain velocity with a surface will be reflected from it. In kinetic theory analysis, the gas-surface interaction models are employed as boundary conditions

for the Boltzmann equation. The GSI model, also called the scattering kernel, is defined by the conditional probability density function $f(\boldsymbol{v}|\boldsymbol{v'})$, which represents the probability density that an impinging gas molecule with velocity $\boldsymbol{v'}$ is rebounced with velocity $\boldsymbol{v}$ from the surface. The CLL model is the most accurate phenomenological gas-surface collision model because of its relatively well-defined mathematical framework leading to the satisfactory prediction of the scattering process. In the CLL model the gas molecule velocity components in the tangential (v_t) and normal (v_n) directions are assumed to be independent from each other, and they can be determined through Eqs. (3.1) and (3.2), respectively, as follows:

$$f^{CLL}(v_t|v_t') = \frac{1}{\sqrt{\pi\alpha_t(2-\alpha_t)}}\exp-\frac{[v_t-(1-\alpha_t)v_t']^2}{\alpha_t(2-\alpha_t)}, \tag{3.1}$$

$$f^{CLL}(v_n|v_n') = \frac{2v_n}{\alpha_{NE}}\exp\left[-\frac{v_n^2+(1-\alpha_{NE})v_n'^2}{\alpha_{NE}}\right]I_0\left[\frac{2(\sqrt{1-\alpha_{NE}})v_n v_n'}{\alpha_{NE}}\right], \tag{3.2}$$

where α_t and α_{NE} are the accommodation coefficients corresponding to the tangential momentum and normal kinetic energy, respectively. I_0 denotes the modified Bessel function of the first order and zeroth order. Here the velocities are normalized by $\sqrt{2RT_w}$, where R is the specific gas constant ($\frac{k_B}{m_g}$) and T_w is the wall temperature. In this model, a fixed empirical correlation is imposed between α_t and the tangential energy accommodation coefficient α_{TE}, that is, $\alpha_{TE}=\alpha_t(2-\alpha_t)$. To generate new velocities after collision according to the CLL scattering kernel the algorithm provided in the paper by Peddakotla et al. [110] is employed, in which including a separate AC for each spatial direction assists the comparison with MD simulation results.

Accommodation coefficients used normally to characterize the energy and momentum exchange in the gas-solid interface are the main input parameters required in the CLL model. In this work, to compute ACs the method proposed by Spijker et al. [66], which also allows evaluating ACs for an isothermal system, is employed. In this approach, the slope of the best least-squares linear fit of the MD obtained collisional data is used for computing various ACs

$$\alpha_q = 1-\frac{\sum_i(Q_I^i-\langle Q_I\rangle)(Q_R^i-\langle Q_R\rangle)}{\sum_i(Q_I^i-\langle Q_I\rangle)^2}. \tag{3.3}$$

where subscript q can refer to various gas molecule kinematic properties such as its momentum or kinetic energy in a certain direction. Q_I^i and Q_R^i represent the precollisional and postcollisional value of the same property for the gas particle, respectively. The notation $\langle . \rangle$ indicates the ensemble average of the molecular property Q.

3.2.3. GAUSSIAN MIXTURE MODEL

The Gaussian mixture model is an unsupervised learning method, which found its applications in speech recognition [111], image, and pattern recognition [112, 113] among many other areas. Generally, the GM model is categorized as a parametric model, but due to the presence of the numerous fitting parameters in the model it retains the flexibility of a nonparametric model. Therefore, it can be considered as a promising tool to model high-dimensional data space. The idea of representing data as a collection of D-dimensional Gaussians (D being the number of features) is desirable in its application to predict scattering velocity distributions, which are also mostly Gaussians. A GM model of the N_G component Gaussians with mean vector $\vec{\mu}_i$ and covariance matrix Σ_i is given by Eq. 3.4.

$$p(\vec{x}) = \sum_{i=1}^{N_G} w_i g\left(\vec{x}|\vec{\mu}_i, \Sigma_i\right) \tag{3.4}$$

where,

$$g\left(\vec{x}|\vec{\mu}_i, \Sigma_i\right) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} exp\left[-\frac{1}{2}(\vec{x}-\vec{\mu}_i)'\Sigma_i^{-1}(\vec{x}-\vec{\mu}_i)\right]$$

The parameters ($\{w_i, \vec{\mu}_i, \Sigma_i\}\ \forall\ i$ in $\{1 \cdots N_G\}$) in the model are optimized using the expectation-maximization (EM) algorithm [114]. All calculations in this work used the GM model package available in SCIKIT-LEARN [115] with all default settings except for the number of Gaussians (N_G). In fact, in the case of training a GM model N_G is the only parameter, which needs to be assigned by the user. It plays the role of a fitting parameter in the model and it has a direct impact on the performance of the model. Therefore, to prevent the occurrence of overfitting or underfitting in the model defining an optimal value of N_G is crucial. A detailed discussion on the effect of N_G is provided in section 3.3. Here, the entire collisional data (i.e., pre and post collision velocities) obtained from MD simulations are considered as the training data for the GM model. This means the interplay between different velocity components is already included in the model. After training the model, using the optimum model parameters ($w_i, \vec{\mu}_i, \Sigma_i$), the conditional form of the scattering kernel required by coarse grained simulation techniques can be obtained as a mixture of Gaussian functions.

3.3. RESULTS AND DISCUSSION

3.3.1. DIFFERENT APPROACHES FOR FEEDING THE TRAINING DATA INTO GAUSSIAN MIXTURE MODEL

Considering the collision data obtained from MD simulations, the tangential velocity components (v'_x, v'_z, v_x, v_z) follow Gaussian distributions. However, the normal velocity components (v'_y, v_y) follow a Rayleigh distribution. In Ref [33] the authors suggested a preprocessing scheme (see $Scheme-I$ in Figure 3.2). They mainly employed such a scheme to transfer the normal velocity components from Rayleigh distribution to a Gaussian distribution, to obtain the same distribution as the tangential velocity components. This leads to a consistency in the overall distribution (Gaussian) with the basis sets (also Gaussian) in the GM model. For this purpose, initially for each velocity pair ($\boldsymbol{v'}, \boldsymbol{v}$) an implicit pair ($-\boldsymbol{v'}, -\boldsymbol{v}$) was added to the data set. Afterwards, implementing the transfer function proposed in Ref [33] and given in Equation3.5 (step 2 in $Scheme-I$) results in the Gaussian distributions for the normal velocity components ($v'^{*,MD}_y$, $v^{*,MD}_y$).

$$T(\zeta) = \sqrt{2\beta}\, erf^{-1}\left[1 - 2\exp\left(-\frac{\zeta^2}{2\beta}\right)\right], \beta = \frac{k_B T_g}{m_g}, \tag{3.5}$$

$$T^{-1}(\zeta) = \sqrt{-2\beta \ln\left[\frac{1}{2} - \frac{1}{2}\, erf(\frac{\zeta}{\sqrt{2\beta}})\right]}, \tag{3.6}$$

where T_g is the gas temperature and can be computed from the average gas kinetic energy.

In Ref [33], it was mentioned that applying the introduced preprocessing step on the input data leads to a better performance of the GM model. However, it was not clarified which were the key factors considered to play a role in analyzing the GM model performance. Having a purpose slightly different than the one in Ref [33], that is in applying the GM model to study the correlation between impinging and outgoing velocity components in each direction rather than for overall kinetic energy, it was important to analyze the impact of applying the transfer function in different steps on the molecular scattering results for different directions. Therefore, to investigate the impact of the proposed scheme on the performance of the GM model in our system, two different approaches for implementing the collision data obtained by MD simulations were employed. In the first case ($Scheme-I$), preconditioning was carried out before logging the collision data into the GM model. In the second case ($Scheme-II$), as it is depicted in Figure 3.2, MD simulations's collisional data were directly fed into the GM model without using any preprocessing. It is noteworthy to mention that in the first approach, all the obtained velocity components by the GM model ($v^{GM}_x, v^{GM}_y, v^{GM}_z$) are Gaussian distributions. Therefore, using the reverse transfer function proposed in Ref [33] and shown in Equation3.6 the normal velocity component is mapped back into the Rayleigh distribution, $v^{*,GM}_y$, (step 5 in $Scheme-I$) to compare with the initial MD results. In $Scheme-II$ the predicted normal velocity component by the GM model resembles a Rayleigh distribution. Therefore the obtained distributions can be used directly for the comparison purpose. Herein, the system with walls at the same temperature ($T_b = T_t = 300\ K$), and without

imposing any external velocity on the walls is used as a benchmark. The resulting velocity correlation distribution and PDF of the outgoing velocity for each molecular velocity component in the case of the He-Au and Ar-Au systems are shown in Figures 3.3 and 3.4, respectively.

As it is depicted in Figure 3.3, for the He-Au system both $Scheme-I$ and $Scheme-II$ approaches are in a good agreement with MD simulation results. In the case of Ar-Au (see 3.4) the only notable difference between MD, $Scheme-I$, and $Scheme-II$ results is in the PDF of the normal velocity component (last column, second row).

For the Ar-Au system, the number of Gaussian functions N_G employed in the GM model was increased to determine whether it can improve the predicted velocities by $Scheme-II$ in the normal direction. In Figure 3.5, the PDFs for the normal velocity component derived by $Scheme-I$ based on $N_G = 100$ (referred as $Scheme-I_{100}$), $Scheme-II$ (referred to as $Scheme-II_{N_G}$, $100 \leq N_G \leq 600$), as well as MD results are represented. It is shown that by increasing N_G the results obtained by $Scheme-II$ converge towards the results obtained by MD simulations and the $Scheme-I$ model, and at $N_G = 600$ it is in excellent agreement with the other two methods.

In addition to the velocity correlations, the energy accommodation coefficient (EAC) was also computed to evaluate the performance of the GM model based on the introduced schemes for feeding the MD data into the model for training purposes. In each scheme N_G was varied from $N_G = 1$ to $N_G = 1000$. In Figure 3.6, it is seen that the variation in the EAC with increasing N_G is rather smoother in the case of $Scheme-I$. In addition, for $N_G = 600$ EAC computed by both GM models are in a perfect match with MD results.

From this analysis, it is deduced that generally $Scheme-I$ has a better performance than $Scheme-II$. Therefore, in order to achieve high precision in the predicted velocity correlations and computed ACs, in the remaining part of this work $Scheme-I$ based on $N_G = 600$ is employed for training the GM model. The training of the model for each case takes within 10 to 12 minutes on a regular laptop computer.

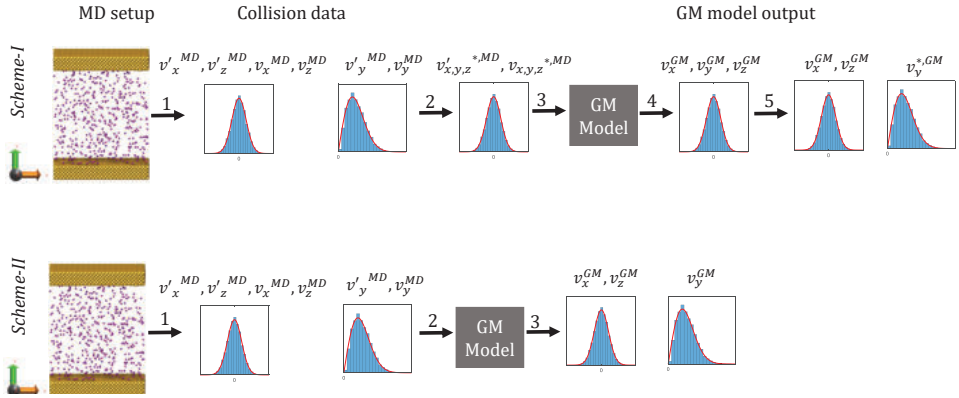

Figure 3.2: Different schemes used to feed the collision data obtained by MD simulations into the GM model; $v'_x{}^{MD}, v'_y{}^{MD}, v'_z{}^{MD}, v_x^{MD}, v_y^{MD}$, and v_z^{MD} are velocity components obtained from MD simulation, respectively; $v'_{x,y,z}{}^{*,MD}$ and $v_{x,y,z}^{*,MD}$ are transformed incoming and outgoing velocity vectors obtained from MD simulations; v_x^{GM}, v_y^{GM}, and v_z^{GM} are velocity components predicted by GM model; $v_y^{*,GM}$ is the normal velocity component predicted by the GM model and mapped back into the Rayleigh distribution.

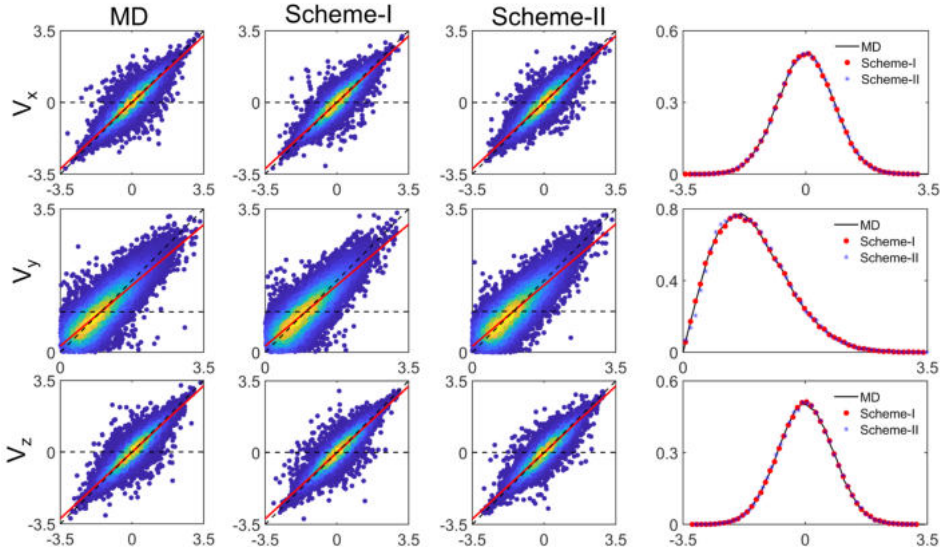

Figure 3.3: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components both in nm/ps for He-Au system. The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for reflecting particles are shown. *Scheme-I*: uses preprocessed MD results. *Scheme-II*: uses MD results without any preprocessing.

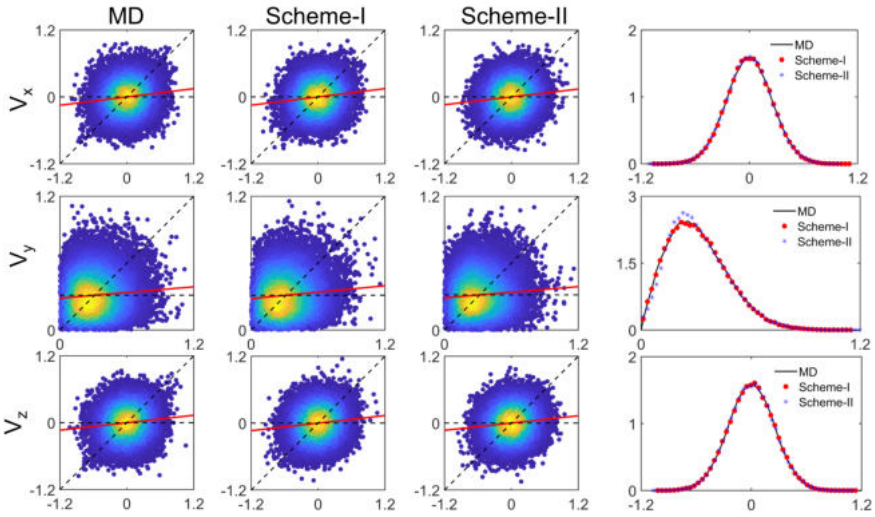

Figure 3.4: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components both in nm/ps for Ar-Au system. The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for reflecting particles has been shown. *Scheme-I*: uses preprocessed MD results. *Scheme-II*: uses MD results without any preprocessing.

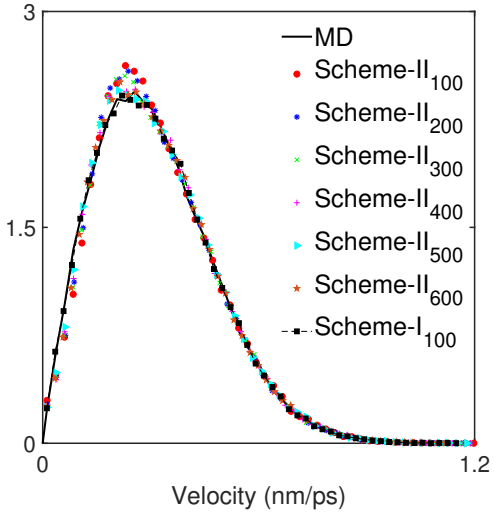

Figure 3.5: Impact of the number of Gaussian functions (N_G)in the predicted velocity distribution in the normal direction by GM model in the case of Ar-Au system.$Scheme-I_{100}$: Uses preprocessed MD results based on $N_G = 100$. $Scheme-II_{100 \leq N_G \leq 600}$: uses MD results without any preprocessing.

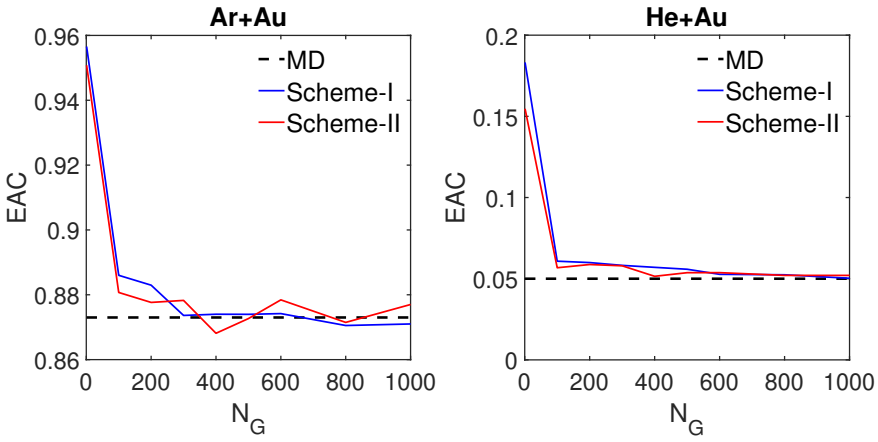

Figure 3.6: Variation of the energy accommodation coefficient obtained from the MD simulation and the GM model using different numbers of Gaussian functions N_G.

3.3.2. PERFORMANCE EVALUATION OF GM MODEL

In the next step, the performance of the GM model in comparison with the CLL model, which is one of the most used stochastic GSI models, and MD simulations results is studied. The comparison was carried out using the velocity correlation between impinging and outgoing velocities, and the PDF of the outgoing velocities in each spatial direction. Besides that, based on the predicted velocities via each of the GSI models, various accommodation coefficients including the tangential momentum accommodation coefficients in the both x and z directions (α_x, α_z), the normal energy accommodation coefficient (α_{NE}), and the energy accommodation coefficient (α_E) are computed and verified against the MD simulations results.

For both Au-Ar and Au-He systems, two kinds of physical problems are investigated. In the first case study, the temperature of both walls are set to 300 K, and in the second case, the temperature of the bottom wall surface T_b is set to 300 K, and that of the top wall surface T_t is set to 500 K. Besides that, in both cases the impact of flow condition on the performance of the GSI models was also investigated. As it is represented in Figure 3.1, to achieve this goal an external velocity is imposed to the bottom and top walls in the $-x$ and x directions, respectively. The following speed ratios are considered in the simulations: $S_w = 0.25$ and $S_w = 0.5$. For each system the most significant results are discussed in the main text and the remaining ones for Ar-Au system can be found in Appendix B (Figures B.1-B.4), and for He-Au can be found in Appendix C (Figures C.1-C.6).

Normally, to derive a machine-learning-based model for a system, two different sets of data are required. The first set, called the training set, is used to derive the model parameters, and the second set, called the test set, is utilized to validate the model. In this study, before gathering the collisional data for computing accommodation coefficients (ACs) or training the GM model, it was ensured that the system has reached equilibrium condition. Therefore, for each MD setup, characterized by a unique set of walls temperatures and velocities, there will not be any noticeable difference in collisional data

obtained from different MD simulations. To prove this, for a specific MD setup, two independent MD simulations were conducted. One used as the training set (MD_{train}), and the other one used as the test set (MD_{test}). For each setup, using a different seed number in the Gaussian velocity generator, a different ensemble of initial velocities was created for the gas atoms. The velocity correlations between impinging and outgoing velocities, the PDF of the outgoing velocities, and the computed ACs are represented in the Appendix A (Figure A.1 and Table A.1). It was shown that the obtained results from the MD_{train}, MD_{test}, and the GM model are in perfect match with each other.

3.3.2.1. *Ar-Au system*

Figure 3.7 shows the velocity correlations and the outgoing velocity PDFs in the case of a no-flow Ar-Au system, in which both walls have the same temperature ($T_b = T_t = 300$ K). It was observed when the gas and the both walls are in equilibrium state with each other, the predicted results by the CLL and GM models are in a good agreement with the ones obtained by MD simulations.

In the next step, the flow velocity was added to the system by imposing a velocity to the walls in the the x-direction. As it is depicted in Figure 3.8, which belongs to the case of $S_w = 0.5$, there is a good agreement between the predicted PDFs in the y and z directions by both GSI models and the MD simulations's results. On the contrary, in the x-direction, which is the direction of the imposed velocity, the results predicted by the CLL model deviate from the MD results while the results from the GM model are still in excellent agreement with the MD results. Herein, the mean value of the reflected velocities obtained from MD simulation and the GM model are 0.122 and 0.124 nm/ps, respectively. While the predicted mean value using the CLL model is 0.00 nm/ps. It is noteworthy to mention that for the system with $S_w = 0.25$ the same kind of trend has been observed (see Appendix B). In this case, while the predicted mean velocities in the x-direction by MD simulation and the GM model are 0.0627 and 0.063 nm/ps, respectively, the value predicted by the CLL model is 0.001 nm/ps.

Figures 3.9 and 3.10 illustrate the resulting velocity correlations of impinging and reflected molecules after the collision with the bottom (cold) and the top (warm) walls in the case, in which there is temperature gradient between the walls ($T_b = 300$ K, $T_t = 500$ K). It is seen that at the both walls the predicted results by the GM model are in perfect match with the MD simulations results. However, the results obtained by the CLL model for the tangential velocity components (v_x, v_z) at both walls slightly deviated from the MD results at the peak of velocity PDFs.

The velocity distributions at the bottom wall for the system with temperature gradient between the walls and flow condition with $S_w = 0.5$ are represented in Figure 3.11. Similar to the earlier cases, the predicted results from the GM model have a very good agreement with MD simulations, while a mismatch is observed between results obtained by MD simulations and the CLL model in the tangential directions. This mismatch is more evident in the direction of imposed velocity (i.e., the x direction).

For the Ar-Au system, the computed ACs in the case of systems with the walls at the same and different temperatures are reported in Tables 3.2 and 3.3, respectively. First of all, by going through the results in the aforementioned tables it can be understood that imposing the external velocity does not have a significant impact on the obtained ACs on the walls. Such behavior was also reported in Ref [116], where a gas kinetic model

was employed to study the GSI in a Couette flow system. Besides that, in the case of the system with the walls at different temperatures (see Table 3.3) all the ACs computed on the top wall are lower than the ones computed on the bottom wall. This behavior is due to the higher surface temperature at the top wall and was also previously reported [66]. Considering the directional accommodation coefficients including α_x, α_z, α_{NE}, it was observed that the results obtained by both GSI models are in good agreement with the MD simulation results. However, it is found that α_E obtained by the CLL model is slightly higher that the MD results, whereas the results obtained by the GM model is always in perfect match with the MD ones.

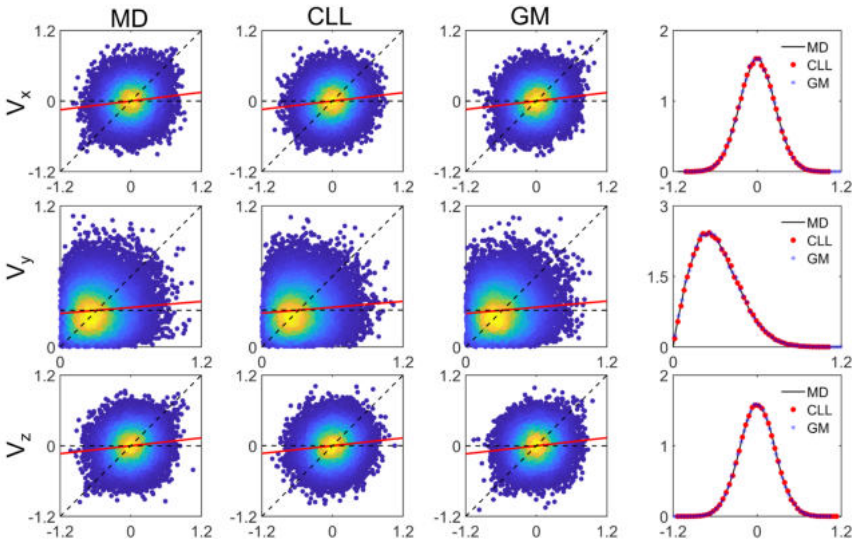

Figure 3.7: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for Ar-Au system (T_b = 300 K, T_t = 300 K, S_w = 0). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

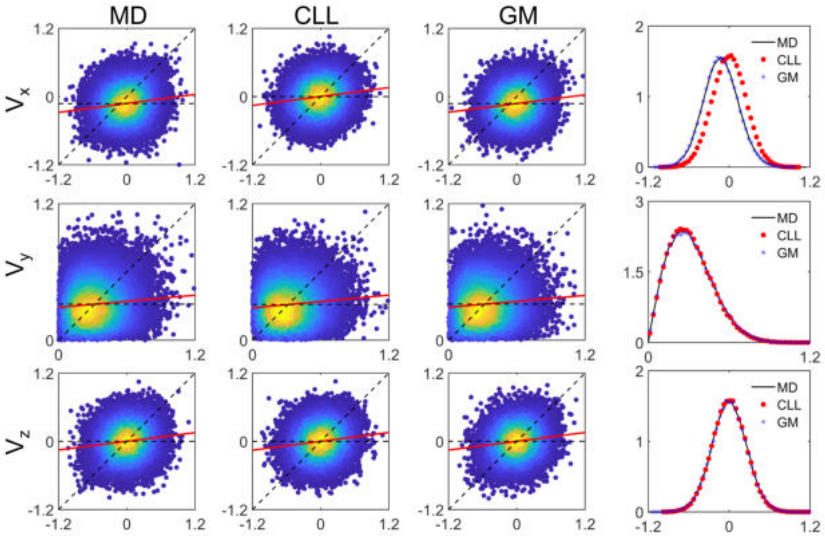

Figure 3.8: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for Ar-Au system at the bottom wall ($T_b = 300$ K, $T_t = 300$ K, $S_w = 0.5$). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

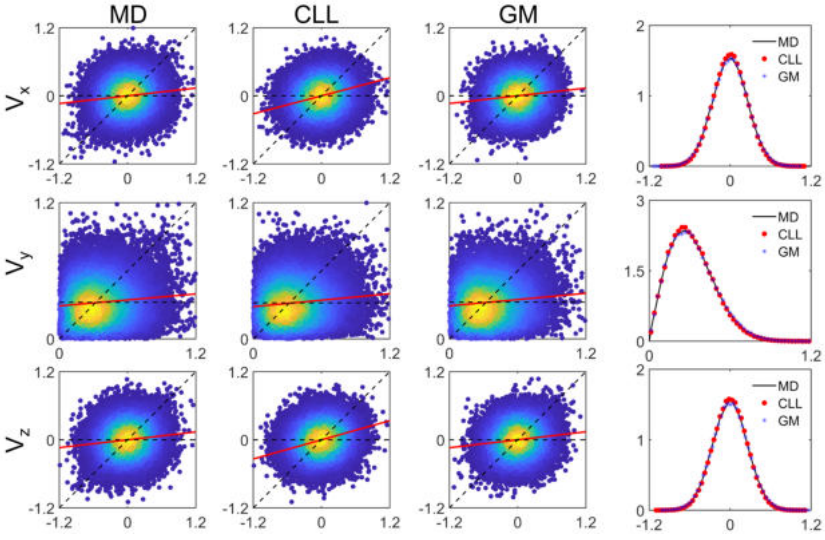

Figure 3.9: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for Ar-Au system at the bottom wall ($T_b = 300$ K, $T_t = 500$ K, $S_w = 0$). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

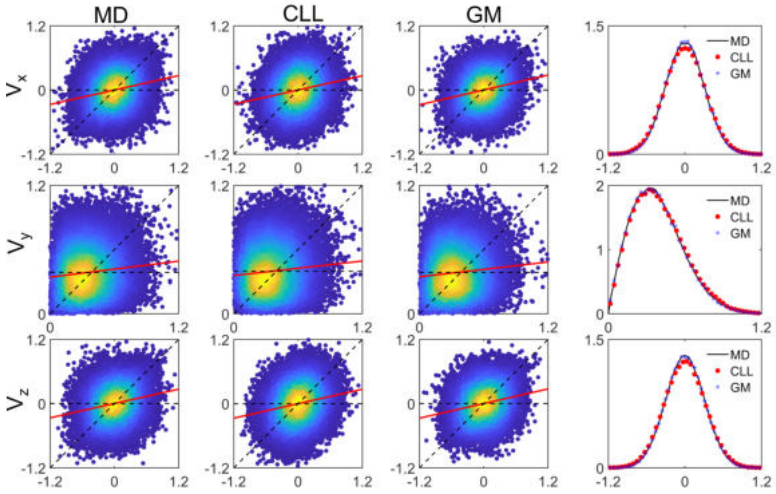

Figure 3.10: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for Ar-Au system at the top wall (T_b = 300 K, T_h = 500 K, S_w = 0). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data.In the last column the corresponding probability density function for the reflecting particles is shown.

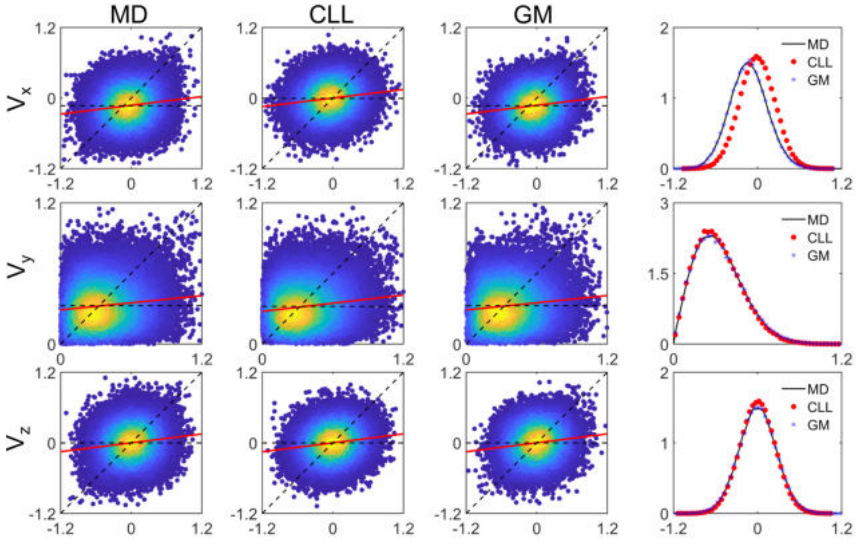

Figure 3.11: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for Ar-Au system at the bottom wall (T_b = 300 K,T_t = 500 K, S_w = 0.5). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

Table 3.2: Tangential momentum (α_x,α_z), normal energy (α_{NE}), and energy (α_E) accommodation coefficients for two parallel walls Ar-Au system with the walls at the same temperature at various speed ratios (S_w), computed using different stochastic approaches: CLL and GM models, and MD simulations.

S_w	Model	α_x	α_z	α_{NE}	α_E
0	MD	0.876	0.888	0.91	0.873
	CLL	0.877	0.886	0.91	0.947
	GM	0.879	0.888	0.909	0.884
0.25	MD	0.879	0.88	0.916	0.878
	CLL	0.873	0.878	0.914	0.95
	GM	0.879	0.877	0.921	0.888
0.5	MD	0.868	0.873	0.903	0.87
	CLL	0.864	0.88	0.901	0.94
	GM	0.872	0.871	0.911	0.891

Table 3.3: Tangential momentum (α_x,α_z), normal energy (α_{NE}), and energy (α_E) accommodation coefficients for two parallel walls Ar-Au system with a temperature gradient between the walls at various speed ratios (S_w),computed using different stochastic approaches: CLL and GM models, and MD simulations.B: bottom wall; T: top wall.

S_w	Model	Wall	α_x	α_z	α_{NE}	α_E
0	MD	B	0.886	0.884	0.91	0.873
		T	0.775	0.777	0.856	0.792
	CLL	B	0.884	0.884	0.908	0.943
		T	0.778	0.777	0.855	0.907
	GM	B	0.887	0.882	0.908	0.882
		T	0.767	0.772	0.871	0.81
0.25	MD	B	0.875	0.876	0.898	0.865
		T	0.771	0.769	0.851	0.778
	CLL	B	0.877	0.881	0.899	0.937
		T	0.772	0.766	0.848	0.901
	GM	B	0.877	0.88	0.903	0.883
		T	0.772	0.773	0.851	0.79
0.5	MD	B	0.876	0.875	0.897	0.869
		T	0.76	0.761	0.844	0.773
	CLL	B	0.875	0.875	0.898	0.94
		T	0.76	0.763	0.843	0.897
	GM	B	0.877	0.873	0.897	0.877
		T	0.76	0.767	0.844	0.789

3.3.2.2. *He-Au system*

The same sort of assessment was also carried out in the case of the He-Au system. The velocity distributions for the system, in which both walls have the same temperature ($T_b = T_t = 300$ K), without and with imposing the external velocity ($S_w = 0.5$) are shown in Figures 3.12 and 3.13, respectively. Comparing the velocity correlations obtained by the molecular simulations and the GSI models, the most significant observed difference is that the results from the CLL model are narrower than those computed by the MD simulations and GM model. Furthermore, inducing wall velocity (see Figure 3.13) does not have any noticeable influence on the system behavior, and as it is also illustrated in the aforementioned Figs, the velocity PDFs obtained by both GSI models are in accordance with the MD results. The resulting velocity distributions at the top wall of a system, in which two walls are kept at different temperatures ($T_b = T_t = 500$ K), and enforced by an external wall velocity ($S_w = 0.5$) are depicted in Figure 3.14. Similar to the previous system with the same wall temperatures, it is seen that also in this system, the only considerable difference between the obtained results from the GSI models and MD simulation is that, in terms of sparsity, the velocity correlation based on the GM model and MD results resemble each other more. It is deduced that in comparison with the Ar-Au system, in the current case introducing various kinds of perturbations, such as the temperature difference between the walls or imposing an external shear flow on the system does not affect the results obtained by MD simulation and GSI models, significantly. This behavior is mainly caused by the essence of relatively weak interaction strength between He gas and Au surface, which leads to rather perfectly elastic reflection of the He atoms from the Au surface. Tables 3.4 and 3.5 present various ACs obtained from the employed GSI models, as well as MD simulations for the He-Au system without and with a temperature gradient between the two walls, respectively. First of all, considering only the MD-obtained results it is observed that the majority of ACs for the He-Au system are one order of magnitude smaller than those obtained for the Ar-Au system for the same system condition. This issue, which has been also reported in Ref [33], confirms further the presence of the weaker interaction potential at the He-Au interface in the comparison with the Ar-Au interface. Furthermore, it was noticed while the calculated directional ACs based on the predicted velocities by both CLL and GM models are in a good agreement with the MD results, the CLL model considerably overpredicts the value of α_E. However, the GM model shows high degree of accuracy in all the computed ACs.

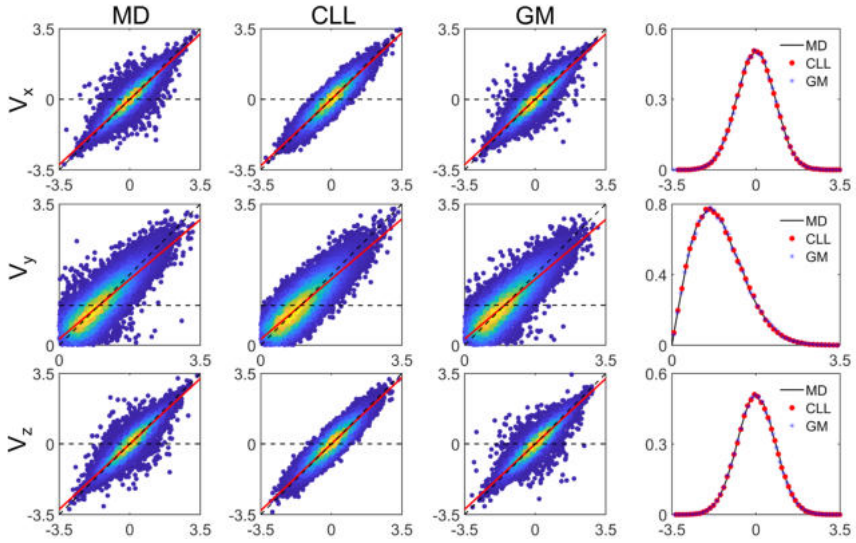

Figure 3.12: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for He-Au system (T_b = 300 K, T_t = 300 K, S_w = 0). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

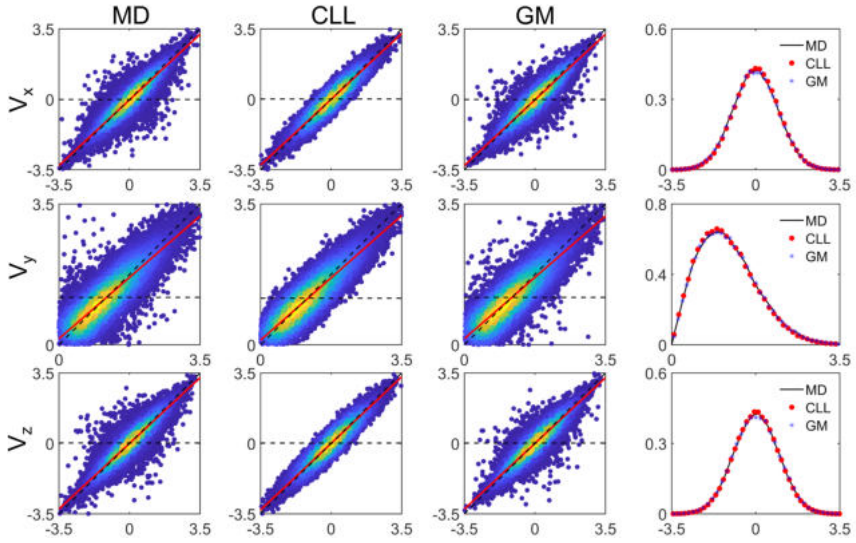

Figure 3.13: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for He-Au system (T_b = 300 K, T_t = 300 K, S_w = 0.5). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

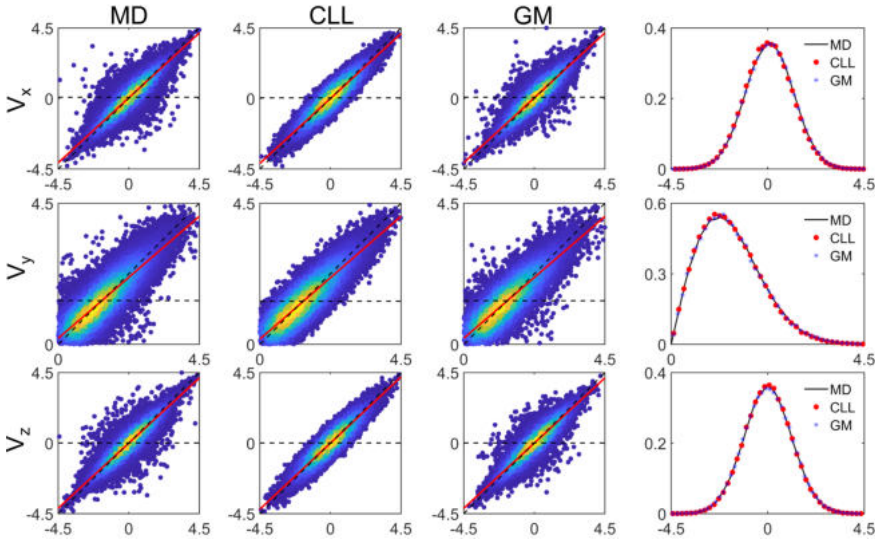

Figure 3.14: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for He-Au system at the top wall (T_b = 300 K, T_t = 500 K, S_w = 0.5). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

Table 3.4: Tangential momentum (α_x,α_z), normal energy (α_{NE}), and energy (α_E) accommodation coefficients for two parallel walls He-Au system with the walls at the same temperature at various speed ratios (S_w), computed using different stochastic approaches: CLL and GM models, and MD simulations

S_w	Model	α_x	α_z	α_{NE}	α_E
0	MD	0.073	0.073	0.132	0.05
	CLL	0.073	0.073	0.131	0.137
	GM	0.074	0.073	0.13	0.052
0.25	MD	0.069	0.0.072	0.126	0.042
	CLL	0.069	0.073	0.125	0.131
	GM	0.0.07	0.0.07	0.131	0.046
0.5	MD	0.067	0.068	0.118	0.041
	CLL	0.067	0.068	0.115	0.123
	GM	0.068	0.068	0.115	0.043

Table 3.5: Tangential momentum (α_x,α_z), normal energy (α_{NE}), and energy (α_E) accommodation coefficients for two parallel walls He-Au system with a temperature gradient between the walls at various speed ratios (S_w),computed using different stochastic approaches: CLL and GM models, and MD simulations. B: bottom wall; T: top wall.

S_w	Model	Wall	α_x	α_z	α_{NE}	α_E
	MD	B	0.072	0.07	0.126	0.045
		T	0.0.079	0.083	0.137	0.046
0	CLL	B	0.07	0.07	0.126	0.131
		T	0.079	0.082	0.133	0.143
	GM	B	0.071	0.072	0.131	0.054
		T	0.079	0.085	0.143	0.046
	MD	B	0.068	0.07	0.122	0.041
		T	0.078	0.08	0.135	0.045
0.25	CLL	B	0.067	0.07	0.121	0.127
		T	0.077	0.08	0.13	0.14
	GM	B	0.066	0.071	0.121	0.042
		T	0.079	0.079	0.131	0.047
	MD	B	0.066	0.067	0.108	0.036
		T	0.079	0.076	0.124	0.039
0.5	CLL	B	0.066	0.067	0.109	0.118
		T	0.079	0.075	0.124	0.135
	GM	B	0.067	0.066	0.109	0.045
		T	0.079	0.075	0.126	0.041

3.4. CONCLUSIONS

In this study, the GM model, an unsupervised machine learning technique, is employed to investigate the interfacial interactions between two noble gases (Ar,He) with Au solid surface. The main advantage of the scattering kernel constructed based on the GM model over the existing stochastic empirical scattering kernels is that its capability is not restricted by the finite number of adjustable parameters, which are required to be known in advance. The GM model uses a superposition of multivariate normal distributions to derive the probability density of a high-dimensional data space. Here the whole collisional data obtained from MD simulation is used for training the model. This guarantees that the obtained model preserves all the important physics in the MD collisional data. The performance of the model in the case of a thermally non-equilibrium, as well as a system with moving walls, which are encountered in various applications involving dilute gas, was evaluated. The assessment has been performed in terms of both statistical and physical characteristics of the system, such as the velocity correlation between incoming and outgoing velocities and the PDF of the outgoing velocities, as well as different accommodation coefficients obtained using the predicted postcollisional velocities by the model. Comparing the results obtained from the GM and CLL models, with the MD results showed the excellent performance of the GM model. Therefore, in the case of monoatomic gases, this model, as a generalized form of the collisional model, can be coupled to the coarse-grained simulation techniques, such DSMC or LBM, and removes the need for the computationally demanding MD simulations. To derive such a generalized scattering kernel, a more extended data set including wider range of the wall temperatures and velocities, as well as different values for gas density is required, and this is in our plan for future studies. It is noteworthy to mention that the high accuracy of the predicted post-collisional velocities by the GM model makes it also a promising candidate for computing various important physical parameters at the gas-wall interface, such as ACs, shear stress, and thermal conductance in the system with highly non-equilibrium gas flow.

4

DEVELOPMENT OF A SCATTERING MODEL FOR DIATOMIC GAS- SOLID SURFACE INTERACTIONS BASED ON GAUSSIAN MIXTURE APPROACH

This work proposes a new stochastic gas-solid scattering model for diatomic gas molecules constructed based on the collisional data obtained from Molecular Dynamics (MD) simulations. The Gaussian Mixture (GM) approach, which is an unsupervised machine learning approach, is applied to H_2 and N_2 gases interacting with Ni surfaces in a two parallel walls system under rarefied conditions. The main advantage of this approach is that the entire translational and rotational velocity components of the gas molecules before and after colliding with the surface can be utilized for training the GM model. This creates the possibility to study also highly nonequilibrium systems, and accurately capture the energy exchange between the different molecular modes that cannot be captured by the classical scattering kernels. Considering the MD results as the reference solutions, the performance of the GM-driven scattering model is assessed in comparison with the Cercignani-Lampis-Lord (CLL) scattering model in different benchmarking systems: the Fourier thermal problem, the Couette flow problem, and a combined Fourier-Couette flow problem. This assessment is performed in terms of the distribution of the velocity components and energy modes, as well as accommodation coefficients. It is shown that the predicted results by the GM model are in better agreement with the original MD data. Especially, for H_2 gas the GM model outperforms the CLL model. The results for N_2 molecules are relatively less affected by changing the thermal and flow properties of the system which is caused by the presence of a stronger adsorption layer.

This chapter is accepted as a peer-reviewed article. Mohammad Nejad, S., Nedea, S., Frijns, A. and Smeulders, D. Development of a scattering model for diatomic gas-solid surface interactions by an unsupervised machine learning approach. *Physics of Fluids, 34(11), 117122, 2022*

4.1. INTRODUCTION

Rarefied gas dynamics or the methods of discrete molecular gas dynamics are applied in various cutting edge technologies, such as aerodynamics of the hypersonic vehicles, semiconductor industry, micro-electro-mechanical systems, and micro/nanoporous media [1]. In the aforementioned applications usually the mean free path (λ) of the gas molecules and the characteristic dimension (d) of the system are in the same order of magnitude. This implies that knowing the details of discrete molecular behavior becomes of crucial importance in understanding flow physics. By increasing the degree of rarefaction, quantified by means of the Knudsen number ($Kn = \frac{\lambda}{d}$), interactions between gas molecules and the adjacent solid surfaces become more dominant than intermolecular gas-gas collisions in determining the macroscopic gas flow characteristics [4, 117]; Accordingly, it is well established that for $Kn > 0.1$ gas experiences significant nonequilibrium phenomena, such as velocity slip and temperature jump at the vicinity of the surface in the region known as the Knudsen layer. As a result, the continuum Navier-Stokes equations are not applicable anymore to describe flow properties [117, 118]. Particle-based numerical techniques, such as direct simulation Monte Carlo (DSMC) method [11] and lattice Boltzmann method (LBM) [10] are being extensively employed for simulating rarefied gas flows. In addition to gas-gas molecular collision models, the successful application of these numerical techniques depends upon the development of accurate gas-surface interaction (GSI) models. However, despite devoting considerable efforts over the last century towards developing realistic GSI models [15–18, 21, 23–25, 119, 120], the physics of GSI phenomena are still not well understood, and reliable and generalized GSI models to describe complex surface interactions and highly nonequilibrium flow in modern engineering applications are still lacking.

In accordance with the probabilistic essence of the particle-based numerical approaches, all commonly used GSI models have also been presented in probabilistic forms describing the relationship between the state of gas molecules before and after colliding with the surface [4]. Generally, according to these models, it can be assumed that the state of scattered molecules falls between two extreme situations: purely specular and purely diffuse reflections. The proportion of diffuse scattering is determined by the so-called accommodation coefficient (AC), which is a parameter to quantify the amount of momentum or thermal energy exchange between gas and solid. The Maxwell model [15] is the simplest empirical GSI model, in which only the tangential momentum AC is used in the model. It assumes that a fixed amount of gas molecules undergoes a fully diffuse reflection, while the remaining part reflects specularly. However, Cercignani and Lampis [16] observed that the Maxwell model was not able to capture the lobular patterns seen in the molecular beam experiment. To tackle this deficiency, they used the normal energy AC alongside the tangential momentum AC, and established a more realistic and mathematically more robust GSI model that better matched with experimental data. The so-called Cercignani–Lampis (CL) model was extended by Lord [119], who proposed the well-known Cercignani–Lampis–Lord (CLL) model and applied it into the DSMC algorithm. Later, Lord extended the CL model even further to describe the scattering of diatomic molecules from a solid surface [17].

Molecular Dynamics (MD) simulations is considered as a powerful computational tool to study gas-solid interactions at the atomistic level. In fact, using MD simulations it

is possible to directly track individual molecules under a wide range of conditions, and gain a detailed understanding of momentum and energy exchange mechanism at a gas-solid interface [49,65,121,122]. However, even with the help of fast computers it becomes very difficult to use a full MD approach for modelling gas-solid interactions in a physical system due to the huge computational costs.

Instead, recent approaches use the insight gained by the MD simulations to construct more elaborated GSI models inspired by the aforementioned empirical scattering models. Nevertheless, the performance of these models relies significantly on the values of the ACs used in these models, which act as calibration parameters. For instance, Yamamoto et al. [21] extended the Maxwell GSI model based on the MD simulations results for N_2-Pt system. Assuming no dependency between translational velocity components, as well as the translational and rotational energy modes, they proposed using the tangential and normal momentum , and the translational and rotational energy ACs to reproduce corresponding velocity and energy components. The main drawback of the CLL model [17] is that the energy exchange between internal energy modes (e.g. the rotational energy mode) and the translational energy mode cannot be described by the model. To overcome this deficiency, Gorji and Jenny [24] extended the CLL model and proposed a new scattering model that could account for energy transfer between the internal and the translational energy modes. Based on the data obtained form the molecular beam experiment, they showed that energy exchange between different energy modes can happen at highly nonequilibrium situations. Using MD simulations data for calibration, they proposed a model with a relatively complex mathematical form, which resulted from coupling a fully diffuse and the CLL model, and depends on five different ACs.

In general, the accuracy of the empirical GSI models and their derivatives highly depends on the values of ACs, and this issue has been addressed as the main disadvantage of these models [41]. In fact, numerous parameters such as gas/solid materials, surface cleanness and roughness, as well as gas/surface temperature can influence the obtained results for the ACs. Therefore, to measure the ACs, accurately, all these parameters need to be controlled properly at atomistic level. This task is considered highly challenging either in an experimental study or MD simulation [38, 55, 123].

In the case of rarefied gas flow systems with complex flow conditions that are far from thermal equilibrium a vast range of physical phenomena could happen at the gas-solid interface that cannot be fully illustrated using a limited number of constant parameters. To overcome such deficiency a new family of scattering kernels known as nonparametric scattering kernels has been introduced recently [31, 32, 42]. Generally, in these works as the first step based on the molecular beam setup, MD simulations have been carried out to study the interactions between a specific gas-solid combination. Afterwards, based on the collisional data gathered from the MD simulations and without using any ACs as an intermediate parameter a specific probability density function (PDF) estimation approach was employed to drive the probabilistic form of the scattering kernel.

Artificial intelligence and machine learning are other promising tools than can be employed to directly construct a probabilistic GSI model from the MD simulations data [33, 34, 44]. As an example the Gaussian Mixture (GM) approach, which is an unsupervised machine learning method was used by Liao et al. [33] to construct a scattering

kernel for monoatomic gases based on the MD data obtained from the molecular beam simulation. In our previous work [34], due to the crucial importance of the gas absorbed layer in the early transition regime (0.1 < Kn <1) on the energy and momentum transfer at gas/solid interface, instead of the molecular beam approach, we used a two parallel walls MD setup as a reference system. We employed the GM model to construct a GSI model for monoatomic gases. Comparison of the obtained results from the GM model against the CLL scattering model in different physical conditions proved the superiority of the GM model. Wu et al. [44] introduced a GM model for N_2-Pt interactions using a molecular beam approach. Neglecting gas-gas interactions, they studied the effect of wall roughness on ACs and reflected velocity distributions. To the best of our knowledge, there is no GM-based scattering kernel for diatomic gases that, including gas-gas interactions, can deal with adsorption-related problems.

In this study, the GM model is used to construct a GSI model for diatomic gas molecules interactions with a given surface. Here, H_2 and N_2 gases are considered as the case studies, due to their numerous engineering applications, and very different molecular masses. At the first step, MD simulations have been carried out for H_2 and N_2 gases confined between Ni walls. Afterwards, the pre and postcollisional translational and rotational velocities of the gas molecules are utilized for the training the GM model. Doing so, the interplay between different velocity components, as well as the energy exchange between the translational and rotational energy modes that are likely to occur in nonequilibrium situations are implicitly taken into account in the model. The performance of the GM based scattering model is examined against the CLL model in some important benchmark thermal problems such as the Fourier thermal problem, the Couette flow problem, and the combined Fourier-Couette flow problem. The examination is carried out in terms of various statistical and physical parameters.

4.2. METHODOLOGIES

4.2.1. MD SIMULATION MODEL

The MD simulation setup used in this work consisting of Ni walls in contact with H_2 or N_2 gas, is presented in Figure 4.1. Each Ni wall is formed by five layers of FCC planes of solid Ni with a cross sectional area of 10.8 nm by 10.8 nm. The distance between the bottom and the top walls, d, is 30 nm. The outermost layer in each wall is kept fixed in order to prevent the translational movement of the walls in the y-direction. Diatomic gas molecules are considered as rigid rotors with a fixed bond length of 0.7414 Å and 1.097 Å for H_2 and N_2, respectively [124]. The total number of gas molecules confined between the two Ni walls is decided in such a way that the mean free path, λ, of gas molecules remains around 10.5 nm, which results in the Knudsen number of $Kn = 0.35$ for both systems.

Periodic boundary conditions are applied along x and z directions. The embedded atom model (EAM) potential [125] is employed to model the interactions between Ni atoms. The non-bonded interactions between gas-gas and gas-wall atoms are modeled using COMPASS force field [126], in which a Lennard-Jones (L-J) 9-6 function is employed to describe the interactions. For two atoms of the same kind the potential parameters

$(\epsilon_{ii}, \sigma_{ii})$ can be found in the COMPASS database [127]. However, for a mixture of two different atoms the potential parameters are calculated using the sixth-power Waldman-Hagler combining rules [128]:

$$\epsilon_{ij} = 2\sqrt{\epsilon_{ii}\epsilon_{jj}}\left[\frac{\sigma_{ii}^3.\sigma_{jj}^3}{\sigma_{ii}^6+\sigma_{jj}^6}\right] \tag{4.1}$$

$$\sigma_{ij} = \left[\frac{\sigma_{ii}^6+\sigma_{jj}^6}{2}\right]^{\frac{1}{6}} \tag{4.2}$$

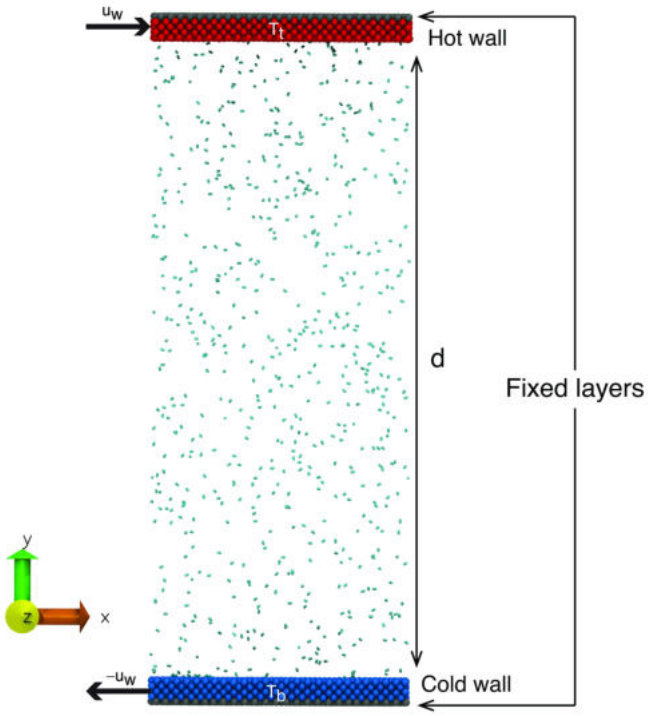

Figure 4.1: Schematic representation of the simulation model; d: distance between the two walls; T_b: Temperature of the bottom wall; T_t: Temperature of the bottom wall; u_w: Prescribed velocity on the walls.

All the gas-gas and gas-wall interatomic potential parameters used in this study are listed in Table 4.1. The cutoff distance for gas-gas and gas-wall interactions are set at $2.5\sigma_{ii}$ and 10 Å, respectively.
In the first step, for each MD simulation setup, the energy minimization is performed by iteratively adjusting atom positions. In the next step, each wall is equilibrated at the desired temperature level using Nose-Hoover thermostat [48] (NVT), while gas molecules

are modeled in the microcanonical ensemble (NVE), and their temperature can only change by exchanging kinetic energy with other molecules in the simulation box via collisions. However, in order to accelerate the equilibration process, the velocity components of gas molecules are initially sampled from a Gaussian distribution with a mean value of 0.0 and a standard deviation of $\sqrt{\frac{k_B T_a}{2m_g}}$, where T_a is the average value of the bottom and top wall temperatures, and m_g is the mass of gas molecule.
To model the Couette flow condition, as it is depicted in Figure 4.1, the top wall moves with the velocity u_w, whereas the bottom wall has the velocity $-u_w$. Herein, the speed ratio S_v is assigned as

$$S_v = \frac{u_w}{\sqrt{2k_B T_b/m_g}} \tag{4.3}$$

where T_b is the temperature of the bottom wall. In order to attain the desired λ and Kn in the systems, all simulation setups are equillibrated for 3 ns to reach the required temperature and pressure with a time step of 0.5 and 1.0 fs for H_2-Ni and N_2-Ni systems, respectively. Afterwards, the production run is followed, which is carried on for 25 ns for H_2-Ni system and 60 ns for N_2-Ni system. All MD simulations are conducted using LAMMPS [129] package.

Table 4.1: Lennard–Jones potential parameters

Atom pair	ϵ (eV)	σ (Å)
Ni-Ni [130]	$6.6e^{-1}$	2.239
H-H	$9.29e^{-4}$	1.421
H-Ni	$1.19e^{-2}$	2.016
N-N	$2.57e^{-3}$	3.8
N-Ni	$1.61e^{-2}$	3.408

4.2.2. SCATTERING KERNEL

In the transitional flow regime ($0.1 < \mathrm{Kn} < 10$) the Boltzmann equation must be employed as the governing equation to solve the problems of gas flows. Having a precise and detailed understating of boundary conditions is crucial in order to guarantee the reliability of the simulation results, while considering flows of gas around bodies or adjacent to a solid surface. The boundary condition defines the relation between the incoming and outgoing gas molecular velocity distributions, $f(\boldsymbol{v'})$ and $f(\boldsymbol{v})$, at the boundary surface. It can be represented generally through the scattering kernel $R(\boldsymbol{v}|\boldsymbol{v'})$ as

$$v_n f(\boldsymbol{v}) = \int_{v'_n<0} |v'_n| R(\boldsymbol{v}|\boldsymbol{v'}) f(\boldsymbol{v'}) d\boldsymbol{v'},\ v_n > 0 \tag{4.4}$$

Here, $\boldsymbol{v'}$ and $\boldsymbol{v}$ are the translational velocities of the center of mass (COM) for the incoming and outgoing molecules, respectively. The velocities v_n and v'_n are the corresponding normal components of the molecular velocities directed into the gas domain and normal to the surface. The scattering kernel $R(\boldsymbol{v}|\boldsymbol{v'})$ denotes the probability that a stream

of incoming gas molecules with velocity in $[\boldsymbol{v'}, \boldsymbol{v'} + d\boldsymbol{v'}]$ will be bounced off with a new velocity in the interval of $[\boldsymbol{v}, \boldsymbol{v} + d\boldsymbol{v}]$. However, in the case of a diatomic gas molecule, in addition to the translational velocity components, the rotational velocity vectors $(\boldsymbol{\omega'}, \boldsymbol{\omega})$ need to be taken into account, as well. Therefore, the scattering kernel must be substituted by $R(\boldsymbol{v}, \boldsymbol{\omega} | \boldsymbol{v'}, \boldsymbol{\omega'})$ and the probability density must be substituted by $f(\boldsymbol{v}, \boldsymbol{\omega})$. The CLL model is the most accurate empirical scattering kernel that commonly used in diatomic rarefied gas flow simulations, and it is given in the following form [17]

$$
\begin{aligned}
R^{CLL}(\boldsymbol{v}, \boldsymbol{\omega} | \boldsymbol{v'}, \boldsymbol{\omega'}) = \frac{2 v_n}{\pi^2 \alpha_t (2-\alpha_t) \alpha_n \alpha_{ro}} \exp\left[-\frac{[\boldsymbol{v_t} - \sqrt{1-\alpha_t}\boldsymbol{v'_t}]^2}{\alpha_t (2-\alpha_t)} \right] \\
\times \exp\left[-\frac{v_n^2 + (1-\alpha_n) {v'_n}^2}{\alpha_n} \right] I_0 \left[\frac{2(\sqrt{1-\alpha_n}) v_n v'_n}{\alpha_n} \right] \\
\times \exp\left[-\frac{(\boldsymbol{\omega} - \sqrt{1-\alpha_{rot}}\boldsymbol{\omega'})^2}{\alpha_{rot}} \right]
\end{aligned}
\tag{4.5}
$$

where α_t, α_n, and α_{rot} are the accommodation coefficients corresponding to the tangential momentum, normal translational kinetic energy, and rotational energy, respectively. $\boldsymbol{v_t}$ represents the tangential velocity vector and I_0 is the modified Bessel function of the first kind and zeroth order. The translational $(\boldsymbol{v'}, \boldsymbol{v})$ and rotational $(\boldsymbol{\omega'}, \boldsymbol{\omega})$ velocity vectors are normalized by $\sqrt{\frac{2k_B T_w}{m_g}}$ and $\sqrt{\frac{2k_B T_w}{I}}$ respectively. Here, T_w describes the wall temperature and I is the mass moment of inertia of the diatomic gas molecule. The algorithm and formulas used in this work to generate post-collisional velocities according to the CLL scattering kernel are given in Ref. [24]. The required accommodation coefficients in Equation 4.5 are calculated using the proposed method in Ref. [49], which is based on the least-squares approximation on the collisional data containing impinging Φ_I and reflected Φ_R quantities

$$
\alpha_\phi = 1 - \frac{\sum_i (\Phi_I^i - \langle \Phi_I \rangle)(\Phi_R^i - \langle \Phi_R \rangle)}{\sum_i (\Phi_I^i - \langle \Phi_I \rangle)^2} \tag{4.6}
$$

where ϕ can be related to the different kinetic properties of the gas molecules such as its center of mass velocity in a certain direction, as well as its translational $E_{tr} = \frac{1}{2} m_g \boldsymbol{v}^2$, rotational $E_{rot} = \frac{1}{2} I \boldsymbol{\omega}^2$ or total $E_{tot} = E_{tr} + E_{rot}$ kinetic energy. Collisions are tracked by defining a virtual plane at a distance of one gas-wall interaction cutoff radius away from the walls ($r_{plane} = r_{cutoff} = 10\,\text{Å}$). Collisional data are recorded when the geometric center of mass of the molecule crosses the virtual plane. The final data sets gathered from MD simulations of different systems include between 100,000 to 150,000 data points.

4.2.3. GAUSSIAN MIXTURE MODEL

Gaussian mixture (GM) model is a probabilistic clustering approach for illustrating normally distributed subset within an overall dataset. The GM model typically does not require knowing from which subset a data point comes, and the model learns the subsets automatically. Having this feature, the GM model is identified as an unsupervised machine learning technique. Due to its robustness and flexibility, the GM model has

been utilized in a wide variety of important practical situations such as language identification, anomaly detection, pattern recognition, signal processing, and object tracking of multiple objects [131–134].
In the GM model a superposition of M-dimensional Gaussian functions is employed to estimate the probability density function of data $\boldsymbol{X}$ as:

$$p(\boldsymbol{X}) = \sum_{i=1}^{K} \rho_i N\left(\boldsymbol{X}|\vec{\mu}_i, \Sigma_i\right) \tag{4.7}$$

where K is the number of Gaussians. $\rho_i, i = 1,2,...,K$ are the mixture component weights with the constraint that $\sum_{i=1}^{K} \rho_i = 1$, $\vec{\mu}_i$ is the mean vector, and Σ_i is the covariance matrix. $N\left(\boldsymbol{X}|\vec{\mu}_i, \Sigma_i\right)$ are the components density functions, which are indeed M-dimensional normal distributions given by

$$N\left(\boldsymbol{X}|\vec{\mu}_i, \Sigma_i\right) = \frac{1}{(2\pi)^{M/2}|\Sigma_i|^{1/2}} exp\left[-\frac{1}{2}(\boldsymbol{X}-\vec{\mu}_i)'\Sigma_i^{-1}(\boldsymbol{X}-\vec{\mu}_i)\right] \tag{4.8}$$

The model parameters ($\psi = \{\rho_i, \vec{\mu}_i, \Sigma_i\}\ \forall\ i$ in $\{1\cdots K\}$) are determined using the expectation-maximization (EM) optimization algorithm [135]. However, before implementing the EM algorithm to estimate the GM model parameters, it is essential to define the model attributes such as the type of covariance matrix [33] (e.g. spherical, diagonal, tied, or full) and the number of Gaussians. In this work, the full covariance matrix has been applied that has the best performance, and it is taken as the default covariance matrix in the SCIKIT-LEARN [115] package used to implement the GM model. The number of Gaussians is a crucial parameter that is required to be specified adequately by the user to avoid underfitting or overfitting in the model, and it directly affects the accuracy and the computational cost of the model. To determine the optimal K, a sensitivity analysis has been carried out. It is discussed in detail in the Appendix D. From this analysis, in this work $K = 500$ and $K = 450$ are used as the number of Gaussians for H_2-Ni and N_2-Ni systems, respectively.
The collisional data used to train the GM model is a 10-dimensional dataset including impinging and outgoing translational velocities of the COM of the molecules (v_x', v_y', v_z', v_x, v_y, v_z), as well as impinging and outgoing rotational velocities of the molecules (ω_1', ω_2', ω_1, ω_2). Including both translational and rotational degrees of freedom for training purpose indicates that the possible energy transfer between translational and rotational modes in the case of nonequilibrium condition is taken into account in this model. From each dataset obtained from a specific MD simulation, 75% of the trajectories (consisting of translational and rotational velocities), along with a desired number of Gaussians, are fed into the GM model for training. The remaining data is used for validation purpose.
Here, except the normal velocity components (v_y', v_y) that follow a Rayleigh distribution, the other velocity components follow a Gaussian distribution. As it has been comprehensively discussed in our previous work [34], the GM model performs better when all the components of the original MD data have Gaussian distributions. Therefore, exploiting the preprocessing scheme proposed by Liao et al. [33], the normal velocity components have been transferred from Rayleigh to Gaussian distributions. Firstly, for each normal velocity pair (v_y', v_y), its equivalent inverse ($-v_y'$, $-v_y$) is added to the dataset.

Secondly, using the following scheme (Equation. 4.9) the distributions of the normal velocity components are transferred from Rayleigh to Gaussian:

$$T(\theta) = \sqrt{2\beta}\, erf^{-1}\left[1 - 2\exp\left(-\frac{\theta^2}{2\beta}\right)\right], \beta = \frac{k_B T_g}{m_g} \tag{4.9}$$

where T_g refers to the gas temperature and can be computed using the average translational kinetic energy of the gas. The first step of preprocessing scheme doubles the size of velocity data in the normal direction. Therefore, half of the final obtained data is chosen and used for training alongside with the other components in the dataset. In accordance with the distribution of the final data used for training the GM model, the obtained velocities from the model also have Gaussian distribution. Hence, to compare the model predictions with the initial MD results, the predicted velocity components in the normal direction are transformed back into the Rayleigh distribution as follows:

$$Z(\theta) = \sqrt{-2\beta \ln\left[\frac{1}{2} - \frac{1}{2}\, erf\left(\frac{\theta}{\sqrt{2\beta}}\right)\right]} \tag{4.10}$$

4.3. RESULTS AND DISCUSSION

This section is divided into three main parts. The first part is devoted to validate the MD simulations approach used in this work to study the diatomic gas-solid surface interactions. In the second part the performance of the GM and CLL scattering models in three thermal problems, commonly encountered in rarefied gas flow systems, are investigated. Finally, in the last part based on the density profiles extracted from the MD simulations, the reasons for different behavior of H_2 and N_2 gases in terms of thermal behaviors and flow properties are discussed.

4.3.1. MD SCHEME VALIDATION

Using experimental results, our aim is to validate first the force fields used in the simulations and secondly to prove the suitable methods to extract ACs from MD data. Here, the tangential momentum accommodation coefficient (α_x), and the total thermal energy accommodation coefficient (α_{tot}) for H_2-Ni and N_2-Ni systems obtained from the MD simulations are compared with the corresponding experimental values in the literature. For this, we replicated the experimental conditions using a two parallel plates setup [136] ($T_b = 288$ K, $T_t = 308$ K, $S_v = 0$) in the MD simulations.

From MD simulations for H_2-Ni system the values of α_x and α_{tot} computed on the bottom wall were 0.953 and 0.316, respectively. With regard to α_x, no measurement on Ni surface was found in the literature. Knowing that surface material has minor impact on the value of α_x [39], the results on the other metal surfaces have been used for the comparison purpose. Herein, the reported experimental value of α_x for H_2 on Bronze surface is 0.94 [55], which is consistent with the value computed in this work. Besides that, for H_2-Ni an empirical study showed that $\alpha_{tot} = 0.293$ [137], which is in good agreement with the MD obtained value.

In the case of N_2-Ni, the results obtained from MD simulations are $\alpha_x = 0.87$ and $\alpha_{tot} =$

0.869. The reported values for α_x on different surfaces are in the range of $0.84 < \alpha_x < 1$ [39]. Referring to α_{tot}, the value reported by Amdur et al. [137] for N_2-Ni is 0.823, which is consistent with the value computed in this work. Based on the presented results it can be deduced that the choice of the intermolecular potentials and the atomic schemes to study GSIs in this work are quite reliable for studying further the three thermal problems we want to investigate.

4.3.2. GM MODEL ASSESSMENT

Using the MD simulations results for H_2-Ni and N_2-Ni as the reference solutions, we aim to assess the performance of the GM and CLL scattering models in predicting the gas velocities after the collision with the solid walls.

Three different benchmark problems, namely the Fourier thermal problem, the Couette flow problem, and the combined Fourier-Couette flow problem are examined. In the case of the Fourier thermal problem, the bottom wall has the temperature T_b =300 K, while the temperature of the top wall is T_t =300 K (named the isothermal walls system) or 500 K (named non-isothermal walls system). For the Couette flow problem, the considered velocity ratios are S_v =0.2 or 0.4, and both walls have the same temperatures $T_b = T_t$ =300 K. For the combined Fourier-Couette flow problem, using the non-isothermal wall system as the initial system, different velocity ratios (S_v =0.2 or 0.4) are imposed on the walls. The training of the GM model in each case study is carried out separately. For each case study the most significant results are addressed in the main text, while the remaining ones for H_2-Ni and N_2-Ni can be found in Appendix E and Appendix F, respectively.

To generate the outgoing velocities based on the CLL model, first of all, the required ACs including TMAC in x and z-directions (α_x,α_z), α_n, and α_{rot} are computed using the collected MD collisional data from each specific case study. Afterwards, the post-collisional velocities are generated by using the computed ACs according to the method described in Ref. [24]. In the case of the systems with moving walls, the value of the imposed velocity at the wall (u_w) is added to the all velocity components generated by the CLL model in the direction of wall movement ($V_{x,final} = V_{x,CLL} + u_W$).

The evaluation of the aforementioned stochastic GSI models has been carried out in terms of the correlation between the incoming and outgoing translational velocity components of COM (V_x, V_y, V_z), different energy modes of the gas molecules (E_{tr}, E_{rot}, E_{tot}), as well as the PDF of the outgoing velocity components or the energy modes. To compare the resulted PDFs in a quantitative manner, the Kullback-Leibler divergence (KLD) [138] also known as information divergence or relative entropy is employed. KLD, as the most commonly used information criterion for evaluating the model discrepancy, is being used frequently in machine learning studies to measure the difference between the actual and observed or predicted probability distributions. Suppose $p(x)$ and $q(x)$ are two PDFs on the same probability space χ, the KLD is defined as

$$KLD(p(x)||q(x)) = \sum_{x\epsilon\chi} p(x) ln \frac{p(x)}{q(x)} dx, \tag{4.11}$$

where x can be referred to the velocity or energy data obtained from either MD simulations or GSI models. To calculate $p(x)$ and $q(x)$, the related data space is evenly seg-

mented into n bins. The density in each bin is computed by counting the number of samples in the bin. In addition to the correlations and PDFs, the corresponding ACs for the aforementioned kinetic features of the gas molecules have been used for the assessment purpose.
All the aforementioned properties are tested because, as it will be shown in the following section, in some case studies even if PDFs or ACs are in agreement, velocity correlations can be different affecting the heat transfer predictions.

4.3.2.1. *H_2-Ni system*
The velocity and energy distributions determined by the atomistic simulations and different GSI models for the isothermal Fourier thermal problem are plotted in Figure 4.2. It is seen that in this case study, which resembles the fully equilibrium situation in the system, there is a perfect match between the correlation plots and PDFs of the partial translational velocity components, and the rotational energy mode. However, E_{tr} clouds of the MD simulation and the GM model are slightly narrower than the CLL model. Since $E_{tot} = E_{tr} + E_{rot}$, this mismatch is propagated into the E_{tot} cloud, as well. This issue can be also deduced from the values of the ACs in Table 4.2, in which various ACs of the Fourier thermal problem for the H_2-Ni system are presented. It is observed that there is a perfect match between the values of α_x, α_y, α_z, and α_{rot} obtained from the GM and CLL scattering models with the MD simulations, while the values of α_{tr} and α_{tot} obtained from the wall scattering models are larger than the values obtained directly from the MD simulations. However, in comparison with CLL results, the values obtained from the GM scattering model are much closer to the MD results.

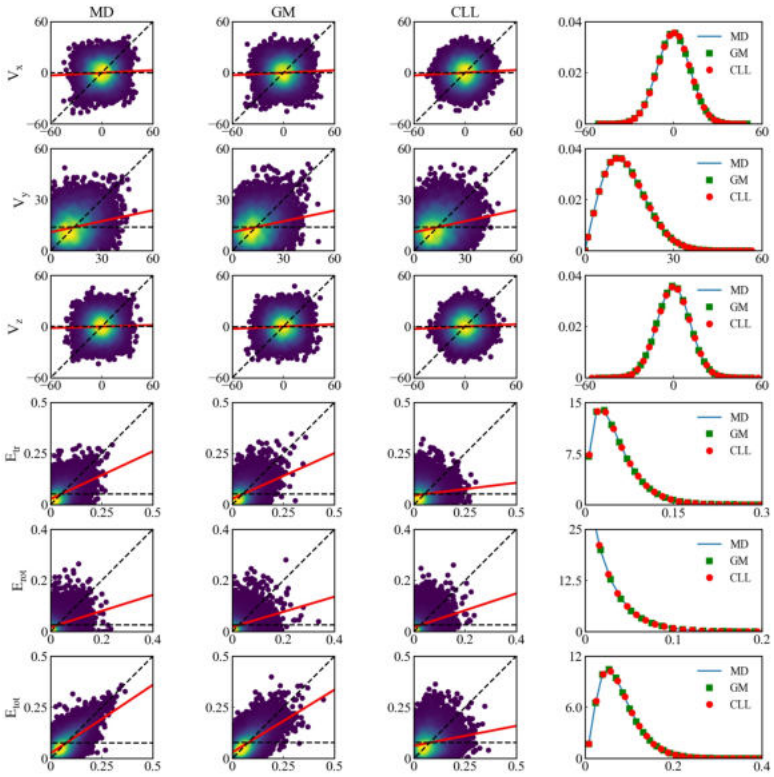

Figure 4.2: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the isothermal Fourier thermal problem for H_2-Ni system at the bottom wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of translational velocity components and energy modes for the reflecting particles are presented

The correlation plots and the PDFs at the bottom wall of the non-isothermal Fourier thermal system are shown in Figure 4.3. While the correlation plots acquired form the GSI models for the translational velocity components are in good agreement with the MD simulations results, a deviation is observed around the peak values in the PDFs obtained from the CLL model(e.g. see the PDF of E_{tr} in Figure 4.3). Since the peak value in such distributions provides a measure of temperature, predicting higher peaks by the CLL model implies that gas molecules accommodate more to the surface, and this leads to a lower temperature for the reflected gas molecules. The predicted outgoing translational velocities by the statistical GSI models, and the original MD results were employed to compute the temperature of the gas after colliding with the surface. The post-collisional gas temperature according to the MD results and the GM model were 377.51 and 374.23 K, respectively. However, the CLL model gave the outgoing temperature of 315.1 K. This underprediction of the temperature confirms the aforementioned statement. With regard to the various energy modes of the gas molecules, the results from the

GM model are always in a reasonable agreement with the MD simulation results. However, the correlation plots and the PDFs obtained based on the CLL model for E_{tr} and E_{tot} deviate from the MD results, which predict the reflection in a more diffuse manner according to the CLL model. On the contrary, the predicted E_{rot} by the CLL model matches well with the MD data, although looking to the correlation plots, the MD data are more dispersed than the CLL scattering model results. Based on the predicted translational velocity components and the energy modes by the GM and CLL GSI models, the KLD coefficient between these models and the original MD results were computed (see Figure 4.4a). It is seen that the deviation between the CLL model and the MD results are 2 orders of magnitudes larger than the deviation between the GM model and MD results. By looking at the reported ACs of this case study (see Table 4.2) it is inferred that similar to the previous case study the results from the GM model in comparison with the CLL model are overall in a better agreement with the MD simulations results. Comparing the values of the ACs in this case study with the isothermal Fourier problem, it is noted that all the computed ACs are slightly lower. This is inline with the previously observed trend of decreasing AC by increased kinetic energy of the gas molecules [37,64] that in this case study is caused by having higher temperature at the top wall.

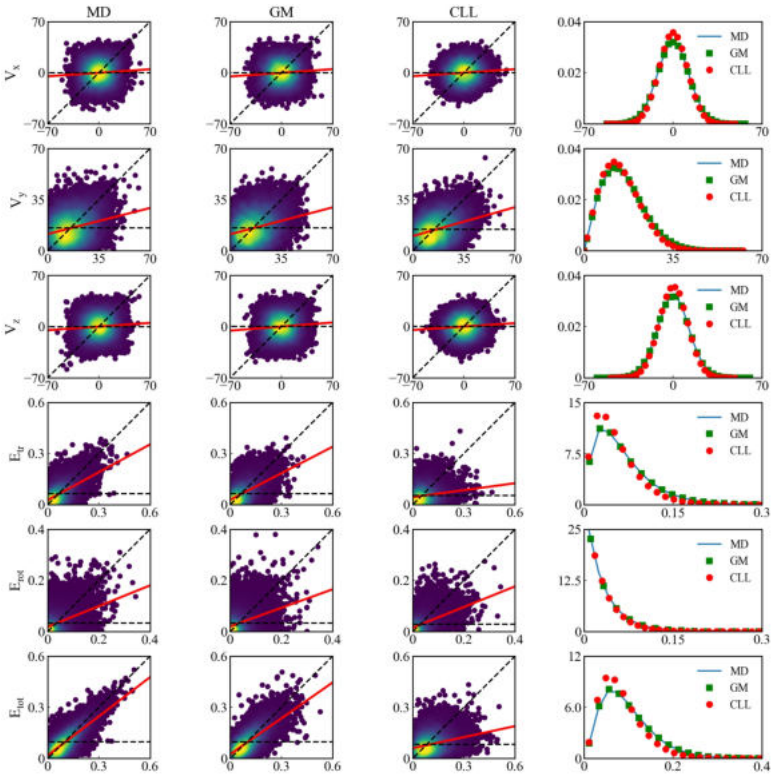

Figure 4.3: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the non-isothermal Fourier thermal problem for H_2-Ni system at the bottom wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

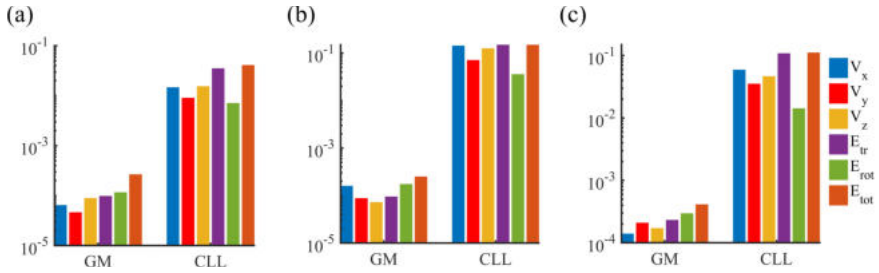

Figure 4.4: The Kullback Leibler divergence of the translational velocity components and different energy modes of H_2-Ni system determined by GM and CLL scattering models; (a) non-isothermal Fourier thermal problem at the bottom wall; (b) Couette flow problem (S_v = 0.4) at the bottom wall; (c) combined Fourier-Couette flow problem (S_v = 0.4) at the top wall.

Table 4.2: Tangential momentum (α_x,α_z), normal momentum (α_y), translational (α_{tr}), rotational (α_{rot}), and total (α_{tot}) energy accommodation coefficients of the Fourier thermal problem for H_2-Ni system (T_b = 300 K), computed using different scattering kernels: GM and CLL models, and MD simulations.B: bottom wall; T: top wall.

T_t	Model	Wall	α_x	α_y	α_z	α_{tr}	α_{rot}	α_{tot}
300	MD	B	0.958	0.797	0.955	0.526	0.666	0.326
	GM	B	0.958	0.798	0.954	0.558	0.700	0.379
	CLL	B	0.959	0.783	0.955	0.880	0.666	0.810
500	MD	B	0.928	0.748	0.820	0.452	0.600	0.234
		T	0.923	0.777	0.924	0.473	0.600	0.274
	GM	B	0.927	0.747	0.933	0.481	0.634	0.296
		T	0.921	0.776	0.922	0.507	0.636	0.340
	CLL	B	0.929	0.713	0.930	0.859	0.601	0.776
		T	0.922	0.774	0.926	0.864	0.595	0.775

Figure 4.5 shows the scattering results based on the MD simulations, and the employed stochastic GSI model for Couette flow problem with S_v =0.4. It is seen that the GM model results are in good agreement with the MD simulation results. However, in the case of the CLL model, except for the rotational energy, the other energy modes (E_{tr},E_{tot}), as well as the partial velocity components (V_x,V_y,V_z) significantly deviate from the MD results. For the tangential velocity components the obtained ACs are similar. This can be understood by comparing the slope of the red lines in the correlation graphs obtained from the MD simulations and the scattering models in the tangential directions, as well as the values of α_x and α_z reported in Table 4.3. However, the shape of the resulted velocity clouds obtained from the CLL model are different from the ones resulted from the original MD data and the GM model. In both tangential directions the CLL results have ellipsoidal shapes, and look symmetrical around the horizontal dashed line. However, MD simulation correlation graphs in the tangential directions are more concentrated around the diagonal lines. In fact, MD results show that in case the magnitude of the incoming velocity is large, there is a strong correlation between incoming and outgoing velocities. This causes that the velocity clouds become narrower in their ending points around the diagonal line at high velocities (see Figure 4.5). Comparison of the tangential velocity clouds from the employed GSI models with the MD results indicates that the aforementioned behaviour can be captured by the GM model, and not by the CLL model.

KLDs based on the obtained scattering results in this case study are shown in Figure 4.4b. Comparing the GM results with the non-isothermal Fourier system (see Figure 4.4a), the most noticeable difference is a higher KLD value for V_x component in the current case study (2.5 times higher), indicating relatively poor performance of the GM model here. However, the deviation of the CLL model and the MD results are much more noticeable in this case study. The ACs obtained from MD simulations, and different GSI models for Couette flow problem are reported in Table 4.3. It is seen that α_x, α_y, and α_z obtained from both scattering models agree well with the MD simulation results. With regard to the energy ACs, α_{rot} obtained from the CLL model has a slightly better match with MD results than the value obtained from the GM. However, while the values of α_{tr} and α_{tot} obtained from CLL are significantly larger than the MD results, the results based on the GM model are in a reasonable agreement with the MD ones. Comparing the ACs presented in Table 4.3 with the ones for the fully equilibrium system in Table 4.2, it can be deduced that imposing a velocity at walls resulted in decreasing the value of all the ACs. In addition, increasing the velocity ratio from S_v =0.2 to 0.4, the value of α_{tot} decreased by half. This inverse correlation was also observed by Uene et al. [139].

Table 4.3: Tangential momentum (α_x,α_z), normal momentum (α_y), translational (α_{tr}), rotational (α_{rot}), and total (α_{tot}) energy accommodation coefficients of the Couette flow problem for H_2-Ni system at different speed ratios (S_v), computed using different scattering kernels: GM and CLL models, and MD simulations

S_v	Model	α_x	α_y	α_z	α_{tr}	α_{rot}	α_{tot}
0.2	MD	0.942	0.743	0.933	0.432	0.610	0.220
	GM	0.940	0.742	0.931	0.466	0.651	0.284
	CLL	0.941	0.713	0.932	0.855	0.615	0.777
0.4	MD	0.854	0.655	0.846	0.307	0.477	0.105
	GM	0.853	0.660	0.844	0.351	0.526	0.183
	CLL	0.854	0.570	0.845	0.809	0.477	0.700

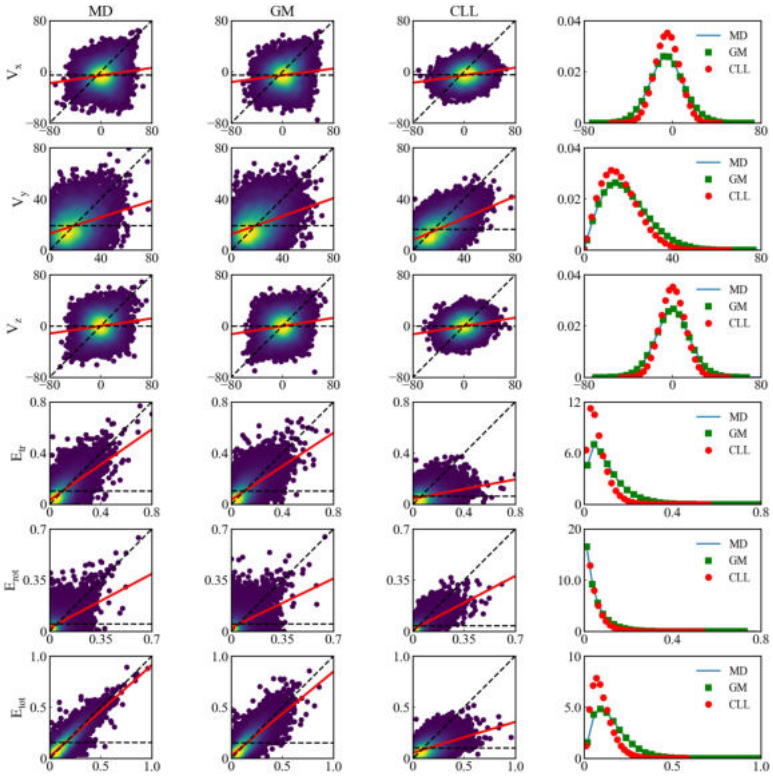

Figure 4.5: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of Couette flow problem (S_v =0.4) for H_2-Ni system at the bottom wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

The scattering plots obtained from the GSI models together with the reference MD solutions for the combined Fourier-Couette flow problem (S_v =0.4) at the top wall of the system are depicted in Figure E.5. Again, it is seen that the results from the GM model are in a good agreement with the MD results. On the other hand, the CLL model except for the rotational energy mode, does not show a good performance in predicting the other energy modes, as well as the translational velocity components. Similar to the previous case study, for the tangential velocity components, the sharper tips of the velocity clouds along the diagonal lines observed in MD data are captured just by the GM model, and can not be seen in the CLL model results.

The computed KLDs for this case are presented in Figure 4.4c. Significant difference between the performance of the employed GSI models has been observed also here. For example, KLDs for E_{tr}, and E_{tot} obtained from CLL are 3 order of magnitudes larger than the corresponding values obtained from the GM model. With regard to V_x, V_y, V_z, and E_{rot}, the computed values based on the CLL model are 2 order of magnitudes larger than GM results.

The ACs associated with different velocity components and energy modes for the combined Fourier-Couette flow system are listed in Table 4.4. It is shown that, similar to the previous case studies, both GSI models have an acceptable accuracy in predicting α_x, α_y, α_z, as well as α_{rot}. However, while the GM model results for α_{tr} and α_{tot} are just slightly higher than the MD results, there is a significant discrepancy between the results obtained from the CLL model and MD simulations. Comparing the results shown in Table 4.4 with those in Tables 4.3 and 4.2, the first observation is that in general the ACs in the combined Fourier-Couette flow case studies are lower. In fact, having the top wall at higher temperature together with imposing a wall velocity transfers considerable amount of kinetic energy to the gas domain, which leads to less accommodation of the gas molecules at the walls surface. The second observation is related to the performance of the GM model across these three case studies. Comparing the values of α_{tr} and α_{tot} obtained from the MD simulation and the GM model, it is seen that as the system goes towards higher non-equilibrium state, the deviation between the GM model and the original MD results increases.

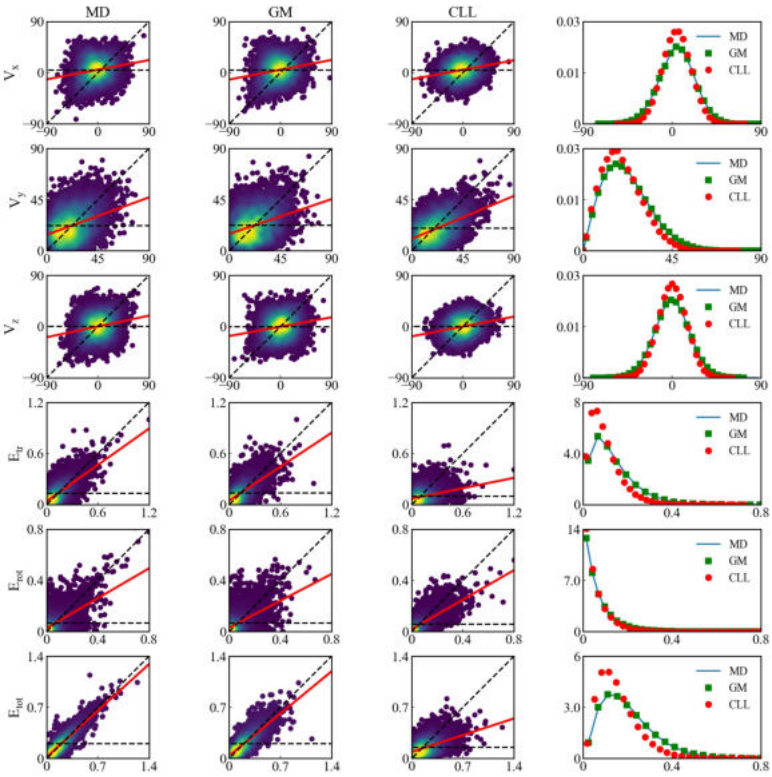

Figure 4.6: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the combined Fourier-Couette flow problem for H_2-Ni system at the top wall ($S_v = 0.4$). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

Table 4.4: Tangential momentum (α_x,α_z), normal momentum (α_y), translational (α_{tr}), rotational (α_{rot}), and total (α_{tot}) energy accommodation coefficients of the combined Fourier-Couette flow problem for H_2-Ni system at different speed ratios (S_v), computed using different scattering kernels: GM and CLL models, and MD simulations.B: bottom wall; T: top wall.

S_v	Model	Wall	α_x	α_y	α_z	α_{tr}	α_{rot}	α_{tot}
0.2	MD	B	0.913	0.705	0.904	0.379	0.544	0.155
		T	0.913	0.726	0.896	0.397	0.556	0.190
	GM	B	0.912	0.709	0.908	0.421	0.593	0.232
		T	0.916	0.726	0.892	0.435	0.590	0.259
	CLL	B	0.913	0.647	0.904	0.831	0.544	0.738
		T	0.908	0.718	0.893	0.838	0.559	0.745
0.4	MD	B	0.812	0.633	0.801	0.274	0.418	0.075
		T	0.808	0.646	0.805	0.286	0.420	0.092
	GM	B	0.807	0.633	0.802	0.314	0.479	0.155
		T	0.813	0.649	0.807	0.328	0.480	0.167
	CLL	B	0.812	0.527	0.800	0.791	0.419	0.667
		T	0.808	0.591	0.804	0.796	0.420	0.670

4.3.2.2. N_2-Ni system

The resulting scattering plots of the non-isothermal Fourier thermal problem for the N_2-Ni system at the top wall are depicted in Figure 4.7. It is seen that the correlation graphs and the PDFs of the reflected translational velocity components, as well as the energy modes obtained from both GSI models match with the atomistic simulation results. The computed KLDs based on the employed scattering models are presented in Figure 4.8a. It is seen that unlike H_2-Ni system, KLDs obtained from both GSI models are in the same order of magnitude. However, the results of the GM model still show less deviation in comparison with the CLL model (e.g. for E_{tot}: KLD_{CLL}= 2.2 KLD_{GM}). Different ACs obtained for this case study are listed in Table 4.5. It is observed that the ACs based on the GM and CLL scattering models are in good agreement with the reference MD results. However, the value of α_{tot} predicted by the GM model is closer to the MD results in comparison with the CLL model. Furthermore, changing the top wall temperature from T_t =300 K to 500 K, so going from the isothermal to the non-isothermal system, did not have any noticeable impact on the values of the ACs. This means that different from H_2-Ni system, here ACs are not dependent on the temperature in the system.

Figure 4.9 represents the scattering plots for the Couette flow problem (S_v =0.4) at the bottom wall. It is shown that all velocity components, as well as the energy modes predicted by the scattering models agree well with the MD results. However, in the x-direction the original MD and GM results are more skewed, while the correlation graph obtained from the CLL model seems very symmetric around the horizontal dashed line. The computed KLDs for this case study (see Figure 4.8b) also confirm the observed discrepancies in the scattering plots. It is shown that the measured KLD for V_x based on the CLL model is the highest one. The impact of such deviation can be seen also in relatively larger values of E_{tr} and E_{tot} obtained from the CLL scattering model. Different ACs for the Couette flow system are listed in Table 4.6. Similar to the previous case study, quite acceptable performance in terms of predicting ACs was observed from the GSI models. Furthermore, based the reported values in Tables 4.5 and 4.6, it can be deduced that neither imposing velocity at walls nor increasing the velocity ratio from S_v =0.2 to 0.4 have any noticeable impact on the values of the ACs.

The correlations and PDFs graphs for the Fourier-Couette flow problem at the bottom wall (S_v =0.4) are depicted in Figure 4.10. From this figure, the most notable difference between the performance of the employed GSI models is seen just in the shape of the clouds in the tangential directions. It is shown that the tips of the V_x and V_z components obtained from MD simulations and the GM model at high velocities are sharper than the CLL results. In addition, the overall shape of the clouds for V_x from MD simulations and GM are not symmetric, while the V_x component obtained from the CLL model is symmetric around the horizontal dashed line. The KLDs for this case study are presented in Figure. 4.8c. Similar to the previous case studies, superior performance of the GM model can be also noticed here. Comparing the KLDs from CLL in this case study with the ones for the Couette flow problem (see Figure 4.8b), it can be inferred that imposing a wall velocity together with having walls with different temperatures induced larger deviation of the CLL predictions from the reference MD results. The calculated ACs for this case are listed in Table 4.7. Again, it is seen that both GSI models can predict ACs with good accuracy. However, the value of α_{tot} obtained from the GM model has a slightly bet-

ter agreement with the MD results as compared with the value obtained from the CLL model. Comparing the results at different velocity ratios, it can be argued that similar to the Couette flow problem, here also increasing the velocity ratio merely affects the values of the ACs, which was different in the case of H_2-Ni system. Looking to the original MD data, the possible reasons for different thermal behaviors of the reflected H_2 and N_2 molecules are discussed in the following section.

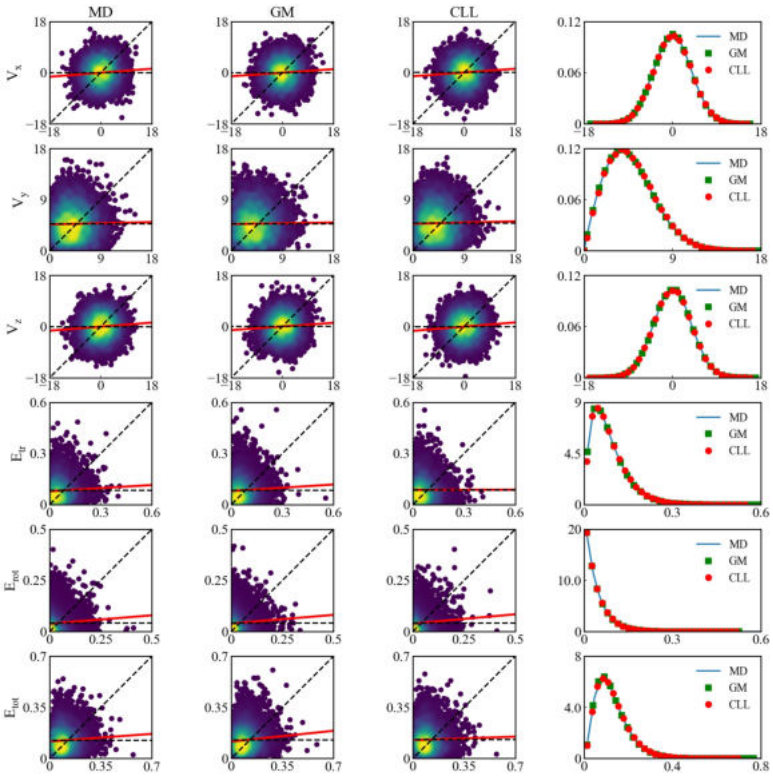

Figure 4.7: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the non-isothermal Fourier thermal problem for N_2-Ni system at the top wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

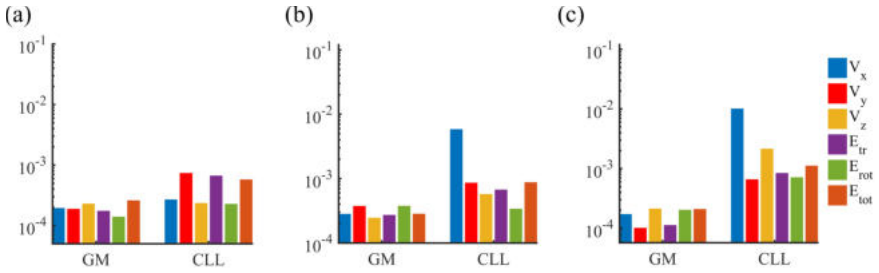

Figure 4.8: The Kullback Leibler divergence of the translational velocity components and different energy modes of N_2-Ni system determined by GM and CLL scattering models; (a) non-isothermal Fourier thermal problem at the top wall; (b) Couette flow problem ($S_v = 0.4$) at the bottom wall; (c) combined Fourier-Couette flow problem ($S_v = 0.4$) at the bottom wall.

Table 4.5: Tangential momentum (α_x,α_z), normal momentum (α_y), translational (α_{tr}), rotational (α_{rot}), and total (α_{tot}) energy accommodation coefficients of the Fourier thermal problem for N_2-Ni system (T_b = 300 K), computed using different scattering kernels: GM and CLL models, and MD simulations.B: bottom wall; T: top wall.

T_t	Model	Wall	α_x	α_y	α_z	α_{tr}	α_{rot}	α_{tot}
300	MD	B	0.873	0.965	0.873	0.912	0.863	0.878
	GM	B	0.869	0.968	0.868	0.920	0.881	0.891
	CLL	B	0.871	0.976	0.871	0.976	0.867	0.941
500	MD	B	0.856	0.960	0.857	0.906	0.804	0.855
		T	0.907	0.959	0.916	0.931	0.892	0.903
	GM	B	0.856	0.959	0.855	0.910	0.821	0.886
		T	0.907	0.957	0.912	0.930	0.907	0.912
	CLL	B	0.856	0.966	0.856	0.975	0.802	0.917
		T	0.911	0.968	0.921	0.978	0.888	0.948

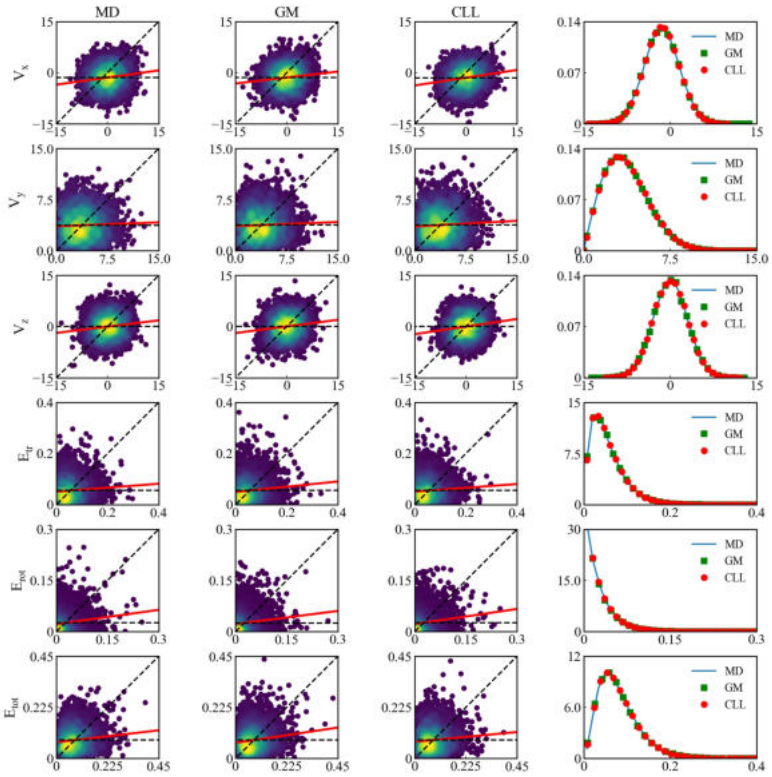

Figure 4.9: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of Couette flow problem (S_v =0.4) for N_2-Ni system at the bottom wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

Table 4.6: Tangential momentum (α_x,α_z), normal momentum (α_y), translational (α_{tr}), rotational (α_{rot}), and total (α_{tot}) energy accommodation coefficients of the Couette flow problem for N_2-Ni system at different speed ratios (S_v), computed using different scattering kernels: GM and CLL models, and MD simulations

S_v	Model	α_x	α_y	α_z	α_{tr}	α_{rot}	α_{tot}
0.2	MD	0.873	0.968	0.873	0.911	0.857	0.873
	GM	0.874	0.967	0.881	0.910	0.873	0.881
	CLL	0.876	0.975	0.873	0.978	0.864	0.940
0.4	MD	0.865	0.968	0.872	0.917	0.862	0.879
	GM	0.860	0.967	0.867	0.907	0.873	0.874
	CLL	0.864	0.973	0.873	0.974	0.859	0.937

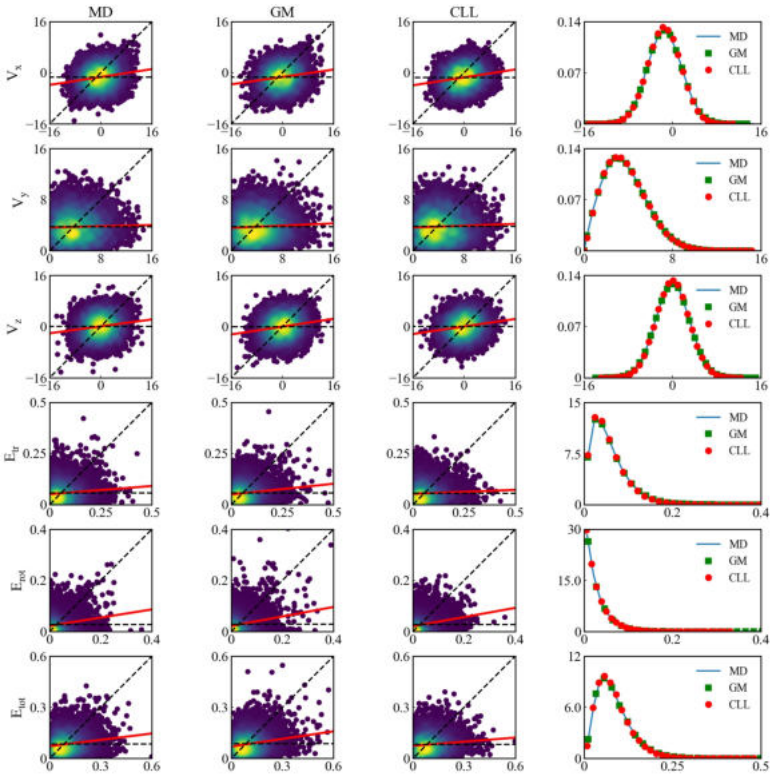

Figure 4.10: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of combined Fourier-Couette flow problem for N_2-Ni system at the bottom wall ($S_w = 0.4$). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

Table 4.7: Tangential momentum (α_x,α_z), normal momentum (α_y), translational (α_{tr}), rotational (α_{rot}), and total (α_{tot}) energy accommodation coefficients of the combined Fourier-Couette flow problem for N_2-Ni system at different speed ratios (S_v), computed using different scattering kernels: GM and CLL models, and MD simulations.B: bottom wall; T: top wall.

S_v	Model	Wall	α_x	α_y	α_z	α_{tr}	α_{rot}	α_{tot}
0.2	MD	B	0.856	0.959	0.855	0.907	0.812	0.856
		T	0.902	0.960	0.908	0.931	0.894	0.903
	GM	B	0.856	0.960	0.856	0.907	0.829	0.863
		T	0.906	0.958	0.916	0.934	0.904	0.904
	CLL	B	0.856	0.961	0.858	0.971	0.814	0.921
		T	0.906	0.954	0.909	0.972	0.899	0.946
0.4	MD	B	0.850	0.962	0.856	0.904	0.812	0.855
		T	0.910	0.967	0.908	0.934	0.897	0.906
	GM	B	0.849	0.962	0.852	0.908	0.831	0.867
		T	0.909	0.965	0.906	0.932	0.905	0.906
	CLL	B	0.853	0.966	0.856	0.975	0.809	0.923
		T	0.917	0.973	0.909	0.980	0.894	0.947

4.3.3. IMPACT OF THE ADSORPTION LAYER ON BEHAVIOR OF REFLECTED GAS MOLECULES

In this part, the MD simulations results for H_2-Ni and N_2-Ni systems are studied more elaborately, in order to gain further insight on the behavior of H_2 and N_2 under different thermal and flow conditions. Here, the variation of the local number density of the gases across the channel height (y-direction) is examined for each benchmark problem. The results for H_2-Ni are shown in Figure 4.11. It is seen that for the systems, in which both walls have the same temperature (see Figure 4.11a), in general, the density near the walls is higher than the bulk density. The density peak at the gas-solid interface indicates the physical adsorption of the gas molecules at the solid surface. This has been also observed previously [74, 83]. It is seen in this figure that by imposing velocity at the walls the density peaks near the walls are diminishing, and at S_v =0.4, the density peaks are lowest. In fact, increasing S_v causes a lower residence time of gas molecules at the surface. For instance, for the isothermal Fourier thermal problem the residence time (t_{RES}) is 1270 ps, while for the Couette flow problem with S_v =0.4, the obtained value is t_{RES}=957 ps. The gas adsorption layer inducing a longer gas-solid interaction time leads to an increased heat transfer at the boundary layer. Higher ACs obtained for the Fourier thermal problem in comparison with the Couette flow problem (see Tables 4.2 and 4.3) also confirms this effect. The number density profiles for H_2-Ni benchmark systems with the walls at different temperatures are presented in Figure 4.11b. It is shown that at the bottom wall (T_b =300 K) the density peak is higher than the density peak at the top wall (T_t =500 K). However, similar to the isothermal walls systems, the reducing effect of the imposed wall velocity on the magnitude of the density peaks adjacent to the walls is also observed here. For the combined Fourier-Couette flow problem (S_v =0.4) at the top wall the value of the number density is even lower than the bulk value. Such behavior has not been observed previously. However, it is noteworthy to mention that, due to high surface temperature in combination with a high wall velocity ($u_W = 629 m/s$), an enormous amount of the kinetic energy is added to the gas molecules. Such an extreme situation can not be studied experimentally, and the main objective of choosing it in this work was the trend study of the scattering models. The obtained ACs for this case are the lowest among the all studied cases for H_2-Ni system.

Figure 4.12 represents the local number density variations of different case studies for N_2-Ni system. The molecular mass of N_2 is considerably higher than H_2 ($m_{N_2} = 14 m_{H_2}$). This fact together with a stronger gas-wall interaction potential between N_2 and Ni leads to more predominant adsorption for N_2 molecules. This can be also deduced from the considerable deviation of the number density in the vicinity of the walls from the bulk number density, as shown in Figure 4.12. Furthermore, the formed layer is strongly attached to Ni surface. Unlike H_2-Ni system, imposing a velocity at the walls does not have any noticeable impact on the density profiles. However, in the case of the system with non-isothermal walls (see Figure 4.12b) the number density at the top hot wall is lower than the bottom cold wall, and this trend is in accordance with the one already observed for H_2-Ni system.

As already mentioned, higher adsorption of gas molecules at the surface causes an enhancement in thermal energy exchange at the gas-solid interface. However, it was observed by Sun et al. [74] that there is a limit for such enhancement. Basically when the

solid surface is fully covered by gas molecules and there is no more space for further adsorption, the energy transport properties do not increase anymore. In fact, under such circumstances gas phase molecules actually collide mainly with the other gas molecules present in the gas layer at the solid surface. Consequently, there is no further energy exchange caused by gas-wall interactions. Our results show that for N_2-Ni system the solid surface is fully saturated with gas molecules, and the interfacial heat transfer is highly dominated by the presence of the gas adsorbed layer. Therefore, changing the characteristics of the system such as imposing a velocity at the wall or increasing the wall temperature does not have any significant impact on the thermal properties of N_2 molecules after colliding with the Ni surface. This is also reflected in the values of various ACs in different case studies for N_2-Ni system reported in section 4.3.2.2, that are very close to each other. On the other hand, in the case of H_2-Ni system the adhesion force between H_2 molecules and Ni surface is very weak, which is mainly due to the relatively low weight of H_2 molecules. Hence, the thermal characteristics of the H_2-Ni interface are very likely to change by imposing some perturbations in the system. This can be also realized by different shapes of the correlation graphs, as well as different values of computed ACs for the various studied benchmark systems in section 4.3.2.1 .

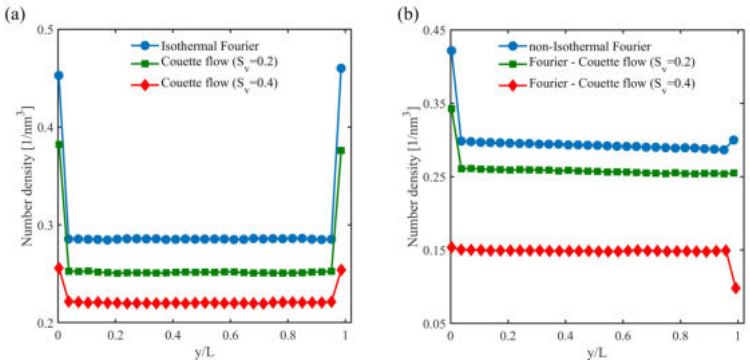

Figure 4.11: Number density profiles for H_2-Ni system with (a) Isothermal walls ($T_b = T_t = 300$ K); (b) non-Isothermal walls ($T_b = 300$, $T_t = 500$ K)

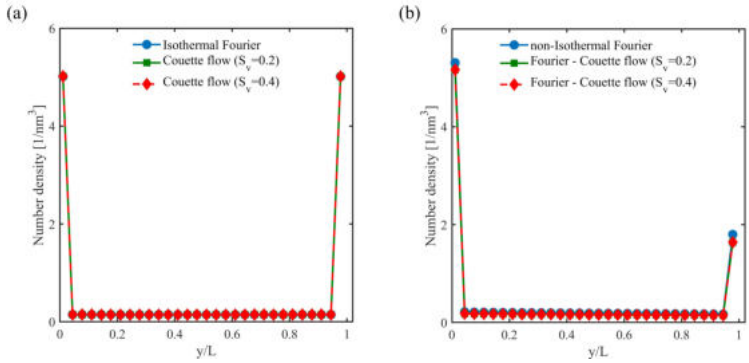

Figure 4.12: Number density profiles for N_2-Ni system with (a) Isothermal walls ($T_b = T_t = 300$ K); (b) non-Isothermal walls ($T_b = 300$, $T_t = 500$ K)

4.4. CONCLUSIONS

In this work, an unsupervised machine learning technique, known as the GM model, is applied to construct a scattering model for diatomic gas molecules (H_2, N_2) interacting with the solid Ni surface. The main reason behind using this specific technique is that GM is considered as one of the most efficient probabilistic machine learning techniques that can be used to cluster a general dataset, in which the subsets follow Gaussian distributions. The GM scattering model is constructed from the superposition of K multidimensional Gaussian functions, each determined by a mean vector and covariance matrix. Therefore, the GM scattering model is categorized as a parametric model. However, since the number of employed parameters is not limited as the classical parametric scattering kernels, the GM model is much more flexible. Here, the entire incoming and outgoing translational and rotational velocity components obtained from the MD simulations are used for training of the model. This guarantees that the model is able to describe all possible phenomena caused by inelastic gas-surface collisions when the system is in a highly non-equilibrium condition.

Using the original MD simulation results for H_2-Ni and N_2-Ni as reference solutions, the performance of the GM scattering model is compared with the CLL scattering model in different thermal problem that are commonly faced in rarefied gas flow systems. It is observed that for H_2-Ni changing the wall temperature or imposing an external velocity at the walls considerably affects the behaviour of postcollisional gas molecules. While the GM scattering model can predict such behavior with good accuracy, the results of the CLL model in non-equilibrium conditions highly deviate from the MD results. Nevertheless, by going toward extremely non-equilibrium situations the performance of the GM model slightly degrades, but it still outperforms the CLL model.

On the other hand, for N_2-Ni case changing the system characteristics does not have a noticeable impact on the kinetic features of N_2 molecules after reflecting from the surface. Here, the shape of the correlation graphs, the PDF of the outgoing translational velocity components and energy modes, as well as the computed ACs in different benchmark systems remain more or less the same. In this case, in general, the predictions from both the employed stochastic scattering models are in a good agreement with the MD results. However, also here the GM results are slightly in better agreement with the MD data.

To shed some lights on the possible reasons behind very different behaviors of H_2 and N_2 gases in the studied systems, the variation of the local number density of the gas molecules in each system obtained from the conducted MD simulations have been investigated. It is observed that in the case of N_2-Ni the solid surface is saturated with gas molecules and applying different perturbations at the walls, such as increasing the temperature or imposing an external velocity, does not have a considerable impact on the amount of adsorbed gas molecules. Consequently, the incoming gas molecules from the bulk of the gas towards the surface mainly encounter other gas molecules that are stuck at the surface, and they do not really exchange energy with Ni molecules. However, for the H_2-Ni the adsorption layer is less dominant. As a result, changing wall features (temperature or velocity) has relatively more impacts on the behavior of reflected gas molecules in the case of H_2 than for N_2.

The observed high precision of the GM predictions indicates that it can be considered a

promising candidate to compute important discontinuity phenomena such as temperature jump and velocity slip in rarefied gas flow systems. In addition, the accuracy of the GM model results indicates the high potential of this approach to construct a generalized scattering kernel for diatomic gas-solid surface interactions. Nevertheless, a more extended dataset, including a wider range of wall temperatures and gas densities, is required to construct such a model. Further studies will be devoted to these problems.

5

GENERALIZED GAUSSIAN MIXTURE WALL MODEL FOR RAREFIED GASES

In Chapters 3 and 4 the GM approach was employed to derive the GSI models for monoatomic and diatomic gases, respectively. However, in the aforementioned chapters, the GM model was trained separately for each set of collisional data gathered from an MD simulation for specific temperatures, pressures and wall velocities. In this chapter, we investigate the capability of the GM approach in deriving a generalized GSI model for monoatomic and diatomic rarefied gases interacting with solid surfaces. Here, two parallel walls systems: Ar-Au and H_2-Ni are chosen as the case studies. In a two parallel walls system, gas pressure and surface temperature are among the potential candidates that can affect the behavior of gas molecules after reflecting from the walls. Therefore, MD simulations at different pressures and surface temperatures are initially performed for each gas-surface pair. Before implementing the GM model, the obtained results from different MD simulations for Ar-Au and H_2-Ni systems are assessed. The main objective of such assessment is to pick up the more influential parameter and train the GM model to predict the impact of that parameter on the post-collisional behavior of gas molecules. In the case of the Ar-Au system, the assessment is carried out by comparing the velocity distributions and total energy accommodation coefficient (EAC) in different case studies. For H_2-Ni, the distributions of translational velocity components and different energy modes, as well as various energy accommodation coefficients (α_{tr},α_{rot},α_{tot}) are used for the assessment purpose.

It is noteworthy to mention that, previously, in Chapter 2, it was shown that increasing pressure causes a slight enhancement in the values of EAC for the Ar-Au system. However, for the sake of completeness of this chapter, a similar kind of assessment is also performed here.

5.1. EFFECT OF GAS PRESSURE ON THE POST-COLLISIONAL BEHAVIOR OF MONOATOMIC GAS MOLECULES

The MD simulation setup used for the Ar-Au system is similar to the one explained in Chapter 3. The only difference is the normal distance between the walls, which is fixed at d = 20 nm in the current case. Two benchmark systems are considered: the isothermal walls system (T_b = T_t = 300 K), and the non-isothermal walls system (T_b = 300 K, T_t = 500 K). For each benchmark system, the number of gas molecules confined between the walls is altered in such a way that the pressure does not exceed critical gas pressure P_{cr} ($P_{cr\text{-}Ar}$ = 4.86 MPa [107]). The velocity correlation between impinging and outgoing velocity components, as well as the PDF of the outgoing velocities for different values of gas pressure related to the isothermal and non-isothermal systems at the bottom wall, are depicted in Figures 5.1 and 5.2, respectively. In these systems, the gas pressure is increased from 0.4 MPa to 4 MPa. Such an enhancement in the pressure leads to a reduction in the degree of rarefaction in the system from Kn = 1 to Kn = 0.1. Looking at the velocity correlations obtained at different pressures, it is seen that at low pressure (P = 0.4 MPa), the velocity clouds in the tangential directions (x,z) are more concentrated around the dashed diagonal line, indicating more specular reflection at the low pressure. However, by increasing the pressure, the shape of the velocity clouds in the tangential directions slightly changes, and gradually they become more condensed around

the horizontal line. Other than that, it is observed that changing pressure does not have any notable impact on the general shape of the velocity correlations in the normal direction, as well as the probability density functions (PDFs) of the outgoing velocities in both isothermal and non-isothermal systems.

The variations of the EAC with the gas pressure in the isothermal and non-isothermal systems are presented in Figure 5.3. It is seen that increasing pressure leads to obtaining higher EAC in both systems. Although that the bottom wall temperature in the isothermal and non-isothermal case are identical, the EAC in the isothermal case is about 2% (at 4MPa) to 9% (at 0.4MPa) larger than in the non-isothermal case. This has probably to do with an increase in average gas temperature, resulting in a lower EAC. With increasing pressure (decreasing Kn) this difference becomes smaller: this is probably due to more intermolecular collisions (and thereby exchanging energy with other molecules) before a gas molecule hits the cold wall after leaving the hot wall. In the case of the isothermal system (see Figure 5.3a), increasing the pressure from 0.4 MPa to 4 MPa causes a 14% enhancement in the value of EAC. Similar behavior is observed in the case of the non-isothermal system (see Figure 5.3b), in which increasing the pressure of the confined gas between the walls from 0.4 MPa to 4 MPa leads to 17% and 14% enhancement in EAC computed at the bottom and top walls, respectively. Such an enhancement of EAC in both case studies, can be also detected just by looking to the correlations plots of the systems (see Figures 5.1 and 5.2). It is shown in these plots that by increasing the pressure, the slope of the red lines for each velocity component gets closer to the horizontal line indicating a higher AC for the system. Similar pressure dependency has been also observed in an empirical study performed by Thomas and Brown [82].

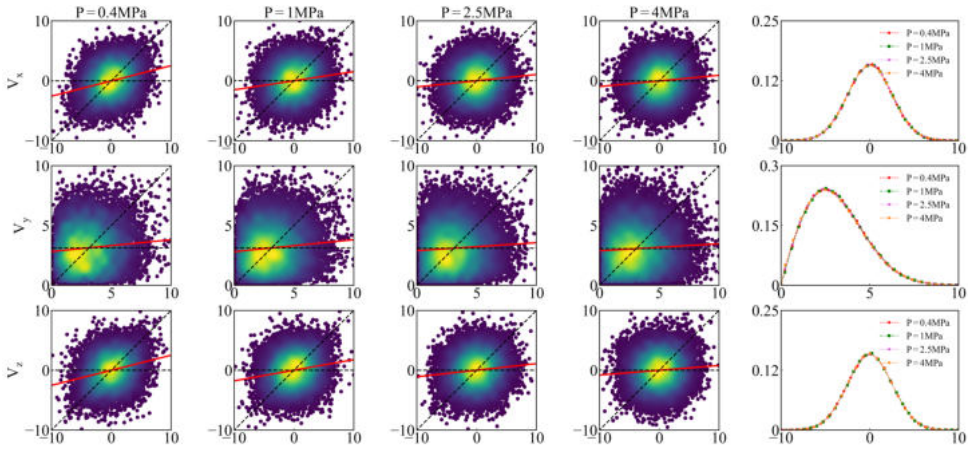

Figure 5.1: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in [Å/ps] for isothermal Ar-Au system at the bottom wall at different pressures. The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density functions for the reflecting particles are shown.

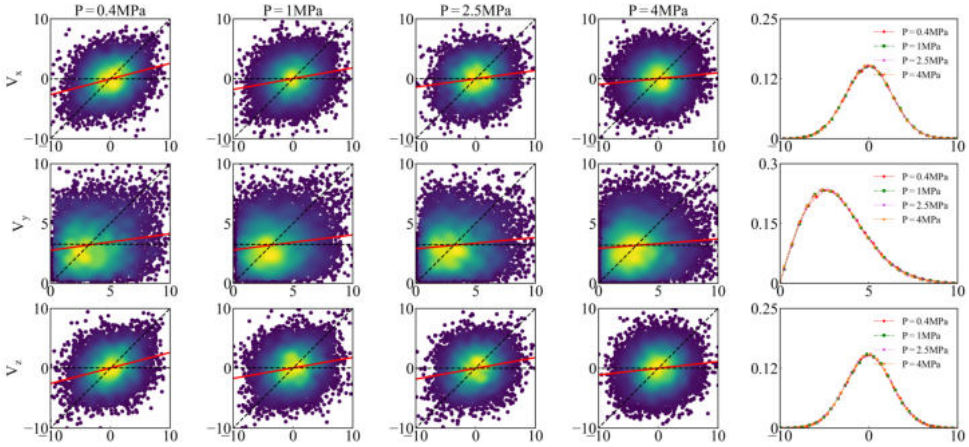

Figure 5.2: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in [Å/ps] for non-isothermal Ar-Au system at the bottom wall at different pressures. The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density functions for the reflecting particles are shown.

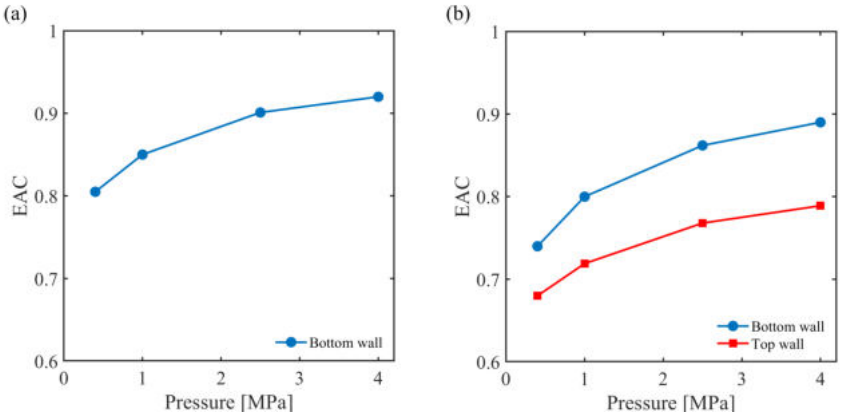

Figure 5.3: The values of energy accommodation coefficient (EAC) at different pressures for: (a) isothermal Ar-Au system (b) non- isothermal Ar-Au system.

5.2. EFFECT OF SURFACE TEMPERATURE ON THE POST-COLLISIONAL BEHAVIOR OF MONOATOMIC GAS MOLECULES

The same MD setup described in Section 5.1 is employed to characterize the impact of surface temperature on the post-collisional behavior of Ar molecules. The bottom wall temperature is kept constant at T_b = 300 K, while the temperature of the top wall (T_t) is changed in the range of 200 K to 600 K. The Kn number in all these simulations is kept constant at 0.25. The obtained velocity distributions at the top wall for different surface temperatures are depicted in Figure 5.4. Looking at the correlations graphs, it is inferred that by increasing the surface temperature, the slope of the red line, hinting at the value of AC, is gradually going towards the diagonal line. This means Ar molecules are less accommodating to the Au surface at higher temperatures. In addition, it is observed that the PDFs of the outgoing velocities at different surface temperatures are quite distinctive. As it was already pointed out in Chapter 4, the pick of the velocity PDF refers to the gas temperature that can be computed based on the average kinetic energy of the gas molecules. In general, a higher pick denotes a lower temperature. The trends observed in the plotted PDfs in Figure 5.4 are in accordance with this fact. The variation of EAC with the surface temperature at the top wall is shown in Figure 5.5. It is observed that increasing the surface temperature from 200 K to 600 K leads to a 22% reduction in the value of EAC. Similar temperature dependency has been also reported in experimental studies [38].

Comparing the impact of gas pressure and surface temperature on the post-collisional behavior of Ar molecules, it is deduced that in the considered range of pressure and temperature, the impact of the surface temperature on EAC is slightly more prominent than the pressure. However, referring to the shown velocity distributions, while changing the gas pressure does not have any significant impact on the PDF of the outgoing velocities obtained from different systems, the surface temperature has a considerable impact on the PDFs of the outgoing velocities. Therefore, in the case of the Ar-Au system, the surface temperature is used as the target variable included in the training of the GM model.

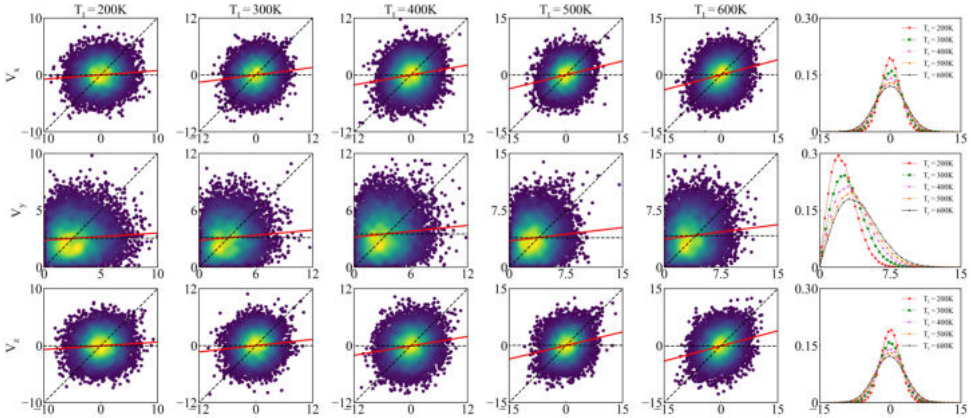

Figure 5.4: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in [Å/ps] for Ar-Au system at the top wall at different surface temperatures. The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density functions for the reflecting particles are shown.

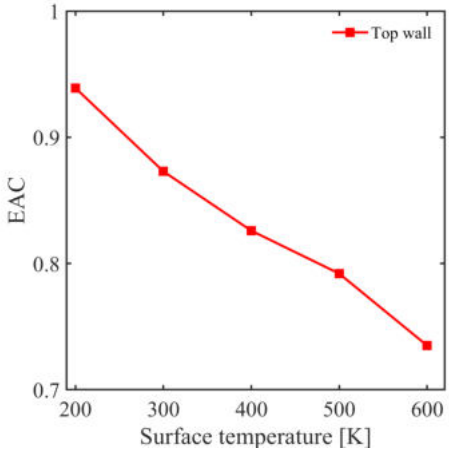

Figure 5.5: Variation of the energy accommodation coefficient (EAC) with surface temperature for Ar-Au system at the top wall.

5.3. EFFECT OF GAS PRESSURE ON THE POST-COLLISIONAL BEHAVIOR OF DIATOMIC GAS MOLECULES

The dimensions of the MD simulation setup used here are the same as the one demonstrated in Chapter 4. The number of H_2 molecules between the Ni surfaces is altered in such a way that the pressure of the gas in bulk does not overpass the value of critical pressure ($P_{cr\text{-}H_2}$=1.3 MPa [107]). Two types of benchmark systems are studied. The isothermal walls system with T_b= T_t = 300 K, and the non-isothermal walls system with T_b =300 K and T_t =400 K. Herein, changing the pressure induces a variation in the Knudsen number in the system from Kn = 0.8 (at P = 0.5 MPa) to Kn = 0.2 (at P = 1.2 MPa). The obtained correlation graphs and PDFs of the outgoing translational velocities and energy modes for the aforementioned systems at different pressures are demonstrated in Figures 5.6 and 5.7, respectively. It is seen that for both benchmark systems changing pressure does not have any considerable impact on the correlation graphs and the corresponding PDFs.

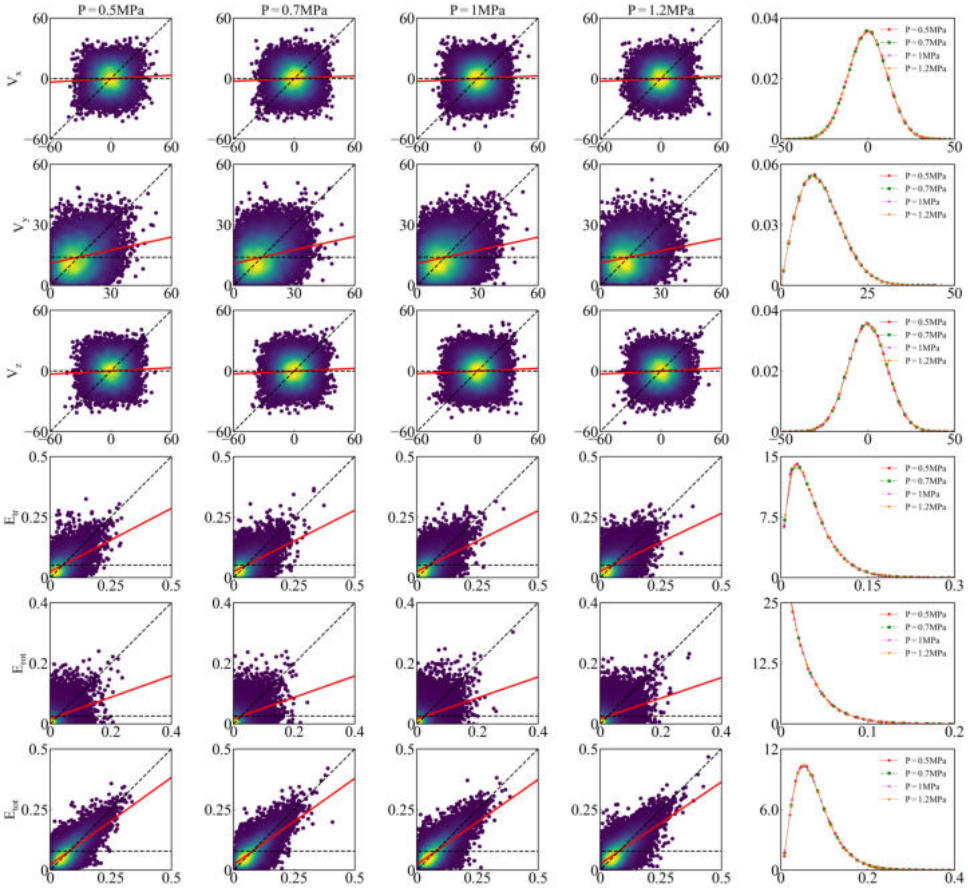

Figure 5.6: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the isothermal H_2+Ni system at the bottom wall at different pressures. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented.

To investigate further the effect of gas pressure on the behavior of reflected H_2 molecules in the considered systems, the obtained ACs for different energy modes are presented in Figure 5.8. It is observed that increasing pressure causes an enhancement in all computed ACs. Nevertheless, the increments in ACs values are very small compared to the Ar-Au system. As an example, for the isothermal system (see Figure 5.8a) the value of α_{rot} is almost constant, while the increments of α_{tr} and α_{tot} are 5% and 10%, respectively. In the case of the non-isothermal system (see Figure 5.8b), the maximum enhancement is related to α_{tot} at the bottom wall, which is 12%.

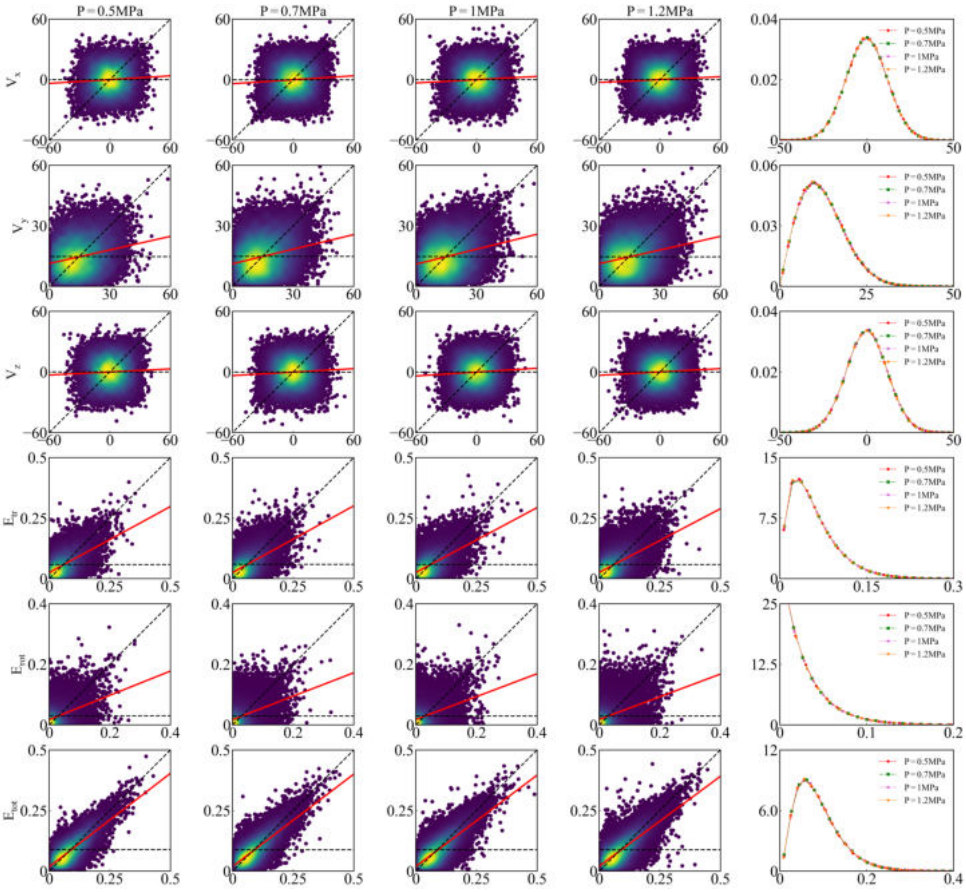

Figure 5.7: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the non-isothermal H_2+Ni system at the bottom wall at different pressures. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented.

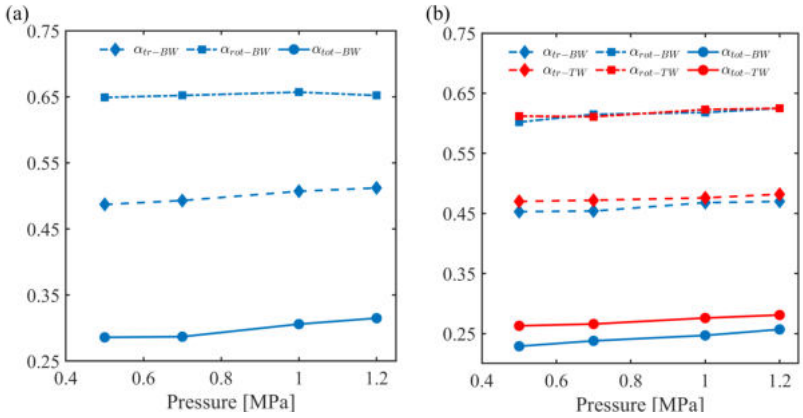

Figure 5.8: Variation of accommodation coefficient for different energy models (translational: α_{tr}, rotational: α_{rot}, and total energy: α_{tot}) with pressure for: (a) isothermal H_2-Ni system (b) non- isothermal H_2-Ni system. BW: bottom wall, TW: top wall.

5

5.4. EFFECT OF SURFACE TEMPERATURE ON THE POST-COLLISIONAL BEHAVIOR OF DIATOMIC GAS MOLECULES

A similar study to the Ar-Au system is conducted to examine the impact of the surface temperature on the kinematic properties of H_2 molecules after reflecting from the Ni surface. The temperature of the bottom wall is kept constant T_b= 300 K, whereas the temperature of the top wall (T_t) is varied in the range of 200 K < T_t < 600 K. The Kn number in all these simulations is kept constant at 0.35. Here, the collisional data collected at the top wall from separated MD simulations are used for assessment purpose. The correlation graphs and PDFs of the outgoing velocities, as well as the energy modes, are depicted in Figure 5.9. Based on the correlation graphs, no clear difference is noted in the shape of velocity or energy clouds obtained from the MD simulations at various surface temperatures. However, in the correlation graphs, it is seen that by increasing the surface temperature, the red line slope becomes slightly closer to the diagonal line, indicating less accommodating of the gas molecules to the surface. On the other hand, a clear difference between the PDFs of the outgoing translational velocities and energy modes is seen in the last column of Figure 5.9. To attain a quantitative understanding of the influence of the surface temperature on various ACs, the trends of ACs related to different energy modes with the surface temperature are presented in Figure 5.10. Similar to the Ar-Au system, increasing the surface temperature leads to a reduction in the values of ACs for H_2-Ni system. However, the impact of surface temperature is even more prominent in the present case. Herein, raising the surface temperature from 200 K to 600 K induces a 20% reduction in the values of α_{tr} and α_{rot}, as well as a 35% reduction in the value of α_{tot}.

Comparing the results presented in this section with the ones from the previous section, it is inferred that the surface temperature is more influential than the pressure in the post-collisional behavior of H_2 molecules. Therefore, it is added to the MD collisional data for the training purpose of the GM model.

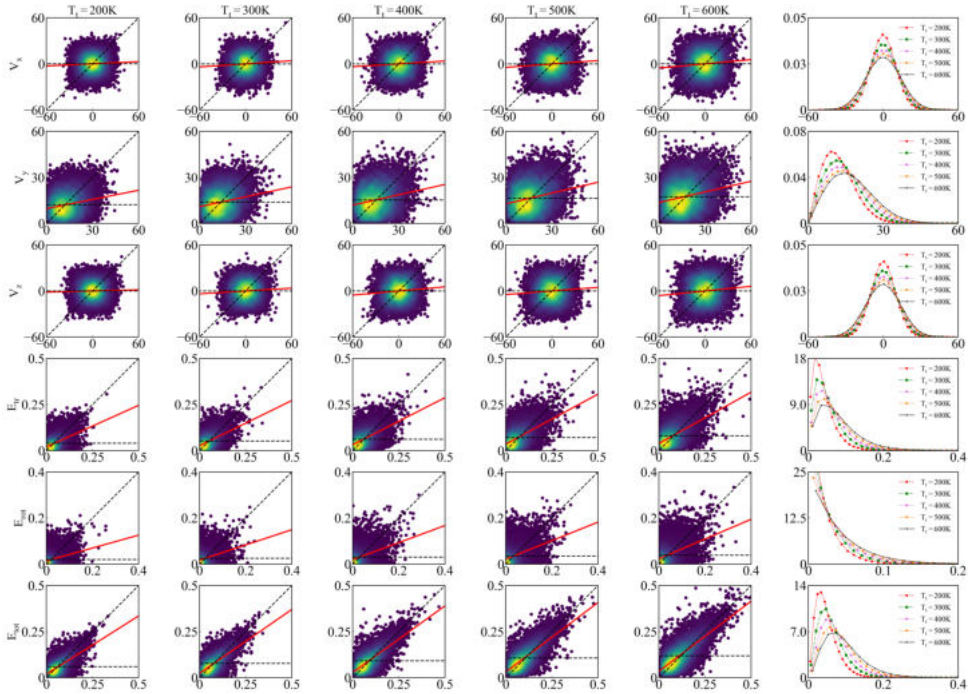

Figure 5.9: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the non-isothermal H_2+Ni system at the bottom wall at different surface temperatures. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented.

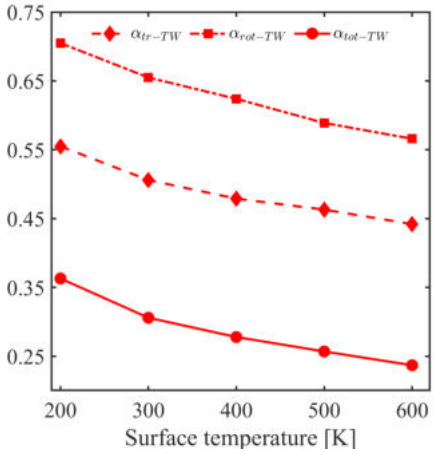

Figure 5.10: Variation of accommodation coefficients related to different energy modes with surface temperature for H_2-Ni system at the top wall.

5.5. CONSTRUCTING A GENERALIZED GM SCATTERING MODEL FOR AR-AU SYSTEM

As addressed in Section 5.2, the surface temperature has relatively more impact on the behavior of Ar molecules reflecting from the Au surface than the gas pressure. Therefore, the temperature data is added to the data set used for the training of the GM model.
As illustrated in Figure 5.11, five independent MD simulations are carried out. The only difference between these simulations is the temperature of the top wall that varies in the range of T_{t1}=200 K to T_{t5}=600 K in steps of 100 K. From each collisional data set gathered at the top wall of the performed MD simulations, 25% is separated as the test set, and the remaining 75% is added to other four collected data sets for the training of the GM model. Herein, $n1$ to $n5$ are the number of training data points gathered from each MD simulation. As it has been already discussed in Chapter 3, normal velocity components (v'_y, v_y) obtained for the MD simulations follow a Rayleigh distribution. Therefore, to get better performance from the GM model, these velocity components are transferred into the Gaussian distribution following the preprocessing scheme explained in Chapter 3 (see Figure 5.11). No preprocessing has been performed for the temperature values, and the same values shown in Figure 5.11 are fed directly into the GM model. In matrix Λ shown in Figure 5.11, the values of T_{t1-1} to T_{t1-n1} are equal to the top wall temperature in the first MD simulation (T_{t1}=200 K). In fact, the same temperature value is repeated $n1$ times in this matrix. The same rule is also applied to the other temperature values in the matrix. It means for a collisional data set including n training data points, the corresponding T_t presents n times in the final training set. Here, the training data is a 7-dimensional data set $(T_t, v'_x, v'_y, v'_z, v_x, v_y, v_z)$. In matrix Λ the values in the first column are one order of magnitude larger than the rest of the training data. Normally under such circumstances, normalization of the data is a necessary step before training the model to have acceptable performance in other machine learning approaches. However, in the case of the GM approach, this issue does not have any notable impact on the model's performance. The obtained covariance matrix for a randomly chosen multivariate Gaussian function after the training of the model is presented in Figure 5.12. It is seen that the covariance values between different T_t and velocity components are considerably smaller than the covariance values between different pairs of velocity components. This induces that the first column, including the top wall temperature, has a minor impact on the predicted velocity components by the model, which are of higher importance for us. After training the GM model, the obtained weights, covariance matrices, and mean vectors are used to generate the new incoming and outgoing velocities at the corresponding wall temperatures. The predicted results by the GM model are compared with the original MD data and those predicted by the CLL scattering model to evaluate the performance of the GM model. Similar to Chapter 3, here also, the assessment is carried out in terms of the distribution of the predicted postcollisional velocities and EAC.

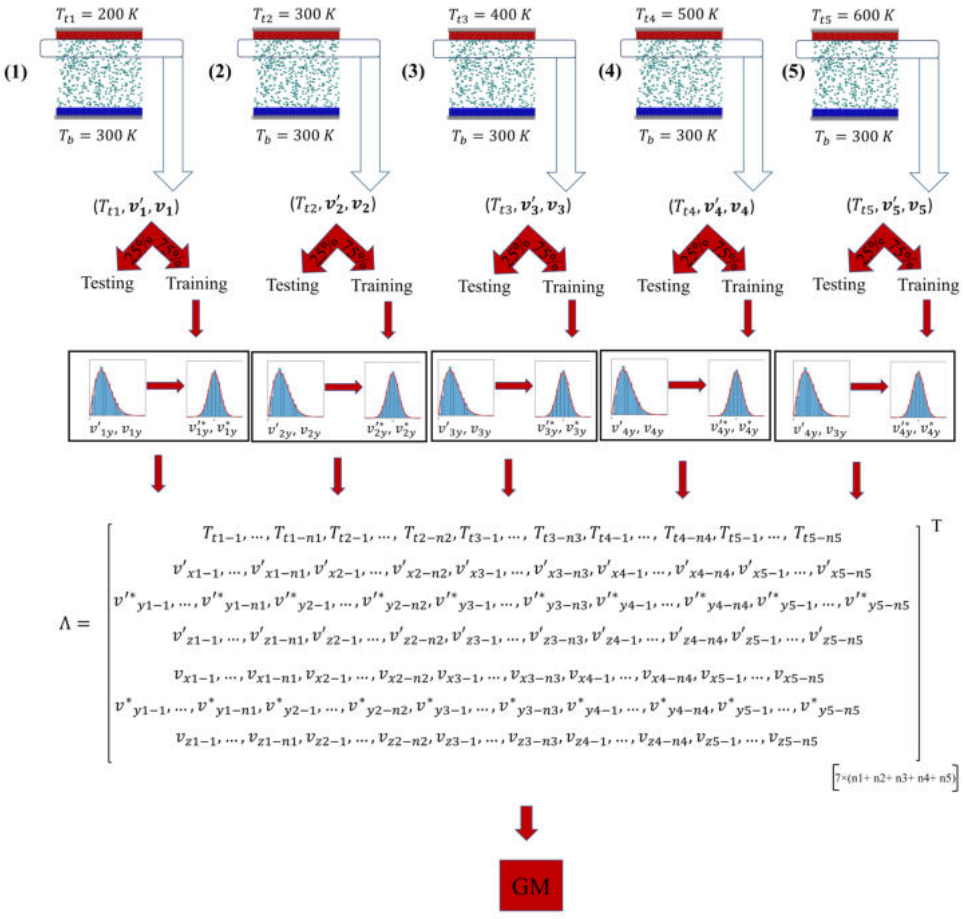

Figure 5.11: Workflow diagram followed to construct the generalized GM scattering model for Ar-Au system. $\boldsymbol{v'}$: (v'_x, v'_y, v'_z) and $\boldsymbol{v}$: (v_x, v_y, v_z) are incoming and outgoing velocity vectors obtained from MD simulations. v'^{*}_y and v^{*}_y are the transformed incoming and outgoing velocity components obtained from MD simulations.

	T_t	$v_{x'}$	$v_{y'}$	$v_{z'}$	v_x	v_y	v_z
T_t	1.00E-06	2.86E-27	5.74E-27	-6.55E-29	-6.85E-27	1.49E-27	4.14E-27
$v_{x'}$	2.86E-27	5.713	-0.590	0.063	-0.505	-0.589	-0.222
$v_{y'}$	5.73E-27	-0.590	4.101	0.077	0.053	0.409	-0.052
$v_{z'}$	-6.64E-29	0.063	0.077	5.151	-0.208	0.305	0.006
v_x	-6.85E-27	-0.505	0.053	-0.208	5.513	0.072	0.557
v_y	1.49E-27	-0.589	0.409	0.305	0.072	5.334	-0.280
v_z	4.14E-27	-0.222	-0.052	0.006	0.557	-0.280	6.174

Figure 5.12: The covariance matrix for a randomly chosen multivariate Gaussian function after training the generalized GM model for Ar-Au system. It is shown that the off-diagonal values for the first column and row are much smaller than the diagonal value and the covariance values between different pairs of velocity components.

Regarding the velocity distributions, it is observed that the results predicted by the GM model are always in good agreement with the MD data (see Appendix G). On the other hand, the CLL model results up to T_t= 400 K are consistent with the MD results, while the CLL model performance degrades by going to the higher wall temperatures. As example, the velocity correlation graphs and the PDFs at T_t= 400 K and T_t= 600 K obtained from the GM and CLL scattering models, as well as the MD simulations are depicted in Figures 5.13 and 5.14, respectively. Looking at the results at T_t= 400 K (see Figure 5.13), no considerable differences can be noted between the results obtained from the GSI models and MD simulation. For the system with T_t= 600 K (see Figure 5.14) the correlations graphs based on both scattering models match with the MD results. However, the predicted PDFs of the outgoing velocities by the CLL model, especially in the tangential directions, deviate from the MD data. To compare the performance of the GSI models in terms of the predicted velocity distributions in a quantitative manner, the Kullback-Leibler divergence (KLD) [138] coefficients between these models and the original MD data are computed (see Figure 5.15). It is shown that the KLD values for the GM model are smaller than the CLL model results, indicating the relatively better performance of the GM model. Besides, it can be realized that by increasing the top wall temperature, the KLDs for the CLL model are also growing, which means the inferior performance of the model.

EAC computed based on the predicted results by the GM and CLL scattering models, and the original MD results are presented in Figure 5.16. The number of Gaussian functions, K, for the GM model varies in the range of $1 < K < 1000$. It is observed that in all case studies, the GM model for $K \geqslant 100$ has better performance than the CLL model. In fact, for $K = 100$ the deviation of the $\mathrm{EAC_{GM}}$ obtained from all the case studies are already within 5% of the MD results ($\mathrm{EAC_{MD}}$). To have a more explicit view of the capability of the utilized stochastic GSI models to predict the variation of EAC with the surface temperature, the trends of EAC with T_t based on applied approaches are depicted in Figure 5.17. It is seen that the GM model can anticipate the inverse relation between T_t and EAC with a high precision, whereas by going to higher temperatures, the deviation between the CLL results ($\mathrm{EAC_{CLL}}$) and MD results ($\mathrm{EAC_{MD}}$) is increasing.

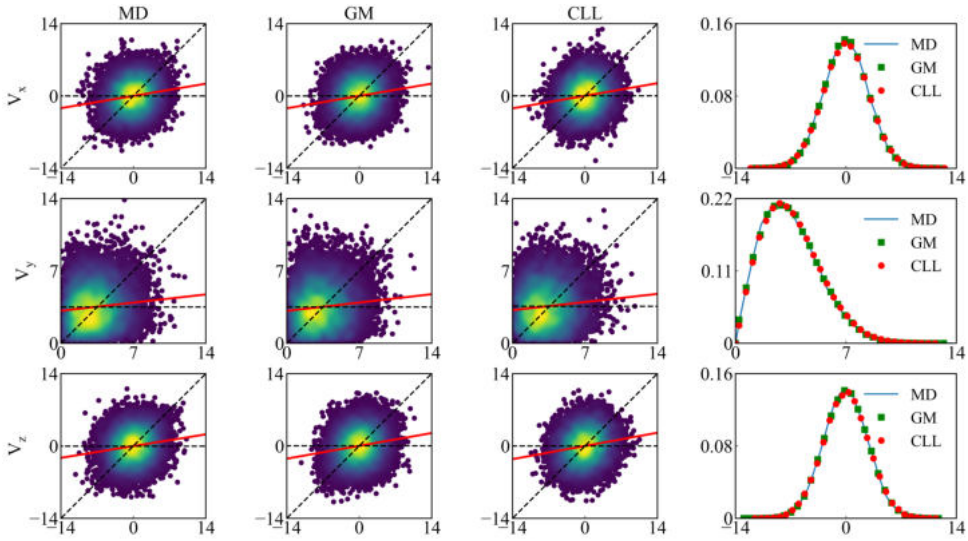

Figure 5.13: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in [Å/ps] for Ar-Au system at the top wall(T_b=300 K,T_t=400 K). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density functions for the reflecting particles are shown.

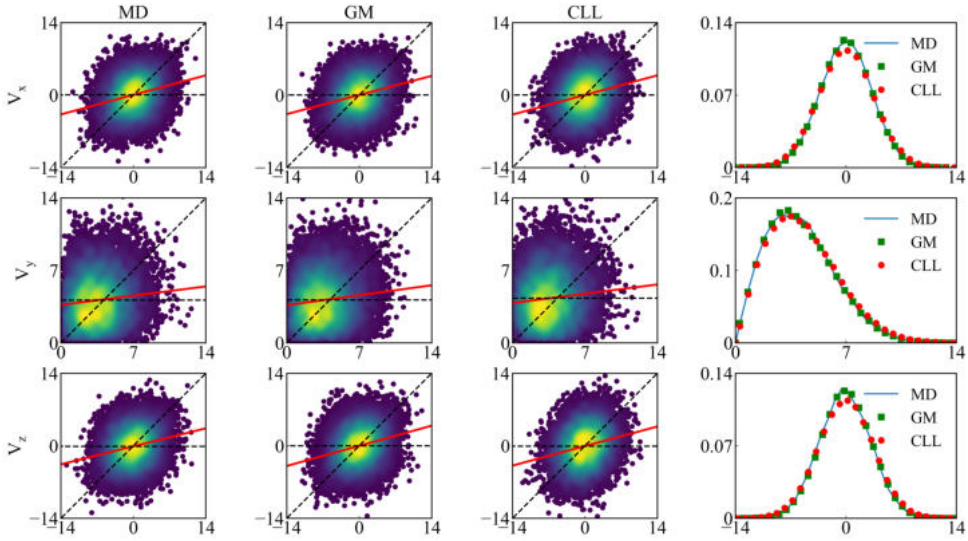

Figure 5.14: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in [Å/ps] for Ar-Au system at the top wall(T_b=300 K,T_t=600 K). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density functions for the reflecting particles are shown.

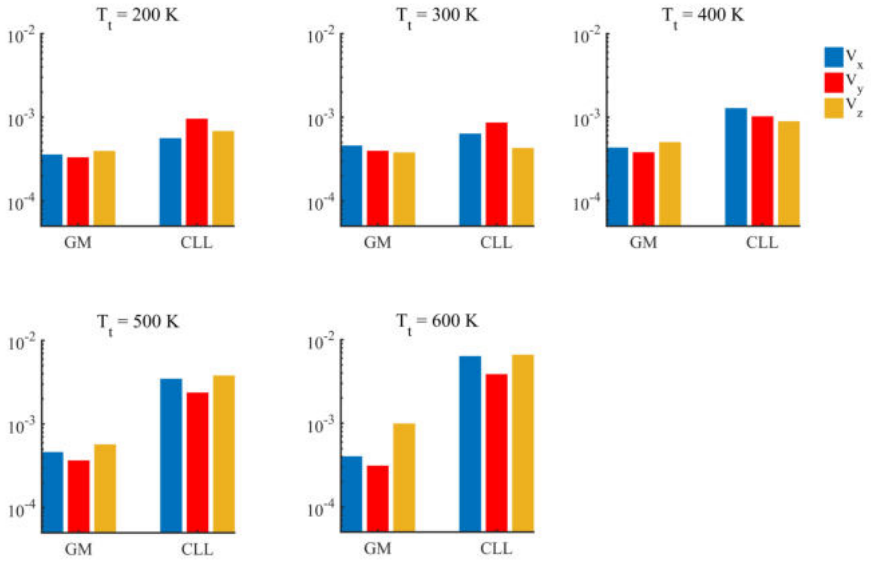

Figure 5.15: The Kullback Leibler divergence (KLD) of the velocity components of Ar-Au system at the top wall determined by GM and CLL scattering models.

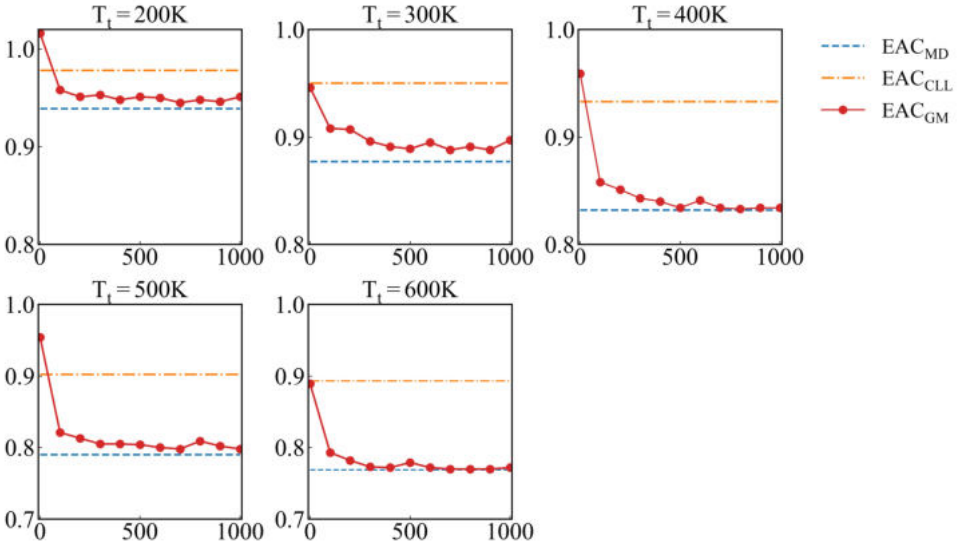

Figure 5.16: Variation of the energy accommodation coefficient (EAC) obtained from the MD simulations, the GM model, and CLL model for two parallel walls systems with different top wall temperatures. Horizontal axis shows the number of applied Gaussian functions in the GM model.

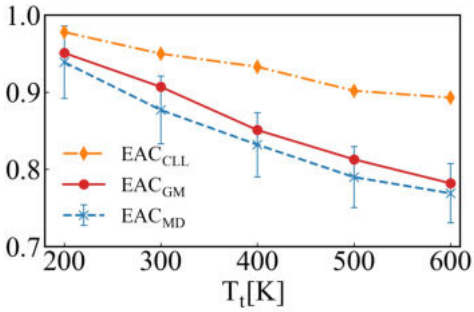

Figure 5.17: The values of energy accommodation coefficient (EAC) obtained from the MD simulations, the GM mode, and CLL model at different surface temperatures. The margins show 5% of the MD values.

5.6. CONSTRUCTING A GENERALIZED GM SCATTERING MODEL FOR H_2-NI SYSTEM

The procedure followed to construct a generalized GSI for H_2-Ni system based on the GM approach is demonstrated in Figure 5.18. In the first step, similar to the Ar-Au system, five separate MD simulations using different top wall temperatures ($200 < T_t < 600$) are carried out. The main difference between the present case study with the Ar-Au system is that since diatomic gas molecules are considered here, the incoming and outgoing rotational velocity vectors of the H_2 molecules obtained from each MD simulation ($\boldsymbol{\xi'} = (\omega', \vartheta'), \boldsymbol{\xi} = (\omega, \vartheta)$ shown in red in Figure 5.18) are also added to the final training data set. Therefore, the final training set (see matrix Λ in Figure 5.18) is a 11-dimensional matrix.

The other important issue is that to get a better performance of the model the translational and rotational velocities in Λ are normalized by $\sqrt{\frac{2k_B T_t}{m_g}}$ and $\sqrt{\frac{2k_B T_t}{I}}$,respectively. Here, T_t describes the corresponding temperature of the top wall. After training the GM model using the obtained model parameters, the incoming and outgoing velocity components are generated for the corresponding values of T_t. The performance of the GM model is examined against the CLL model and the original MD data based on the following criteria: (i) the correlation between the incoming and outgoing translational velocity components of the centre of mass (COM) and different energy modes of H_2 molecules; (ii) the PDF of the outgoing velocity components, and the energy modes; (iii) the ACs related to different energy modes.

It is observed that the velocity and energy distributions predicted by the GM model at different T_t are always in good agreement with MD data (see Appendix H). Regarding the CLL model predictions, except at T_t = 300 K, some discrepancies with the original MD data are always observed. Such discrepancies grow by increasing the temperature difference between the two walls. For instance, the results at T_t = 200 K and T_t = 600 K are presented in Figures 5.19 and 5.20, respectively. Looking to Figure 5.19, it is seen that all correlation graphs and PDFs of the outgoing velocity and energy components predicted by the GM model match well the MD data. Referring to the correlation graphs based on the CLL model, except for V_y and E_{rot}, the other ones have different shapes compared to the MD results. Based on the PDFs of the outgoing velocity components and energy modes, it is deduced that the predicted E_{rot} by the CLL model is consistent with the MD data. However, the other PDFs slightly deviate from the MD results. Similar observations are made in the case of the system with T_t =600 K (see Figure 5.20). However, based on the shown PDFs in the last column, more deviation between the MD and CLL results are detected.

As another statistical criterion, the KLD coefficients computed between the GSI models and the atomistic data are illustrated in Figure 5.21. It is seen that except in the case of the isothermal walls system, in other case studies, KLDs based on the GM model are smaller than the CLL results, indicating the superior performance of the GM model.

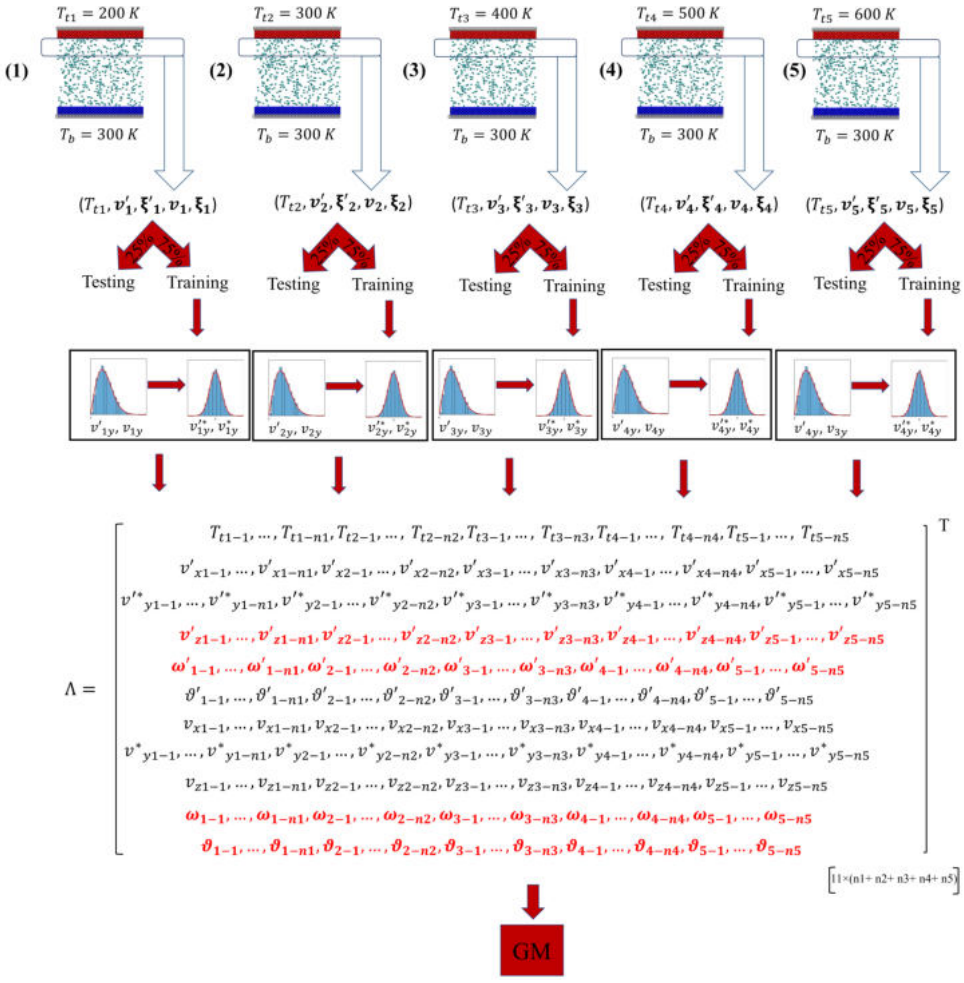

Figure 5.18: Workflow diagram followed to construct the generalized GM scattering model for H_2-Ni system. $\boldsymbol{v'}$: (v'_x, v'_y, v'_z) and $\boldsymbol{v}$: (v_x, v_y, v_z) are incoming and outgoing translational velocity vectors of the center of mass of the gas molecules obtained from MD simulations. $\boldsymbol{\xi'} = (\omega', \vartheta')$ and $\boldsymbol{\xi} = (\omega, \vartheta)$ are incoming and outgoing rotational velocity vectors obtained from MD simulations. v'^{*}_y and v^{*}_y are the transformed incoming and outgoing velocity components obtained from MD simulations.

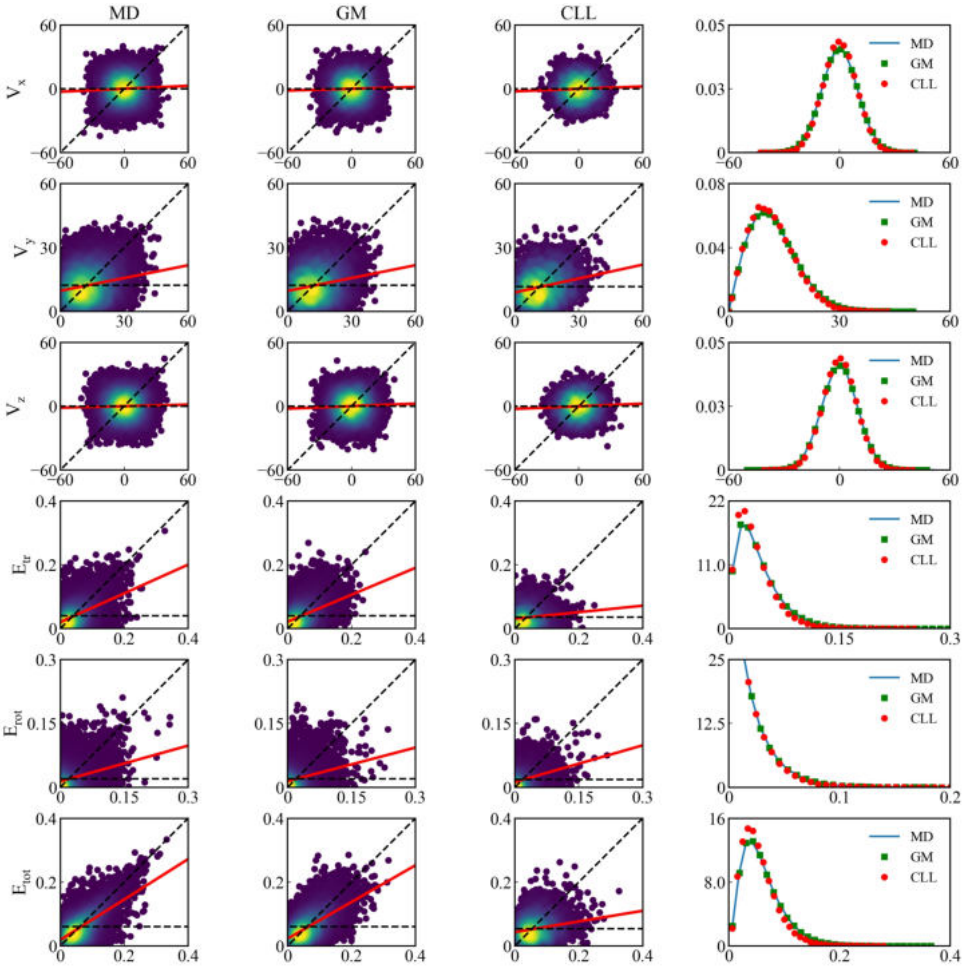

Figure 5.19: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of H_2-Ni system at the top wall (T_b = 300 K, T_t = 200 K). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

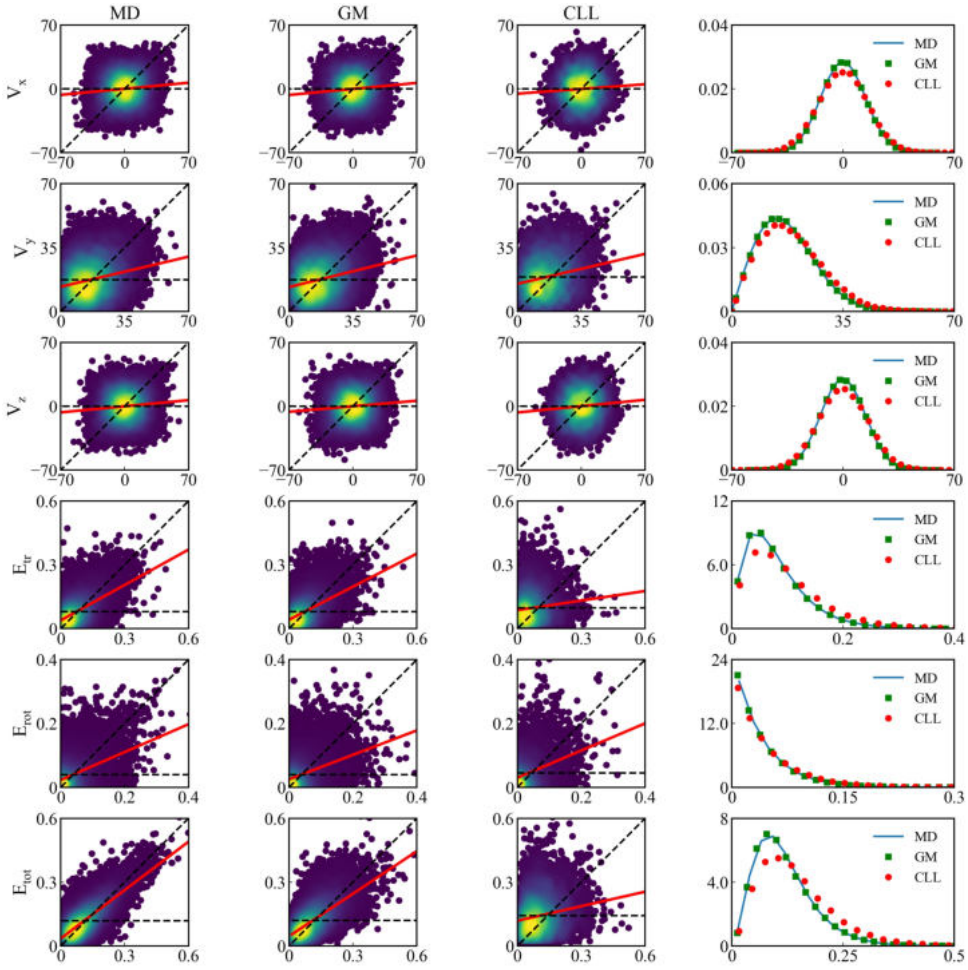

Figure 5.20: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of H_2-Ni system at the top wall (T_b = 300 K, T_t = 600 K). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

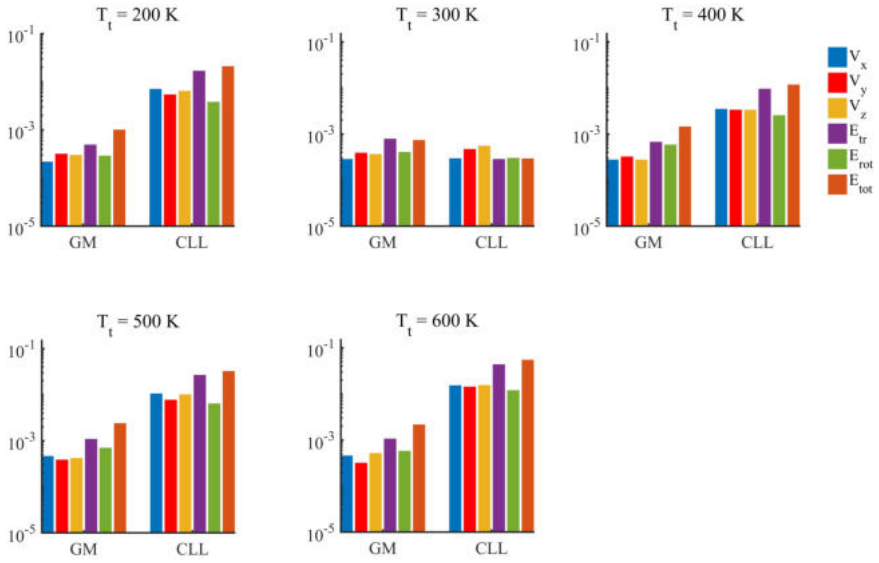

Figure 5.21: The Kullback Leibler divergence of the translational velocity components and different energy modes of H_2-Ni system at the top wall determined by GM and CLL scattering models.

Based on the translational and rotational velocity components predicted by the GM and CLL scattering models, the ACs for different energy modes are computed. The comparisons between these results with the original MD data for the different case studies are depicted in Figure 5.22. It is shown that except for predicting α_{rot}, the GM model considerably outperforms the CLL model in the case of other energy ACs. However, compared with the Ar-Au system (see Figure 5.16), the performance of the GM model has deteriorated here. To be more specific, using 1500 Gaussians in the GM model, the lowest captured deviations are at T_t =200 K that are 6%, 5%, and 18% related to α_{tr}, α_{rot}, and α_{tot}, respectively. On the other, the highest deviations are detected at T_t =600 K that are 12%, 11%, and 44% related to α_{tr}, α_{rot}, and α_{tot}, respectively. Two reasons could cause the spotted deviations: the model performance or the sampling approach used afterward to generate the translational and rotational velocity components. Different sampling approaches are examined in the following to determine the main reason.

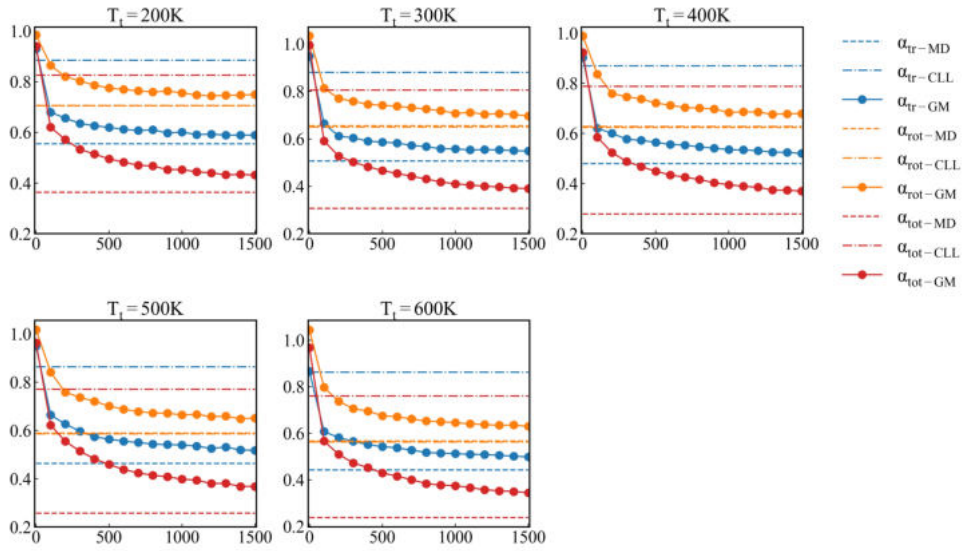

Figure 5.22: Variation of different accommodation coefficients obtained using correlation method for two parallel walls H_2-Ni system at the top wall based on MD data, as well as the GM and CLL models predictions. Horizontal axis shows the number of applied Gaussian functions in the GM model ($1 < K < 1500$).

5.7. ASSESSMENT OF THE SAMPLING APPROACHES USED TO GENERATE GM DATA

To begin the assessment procedure, initially, all the ACs are computed using another approach. It is noteworthy to mention that all the reported ACs in this work are up to now calculated based on the correlation (Cor) approach (see Equation 1.11). This method was chosen in the first place since it is the only method that can be used to compute ACs for an isothermal walls system. In another common approach, ACs are computed as the ratio of the difference between incoming and outgoing gas molecules fluxes to that of the maximum possible flux when gas molecules are fully accommodated to the surface:

$$\alpha_q = \frac{\langle Q_I \rangle - \langle Q_R \rangle}{\langle Q_I \rangle - \langle Q_w \rangle} \tag{5.1}$$

where Q can represent various kinetic properties of gas molecules, such as its velocity in certain direction or a specific energy mode (e.g: E_{tr}, E_{rot}, E_{rot}) of it. The subscripts I and R refer to the incoming and outgoing fluxes, while w denotes the corresponding outgoing property when gas molecules are fully accommodated to the surface. The notation $\langle . \rangle$ indicates the ensemble average of the molecular property Q. The key advantage of the current approach, called the classical (Cla) approach in the remaining part of this work, over the correlation approach is to make it possible to compute standard deviation ($\triangle \alpha_q$) for the ACs. Since $\triangle \alpha_q$ denotes how dispersed the data is in relation to the mean, it can be used as a criterion to assess the performance of the sampling approaches used to generate data in the GSI models. Here, to calculate $\triangle \alpha_q$, the method proposed by P.

Spijker et al. [66] is employed. In this method, the entire data set is initially divided into M subsets of equal sizes. Subsequently, the mean and standard deviation for each subset is computed. Then, the mean values are used to compute the overall mean and the average α_q using 5.1. Finally, $\triangle\alpha_q$ is calculated as follows:

$$\triangle \alpha_q = \frac{|\langle Q_r\rangle - \langle Q_w\rangle|\triangle Q_i + |\langle Q_i\rangle - \langle Q_w\rangle|\triangle Q_r}{(\langle Q_i\rangle - \langle Q_w\rangle)^2} \tag{5.2}$$

where $\triangle Q_i$ and $\triangle Q_r$ are the standard deviations of the mean related to the corresponding incoming and outgoing kinetic properties, respectively.
The mean and standard deviations values of different ACs obtained from the GM and CLL scattering models, as well as the original MD data are presented in Figures 5.23 and 5.24, respectively.
Figure 5.23 reveals that the overall mean values of ACs related to different energy modes obtained from the GM scattering model ($\alpha_{tr-GM-Cla}$, $\alpha_{rot-GM-Cla}$, $\alpha_{tot-GM-Cla}$) perfectly match the MD results ($\alpha_{tr-MD-Cla}$, $\alpha_{rot-MD-Cla}$, $\alpha_{tot-MD-Cla}$). This observation differs from the trends previously shown in Figure 5.22, where noticeable deviations between ACs from the MD and GM results are reported. However, similar to the ACs obtained from the correlation approach, a considerable discrepancy between the CLL results ($\alpha_{tr-CLL-Cla}$, $\alpha_{rot-CLL-Cla}$, $\alpha_{tot-CLL-Cla}$) and MD results is also seen here.
Concerning the computed standard deviations, as it is illustrated in Figure 5.24, the standard deviations of different ACs obtained from the CLL ($\Delta\alpha_{tr-CLL}$, $\Delta\alpha_{rot-CLL}$, $\Delta\alpha_{tot-CLL}$) and MD results ($\Delta\alpha_{tr-MD}$, $\Delta\alpha_{rot-MD}$, $\Delta\alpha_{tot-MD}$) are in the same order of magnitude. Nevertheless, the GM scattering model results ($\Delta\alpha_{tr-GM}$, $\Delta\alpha_{rot-GM}$, $\Delta\alpha_{tot-GM}$) are one order of magnitude higher than the MD ones. To better understand the results demonstrated in the aforementioned figures, values of different ACs obtained for each case study based on different approaches (i.e., the correlation and classic approaches) are listed in Table 5.1. In this table the values presented between the parentheses for ACs based on the Cla approach are the computed standard deviations ($\Delta\alpha_q$) based on Equation 5.2. It is seen that the standard deviations based on the GM model results in the majority of the cases are even higher than the reported mean values. This observation indicates the high dispersion in the sampled data based on the GM model parameters.
Another crucial observation based on the MD results presented in Table 5.1 is the considerable difference between the energy ACs computed using the Cla and Cor methods. Such mismatch, which has also been reported previously [31, 49], is another indicator addressing the limitation of empirical scattering kernels, whose performance depends significantly on the values of applied ACs. While this behavior is also seen in the GM results listed Table 5.1, the energy ACs based on the CLL results are not affected by the computation methods.

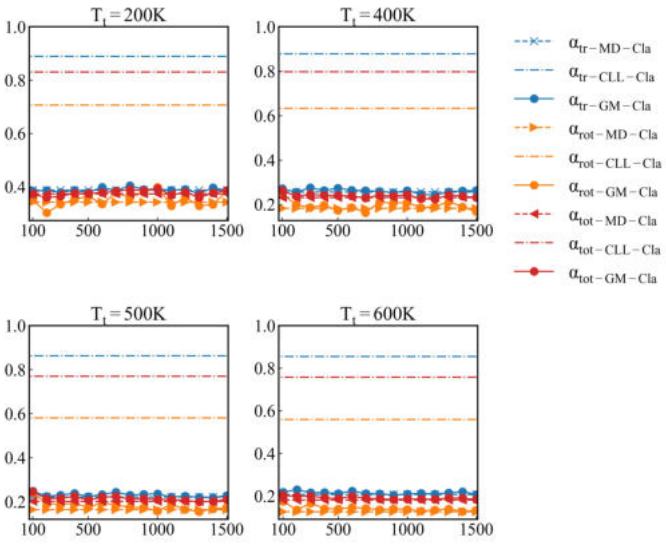

Figure 5.23: Variation of the overall mean of different accommodation coefficients computed with the classical approach for two parallel walls H_2-Ni system at the top wall based on MD data, as well as the GM and CLL models predictions. Horizontal axis shows the number of applied Gaussian functions in the GM model ($100 < K < 1500$).

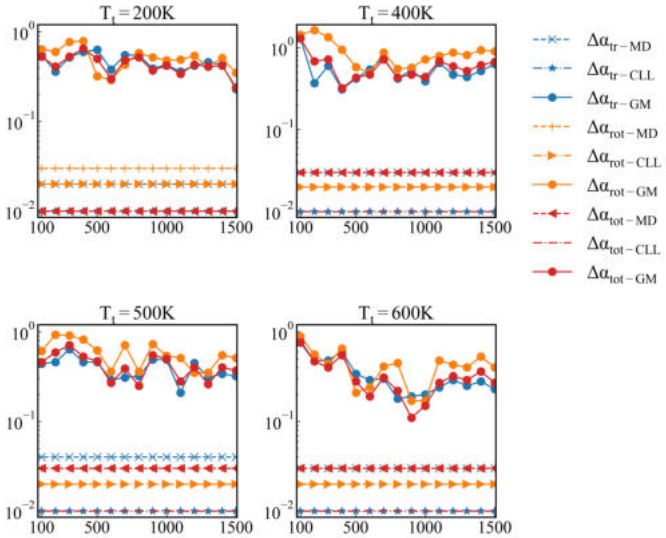

Figure 5.24: Variation of the standard deviation of the mean ($\Delta\alpha_q$) of different energy accommodation coefficients computed with the classical approach for two parallel walls H_2-Ni system at the top wall based on MD data, as well as the GM and CLL models predictions. Horizontal axis shows the number of applied Gaussian functions in the GM model ($100 < K < 1500$).

Table 5.1: Various energy accommodation coefficients obtained based on the MD results, as well as GM (K = 1500) and CLL scattering models for different case studies at the top wall. Cla: Classical approach. Cor: Correlation approach. In the case of the classical approach the numbers within the parenthesis denote the standard deviations ($\Delta\alpha_q$).

T_t	α_{tr-MD}		α_{tr-GM}		α_{tr-CLL}	
	Cla	Cor	Cla	Cor	Cla	Cor
200	0.387(0.02)	0.555	0.372(0.23)	0.586	0.890(0.01)	0.883
400	0.256(0.03)	0.479	0.248(0.63)	0.524	0.863(0.01)	0.870
500	0.221(0.04)	0.463	0.224(0.32)	0.519	0.847(0.01)	0.866
600	0.207(0.03)	0.442	0.206(0.23)	0.502	0.850(0.01)	0.863
	α_{rot-MD}		α_{rot-GM}		$\alpha_{rot-CLL}$	
	Cla	Cor	Cla	Cor	Cla	Cor
200	0.342(0.03)	0.705	0.362(0.35)	0.747	0.711(0.02)	0.704
400	0.182(0.02)	0.624	0.196(0.90)	0.669	0.639(0.02)	0.621
500	0.164(0.03)	0.589	0.156(0.51)	0.654	0.595(0.02)	0.589
600	0.126(0.03)	0.566	0.132(0.40)	0.621	0.566(0.02)	0.568
	α_{tot-MD}		α_{tot-GM}		$\alpha_{tot-CLL}$	
	Cla	Cor	Cla	Cor	Cla	Cor
200	0.373(0.01)	0.363	0.369(0.24)	0.432	0.832(0.01)	0.823
400	0.231(0.03)	0.278	0.231(0.67)	0.371	0.789(0.01)	0.787
500	0.202(0.03)	0.257	0.201(0.37)	0.367	0.764(0.01)	0.773
600	0.181(0.03)	0.237	0.182(0.27)	0.325	0.756(0.01)	0.764

As a further insight into the sampling procedure used to generate the GM model results, consider that the training data is an N by 11 matrix (see Figure 5.18). After training the GM model, its parameters ($\psi = \{\rho_i, \vec{\mu}_i, \Sigma_i\}\ \forall\ i$ in $\{1 \cdots K\}$) are utilized to generate the final results that are in the form of a new matrix, Ω, which has the size of $r \times 11$. Here, r is a positive real number that the user can specify. Matrix Ω includes the new incoming and outgoing kinetic properties (i.e., translational and rotational velocities) for all the studied temperatures. The order of the columns in Ω is the same as the initial training data matrix. Therefore, the reported temperatures in the first column of the matrix are used as the indicators to split Ω into the corresponding sub-matrices related to different surface temperatures (T_t). A similar sampling approach was also used for the Ar-Au system, discussed in Section 5.5. However, in that case, the ACs obtained from the GM model predictions and the original MD results (see Figure 5.16) were in good agreement. Therefore, we did not question the applied sampling approach. The most likely reason behind the observed deficiency of the applied sampling approach in the diatomic system is the higher dimensionality of the training data and having data with different physical meaning in the training data. This issue can also be realized by comparing the results presented in Chapters 2 and 3, in which the GM model was applied to the individual data sets for the Ar-Au and Ni-H_2 systems, respectively. Comparing the reported ACs obtained from the GM model and the original MD results in the aforementioned chapters, the ACs predicted by the GM for the Ar-Au were always in a better agreement with MD data. The reason why such a shortcoming of the sampling approach is more notable in the case of the generalized GM model for Ni-H_2 system is higher scattering in the training data caused by having different kinetic properties (i.e., the translational and rotational velocity components) in the collisional data. To be more specific, in the final data set including the collected collisional data at different T_t for the Ar-Au system, the standard deviations for all velocity components (i.e., columns 2 to 6 in Λ matrix shown in Figure 5.11) are around 3 Å/ps. On the other hand, for the H_2-Ni system the computed standard deviation related to the translational velocity components in the training data set (i.e., columns 2 to 4 and 7 to 9 in Λ matrix shown in Figure 5.18) are around 12 Å/ps. Besides, for the rotational velocity components (i.e., columns 5, 6, 10, and 11 in Λ matrix shown in Figure 5.18) the measured standard deviations are around 32 1/ps. Such a wider range of scattering makes it even more complex to properly predict the correlations between the individual incoming and outgoing components, as they are in the original MD data.

To overcome this problem, another sampling approach known as the conditional sampling approach [33] is examined. The essence of this approach is the same as other commonly used GSI models, such as the CLL scattering kernel. It means the scattering model is utilized to generate only the outgoing velocities. Here, based on the GM model parameters, instead of generating both the incoming and outgoing velocities and the surface temperature, the incoming velocities and the surface temperatures in the testing dataset are applied as the input to the GM scattering model. Afterward, just the corresponding outgoing velocities are generated based on the input data. Using this approach, instead of sampling an 11-dimensional matrix, we need to sample a 5-dimensional matrix, and this issue leads to less dispersion in the final predicted results. The main disadvantage of the current method to the previous one is its relatively high computational expense. For

instance, for a trained model with $K = 100$ Gaussian components, the sampling based on the previous method for all the studied T_t can be carried out within a couple of minutes. However, using the conditional method, generating the outgoing velocities corresponding to each T_t lasts between five to ten minutes.
In the conditional sampling method, assuming $\boldsymbol{x_I} = (T_t, \boldsymbol{v'}, \boldsymbol{\xi'})$ and $\boldsymbol{x_O} = (\boldsymbol{v}, \boldsymbol{\xi})$ as the incoming and outgoing velocity vectors in the training data, in the first step, the obtained mean vector and covariance matrix for each individual i multivariate Gaussian component ($\forall\ i$ in $\{1 \cdots K\}$) are required to be partitioned as follows:

$$\vec{\mu}_i = \begin{bmatrix} \boldsymbol{\mu}_{iO} \\ \boldsymbol{\mu}_{iI} \end{bmatrix},\ \Sigma_i = \begin{bmatrix} \Sigma_{iOO}\ \Sigma_{iOI} \\ \Sigma_{iIO}\ \Sigma_{iII} \end{bmatrix} \tag{5.3}$$

The conditional form of the GM scattering kernel can be written as:

$$P^{GM}(\boldsymbol{X_O}|\boldsymbol{X_I}) = \sum_{i=1}^{K} \tilde{\rho}_i(\boldsymbol{X_I})\, p_i(\boldsymbol{X_O}|\boldsymbol{X_I}) \tag{5.4}$$

where the new set of weights $\tilde{\rho}_i(\boldsymbol{X_I})$ is computed as:

$$\tilde{\rho}_i(\boldsymbol{X_I}) = \frac{\rho_i p_i(\boldsymbol{X_I})}{\sum_{i=1}^{K} \rho_i p_i(\boldsymbol{X_I})} \tag{5.5}$$

The marginal $(\boldsymbol{X_I})$ and the conditional $p_i(\boldsymbol{X_O}|\boldsymbol{X_I})$ distributions presented in equations 5.5 and 5.4 are calculated as [140]:

$$p_i(\boldsymbol{X_I}) = \int p_i(\boldsymbol{X_O}, \boldsymbol{X_I})\, d\boldsymbol{X_O} = N(\boldsymbol{X_I}|\boldsymbol{\mu}_{iI}, \Sigma_{iII}) \tag{5.6}$$

and

$$\begin{aligned} p_i(\boldsymbol{X_O}|\boldsymbol{X_I}) &= \frac{p_i(\boldsymbol{X_O}, \boldsymbol{X_I})}{p_m(\boldsymbol{X_I})} = N(\boldsymbol{X_O}|\boldsymbol{\mu}_{iO|I}, \Sigma_{iO|I}) \\ \boldsymbol{\mu}_{iO|I} &= \boldsymbol{\mu}_{iO} + \Sigma_{iOI}\Sigma_{iII}^{-}1(\boldsymbol{x_I} - \boldsymbol{\mu}_{iI}) \\ \Sigma_{iO|I} &= \Sigma_{iOO} - \Sigma_{iOI}\Sigma_{iII}^{-}1\Sigma_{iIO} \end{aligned} \tag{5.7}$$

The standard deviations of different energy ACs obtained from the GM scattering kernel based on the conditional sampling approach along with CLL, and the MD results are shown in Figure 5.25. It is seen that using the conditional sampling approach considerably reduces the standard deviations of the mean values of ACs obtained from the GM model. In fact, new values of $\triangle\alpha_q$ based on the GM model are in the same order of magnitude as the reference MD results. To give a better overview of the impact of this new sampling approach, the values of different energy ACs obtained from the GM model based on the conditional sampling, the CLL model, and the reference MD results are presented in Table 5.2. Comparing the GM results presented in this Table with the previously reported GM results in Table 5.1, it is seen that the energy ACs obtained from the Cor method based on GM results are not affected by the sampling approach. Besides, the overall means of the energy ACs based on the GM results are also very similar to the

previously reported results (see Table 5.1). On the other hand, it is seen that the conditional sampling approach leads to a significant reduction in the values of standard deviations related to different energy ACs based on the GM results. Therefore, it is deduced that depending on the method applied to compute the ACs, the sampling approach can have considerable impact on the AC results.
These results alongside the previously shown correlation graphs (e.g. see Figure 5.20) confirms the acceptable performance of the GM approach to derive a generalized scattering kernel for the H_2-Ni system.

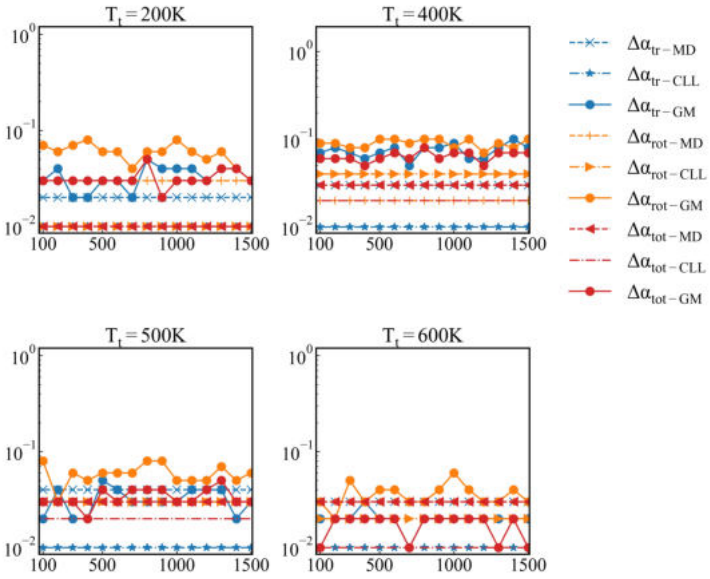

Figure 5.25: Variation of the standard deviation of the mean ($\Delta\alpha_q$) of different energy accommodation coefficients for two parallel walls H_2-Ni system at the top wall, computed with the GM approach based on conditional sampling approach , the CLL model, and MD simulations.

Table 5.2: Various energy accommodation coefficients obtained based on the MD results, GM scattering model based on conditional sampling ($K = 1500$), and CLL scattering model for different case studies at the top wall. Cla: Classical approach. Cor: Correlation approach. In the case of the classical approach the number within the parenthesis denotes the standard deviations ($\Delta\alpha_q$).

T_t	α_{tr-MD}		α_{tr-GM}		α_{tr-CLL}	
	Cla	Cor	Cla	Cor	Cla	Cor
200	0.387(0.02)	0.555	0.386(0.03)	0.592	0.890(0.01)	0.883
400	0.256(0.03)	0.479	0.263(0.08)	0.523	0.863(0.01)	0.870
500	0.221(0.04)	0.463	0.228(0.03)	0.524	0.847(0.01)	0.866
600	0.207(0.03)	0.442	0.207(0.02)	0.504	0.850(0.01)	0.863
	α_{rot-MD}		α_{rot-GM}		$\alpha_{rot-CLL}$	
	Cla	Cor	Cla	Cor	Cla	Cor
200	0.342(0.03)	0.705	0.383(0.03)	0.750	0.711(0.02)	0.704
400	0.182(0.02)	0.624	0.174(0.10)	0.688	0.639(0.02)	0.621
500	0.164(0.03)	0.589	0.167(0.06)	0.668	0.595(0.02)	0.589
600	0.126(0.03)	0.566	0.133(0.03)	0.629	0.566(0.02)	0.568
	α_{tot-MD}		α_{tot-GM}		$\alpha_{tot-CLL}$	
	Cla	Cor	Cla	Cor	Cla	Cor
200	0.373(0.01)	0.363	0.385(0.03)	0.437	0.832(0.01)	0.823
400	0.231(0.03)	0.278	0.231(0.07)	0.371	0.789(0.01)	0.787
500	0.202(0.03)	0.257	0.210(0.03)	0.369	0.764(0.01)	0.773
600	0.181(0.03)	0.237	0.183(0.01)	0.338	0.756(0.01)	0.764

6

A HYBRID GM-DSMC METHOD TO STUDY THE FOURIER THERMAL PROBLEM

In rarefied gas dynamics scattering kernels deserve special attention since they contain all the essential information about the effects of physical and chemical properties of the gas-solid surface interface on the gas scattering process. However, to study the impact of the gas-surface interactions on large scale behaviour of fluid flows, these scattering kernels need to be integrated in larger-scale models like Direct Simulation Monte Carlo (DSMC). In this work, the Gaussian mixture (GM) model, an unsupervised machine learning approach, is utilized to establish a scattering model for monoatomic (Ar) and diatomic (H_2) gases directly from Molecular Dynamics (MD) simulations data. The GM scattering kernel is coupled to a pure DSMC solver to study isothermal and non-isothermal rarefied gas flows in a system with two parallel walls. To fully examine the coupling mechanism between the GM scattering kernel and the DSMC approach, a one-to-one correspondence between MD and DSMC particles is considered here. Benchmarked by MD results, the performance of the GM-DSMC is assessed against the Cercignani-Lampis-Lord (CLL) kernel incorporated into DSMC simulation (CLL-DSMC). The comparison of various physical and stochastic parameters shows the better performance of the GM-DSMC approach. Especially for the diatomic system, the GM-DSMC outperforms the CLL-DSMC approach. The fundamental superiority of the GM-DSMC approach confirms its potential as a multi-scale simulation approach for accurately measuring flow field properties in systems with highly nonequilibrium conditions.

This chapter will be submitted to a peer-reviewed journal. Mohammad Nejad, S., Peter, F., Nedea, S., Frijns, A. and Smeulders, D.,. A hybrid Gaussian Mixture/DSMC approach to study the Fourier thermal problem.

6.1. INTRODUCTION

Rarefied gas dynamics is an active research topic in numerous cutting-edge engineering applications ranging from aerospace to biological applications [1]. At the design stage of such applications, numerical modeling is required to predict flow field properties and thermal energy transport between components to guarantee their optimal performance and lifetime. Depending upon the degree of rarefaction in the system, determined by the Knudsen number (Kn), different modeling approaches can be used to study the rarefied gas dynamics. In the case of low to moderate degree of rarefaction (Kn < 0.1) the continuum approach, based on the Navier-Stokes equations, is the most common simulation approach. On the other hand, for a highly rarefied gas (Kn ≥ 0.1), particle-based simulation techniques, such as Molecular Dynamics (MD) [12], Direct Simulation Monte Carlo (DSMC) [11], and lattice Boltzmann method (LBM) [10], are typically employed to determine flow field properties. Among these techniques, DSMC is the most commonly used one that has been successfully applied in a broad range of engineering applications [1]. Nevertheless, prescribing accurate boundary conditions at the gas-surface interface is a vital parameter to achieving reliable results in DSMC. In fact, going toward a higher degree of rarefaction, the complex physical interactions at the interface become even more dominant than the gas-gas interactions happening at the bulk of the fluid [4, 117]. Therefore, a detailed understanding of such interfacial nanoscale phenomena is of extreme importance in rarefied gas dynamics.

The diffuse reflection scattering model was the first boundary model used to describe the scattering process in rarefied gases [4]. This model assumes that the reflected gas molecules are fully accommodated with the adjacent solid surface, and their outgoing velocity distribution is determined by the Maxwellian distribution based on the wall temperature. However, later experimental studies [141, 142] revealed that the complete diffuse reflection assumption is not always valid. For example, at a very clean or high-temperature surface or under ultra vacuum conditions, gas molecules can experience near specular reflection. Consequently, to establish more comprehensive boundary models, researchers proposed various empirical scattering kernels that could anticipate both specular and diffuse reflections in specific rarefied gas flow applications [15–18, 21, 23–25, 119, 120]. The performance of these scattering kernels normally depends on several parameters, known as accommodation coefficients (ACs). In these scattering kernels, ACs are applied to quantify the accommodating level of different kinetic energy modes of gas molecules on a neighboring solid surface. Cercignani–Lampis–Lord (CLL) scattering kernel is one of the most employed scattering kernels that can be utilized to describe the scattering process for both monoatomic and diatomic gas molecules [17]. Despite the acceptable performance of the classical scattering kernels in the systems near equilibrium conditions, it was shown that these models are incapable of fully capturing the complex physical phenomena, such as a considerable temperature jump, happening in systems in a highly non-equilibrium situation [37, 100]. This shortcoming, alongside the lack of generality of such scattering kernels, raises the need for more elaborated gas scattering kernels.

As the most accurate particle-based simulation approach, MD simulation is a valuable tool for studying interfacial nanoscale phenomena. In an MD simulation, the interactions between individual gas molecules with neighboring solid molecules are modeled

deterministically. This fact makes MD computationally very expensive. Therefore, its application is usually restricted to the nanoscale level. Nevertheless, MD simulation is exploited in various approaches for investigating interfacial physics. Most commonly, MD is used to compute different ACs between a gas-solid pair. These ACs are then fed into one of the previously discussed empirical scattering kernels that are served as boundary conditions for high scale simulation approaches such as DSMC [143].

Other researchers developed hybrid simulation schemes combining the classical MD and DSMC approaches to study rarefied gas flow at mesoscale level [21, 143–146]. In these schemes, the simulation domain is decomposed into smaller regions. To benefit the precision of MD and the speed of DSMC at the same time, the first approach is applied in the vicinity of the solid surfaces, while the latter one is utilized in the bulk of the domain. Although these hybrid schemes are relatively faster than pure MD, since a large number of solid atoms must still be simulated in the MD part, these schemes are computationally expensive. In addition, there are considerable differences in the order of magnitude of time and space between MD and DSMC. Therefore, while applying these hybrid schemes, special attention must be taken into account to ensure particle and energy conservation at the coupling interfaces.

Another category of wall scattering kernels is nonparametric scattering kernels [31, 32, 42]. Unlike the classical empirical scattering kernels, no intermediate calibration based on ACs is needed, and MD data are used directly to construct these scattering kernels. These scattering kernels generally show better performance in the highly non-equilibrium situation compared to the classical scattering kernels [32]. As a first step to constructing these scattering kernels, single gas molecules with independent thermal velocities are beamed onto a specific solid surface in an MD simulation. Then, the discrete gas molecule trajectories, characterized by the system's initial condition, such as gas molecular velocity or surface conditions, are gathered in a database. Based on the relation between the prescattered and postscattered molecular velocity vectors in the MD dataset, a conditional probability density function is derived that can be used for sampling the reflected gas molecular velocity depending on its initial state. Liu et al. [42] developed such a scattering kernel and coupled it to a DSMC code to model hypersonic flow over a rounded wedge. Comparing the simulation results with another case study, in which the CLL scattering kernel was used as the boundary model, showed that the DSMC code based on their proposed model performs better in describing different flow properties. Nevertheless, in all these scattering kernels, the gas-gas interactions that can affect the reflected gas molecules properties in the early transition regime (0.1<Kn<1) are ignored, and these wall models cannot deal with adsorption-related problems.

Machine learning is another promising technique that can be used to establish a gas scattering kernel directly based on the collisional data obtained from MD simulations [33–35, 147]. As an example, in our previous works [34, 147], the Gaussian mixture (GM) approach, an unsupervised machine learning approach, was employed to construct a scattering kernel for monoatomic gases (Ar, He) interacting with the Au surface and diatomic gases (H_2, N_2) interacting with the Ni surface, respectively. While in the existing GSI models, only gas-wall interactions are included in the scattering model, in our approach, both gas-gas and gas-wall interactions are employed to construct the scattering model. This issue guarantees the capability of the model to deal with adsorption-related

problems in the case of systems in which a considerable number of gas molecules are adsorbed on the solid surface. As an example, the surface coverage for the N_2-Ni and H_2-Ni systems are 3.8 and 0.35 $1/nm^3$, respectively. Such a higher value of surface coverage in the case of N_2-Ni system considerably affects the scattering behavior of gas molecules. In our studies, a two parallel walls MD setup was used as the reference system to include the impact of both gas-wall and gas-gas interactions on the postcollisional behavior of gas molecules [34,147]. The physical and stochastic properties predicted by the GM scattering kernel were assessed against the CLL model results and the original MD data in the Fourier thermal and the combined Fourier-Couette flow problems. It was illustrated that the GM model outperforms the CLL model in both benchmark systems. The main advantage of the GM approach-driven scattering kernel over the empirical scattering kernels is that its performance does not rely on a finite number of parameters. In fact, we can adjust the number of fitting parameters to get the best performance from the scattering kernel. In addition, unlike nonparametric approaches whose extension to higher dimensions is limited [43], the GM scattering kernel can handle high-dimensional data sets rather straightforwardly. To the best of our knowledge, the scattering kernel based on the GM approach has not been coupled to a DSMC solver before to study the rarefied gas flow systems.

In this work, the main goal is to investigate the capability of the GM-driven scattering kernels to be used as a boundary model in an DSMC solver. Initially, based on a two parallel walls system, the interactions of Ar gas with Au surface and H_2 gas with Ni surface are studied using MD simulations. Two benchmark problems are considered: an equilibrium gas system confined between isothermal walls and a Fourier thermal problem. For the Ar-Au system, the pre and postcollisional translational velocities of Ar molecules are used for training the GM model. Whereas, for H_2-Ni system, both translational and rotational velocities are employed for the training purpose. Such consideration guarantees the model ability to anticipate the possible energy transfer between the translational and rotational modes at non-equilibrium conditions. Implementing the GM model on the MD collisional data in each case study, a conditional multivariate probability distribution is derived that can be used to generate the postcollisional velocities of gas molecules based on their precollisional states. The GM approach-driven boundary models for the Ar-Au and H_2-Ni are incorporated in pure DSMC simulations based on a one-to-one mapping between the corresponding MD and DSMC simulation setups. The DSMC simulations coupled to the GM scattering kernel (GM-DSMC) results are evaluated against the DSMC simulations coupled to the CLL scattering kernel (CLL-DSMC) and the reference MD solutions. The evaluation is performed based on different physical and stochastic criteria.

6.2. METHODOLOGIES

In this section the most relevant features of MD, the applied gas-surface interaction models, and DSMC related to the hybrid GM-DSMC approach are addressed. Further details of MD, gas-surface interactions models, and DSMC can be found in Ref. [12], Refs. [24, 33, 34, 43, 147] and Ref. [11], respectively.

Table 6.1: Lennard–Jones potential parameters

Atom pair	ϵ (eV)	σ (Å)
Ar-Au	$1.14e^{-2}$	3.819
Ar-Ar	$1.22e^{-2}$	3.35
H-H	$9.29e^{-4}$	1.421
H-Ni	$1.19e^{-2}$	2.016

6.2.1. MD SIMULATION

The exact particle trajectories are calculated in an MD simulation based on the interaction potentials and Newton's law. Thus, MD is considered as the most accurate method for modeling the scattering process and is used to provide benchmark solutions in this work. Our MD setup to study the 1-D Fourier thermal problem in a nanochannel consists of two infinite parallel walls at a distance, L_y, apart from each other and of gas molecules confined between these two walls (see Figure 6.1). Each wall is constructed with five layers of FCC planes. In the case of the Ar-Au system, each wall has a cross-sectional area of 10 nm × 10 nm, while for the H_2-Ni system, the cross-sectional area is 10.8 nm × 10.8 nm. In each wall, the outermost layer is constrained to prevent the translational motion of the wall in the normal direction. The distance between the two walls is fixed at L_y = 20 nm for the Ar-Au system, and at L_y = 30 nm for H_2-Ni system.

For the Ar-Au system the Knudsen number in the bulk and the gas reduced density (η) are Kn = 0.33 and η = 0.008, respectively. The reduced density is defined as $\eta = \pi n a^3/6$, where n is the number density and a is the particle diameter [76]. On the other hand, for H_2-Ni system Kn = 0.37 and η = 0.003. In addition, H_2 molecules considered as rigid rotors with a fixed bond length of 0.7414 Å [124].

In the both studied gas-solid pairs periodic boundary conditions are considered along x and z directions. The interactions between the solid Au and Ni atoms located in the walls are models using the corresponding embedded atom model (EAM) potentials developed by H. Sheng et al. [77] and S. Foiles et al. [125], respectively. For the Ar-Au system, the non-bonded gas-gas and gas-wall interactions are modeled by the Lennard-Jones (LJ) 12-6 potential. However, in the case of the H_2-Ni system, the non-bonded interactions are modeled using COMPASS force field [126], in which an LJ 9-6 function is applied to describe the interactions. All the gas-gas and gas-wall interatomic potential parameters utilized in this work are presented in Table 6.1.

The cutoff distances for gas-gas interactions are set at 2.5σ. Considering the gas-wall interactions, the cutoff distances (r_c in Figure 6.1) is set at 12 Å and 10 Å for the Ar-Au and H_2-Ni systems, respectively. In each MD simulation, after deposition of the target number of gas molecules between the solid walls, energy minimization is carried out by iteratively rearranging atom positions to eliminate the possible overlapping of the neighboring atoms. Afterward, each plate is connected to a Nose-Hoover (NVT) thermostat to maintain its temperature at the desired level. On the other hand, gas molecules are modeled in the microcanonical ensemble (NVE), and their temperature can change via the collision with other atoms in the box. In order to speed up the equilibration pro-

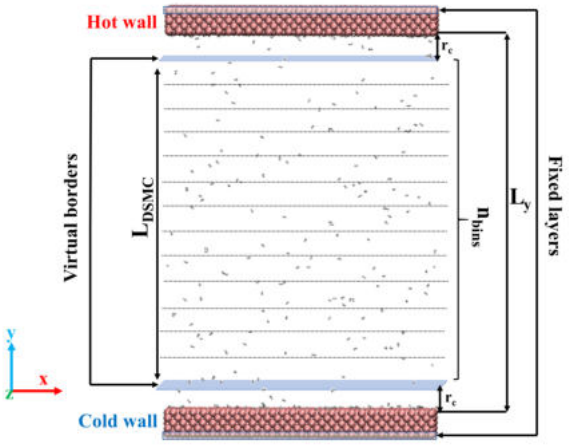

Figure 6.1: Schematic representation of the simulation model.

cess, the velocity components of gas molecules are initially sampled from a Gaussian distribution with a mean value of 0.0 and a standard deviation of $\sqrt{\frac{k_B T_a}{2m_g}}$, where T_a is the average value of the bottom and top plates temperatures, and m_g is the mass of the gas molecule. Each MD setup is equilibrated for 3 ns with a time step of 1 and 0.5 fs for the Ar-Au and H_2-Ni systems, respectively. After the complete thermalization of the system, the production run is started, which is proceeded for 25 ns for each MD system. All MD simulations are carried out using LAMMPS [129] package. The overall computational time for extracting the required MD data is around 13 hours running on a computer with 16 cores.

6.2.2. GAS-SURFACE INTERACTION MODELS

In rarefied gas flow simulations, gas-surface interaction models are used as boundary conditions. Such interactions usually are described in terms of a scattering kernel, $R(\boldsymbol{v}|\boldsymbol{v'})$, representing the probability density that an impinging gas molecule with velocity $\boldsymbol{v'}$ is reflected with velocity $\boldsymbol{v}$. In a general form, the boundary condition for the impinging gas molecular velocity distribution, $f(\boldsymbol{v'})$, can be expressed as [4],

$$v_n f(\boldsymbol{v}) = \int_{v'_n<0} |v'_n| R(\boldsymbol{v}|\boldsymbol{v'}) f(\boldsymbol{v'}) d\boldsymbol{v'}, \; v_n > 0 \tag{6.1}$$

where v'_n and v_n are the normal components of the impinging and reflected gas molecular velocity, respectively. In the case of a diatomic gas molecule, in addition to the translational velocity components of the center of mass $(\boldsymbol{v'}, \boldsymbol{v})$, the rotational velocity vectors $(\boldsymbol{\omega'}, \boldsymbol{\omega})$ need to be taken into account, as well. Therefore, for a diatomic gas molecules the scattering kernel is presented as $R(\boldsymbol{v}, \boldsymbol{\omega}|\boldsymbol{v'}, \boldsymbol{\omega'})$, and the probability density is denoted

by $f(\boldsymbol{v},\boldsymbol{\omega})$.

6.2.2.1. *CLL scattering model*

Among all the proposed empirical scattering kernels, the CLL model is used in this work. We chose this particular model since it is one of the most reliable scattering models for monoatomic and diatomic gases and implementing it in DSMC is rather straightforward [24, 37, 143]. For diatomic gas molecules, the CLL scattering kernel is given as [17]:

$$R^{CLL}(\boldsymbol{v},\boldsymbol{\omega}|\boldsymbol{v'},\boldsymbol{\omega'}) = \frac{2v_n}{\pi^2\alpha_t(2-\alpha_t)\alpha_n\alpha_{rot}} \exp\left[-\frac{[\boldsymbol{v_t}-\sqrt{1-\alpha_t}\boldsymbol{v'_t}]^2}{\alpha_t(2-\alpha_t)}\right] \times \exp\left[-\frac{v_n^2+(1-\alpha_n){v'_n}^2}{\alpha_n}\right] I_0\left[\frac{2(\sqrt{1-\alpha_n})v_n v'_n}{\alpha_n}\right] \times \exp\left[-\frac{(\boldsymbol{\omega}-\sqrt{1-\alpha_{rot}}\boldsymbol{\omega'})^2}{\alpha_{rot}}\right] \tag{6.2}$$

where α_t, α_n, and α_{rot} are the accommodation coefficients corresponding to the tangential momentum, normal translational kinetic energy, and rotational energy, respectively. $\boldsymbol{v_t}$ represents the tangential velocity vector and I_0 is the modified Bessel function of the first kind and zeroth order. The translational ($\boldsymbol{v'}$, $\boldsymbol{v}$) and rotational ($\boldsymbol{\omega'}$,$\boldsymbol{\omega}$) velocities are normalized by $\sqrt{\frac{2k_BT_w}{m_g}}$ and $\sqrt{\frac{2k_BT_w}{I}}$, respectively. Here, T_w describes the wall temperature and I is the mass moment of inertia of the diatomic gas molecule. It is noteworthy to mention that monoatomic gases do not possess rotational energy. Therefore, the last part in Equation 6.2, referring to the rotational velocities ($\boldsymbol{\omega'}$,$\boldsymbol{\omega}$), needs to be eliminated in the case of the Ar-Au system. As a result, only α_t, α_n are needed in the CLL model for the Ar-Au system. The accommodation coefficients needed in Equation 6.2 are computed using the approach proposed by P. Spijker et al. [49], in which the correlation between the corresponding impinging and outgoing kinetic properties (e.g., the normal translational energy or rotational energy) is applied to derive the relevant accommodation coefficients (e.g., α_n or α_{rot}).

The impinging and outgoing kinetic properties, known as the collisional data, that are required to compute the accommodation coefficients and training of the GM scattering model (discussed in the following part), are recorded at the virtual borders located at the distances r_c away from the walls (see Figure 6.1). The number of data points in different gathered collisional data sets varies between 100,000 to 150,000.

6.2.2.2. *GM scattering model*

Gaussian mixture model, a well-known unsupervised machine learning approach, can be exploited to derive a formalism describing gas-surface interactions directly based on MD simulation results. As it has also been addressed in the previous section, the main input required to train the GM model is the MD collisional data. In the case of the Ar-Au system, having only the translational degrees of freedom, the collisional data is a 6-dimensional dataset including the impinging and outgoing translational velocities of the

center of mass (COM) of Ar atoms ($v'_x, v'_y, v'_z, v_x, v_y, v_z$). On the other hand, for the H_2-Ni system, we must also account for the rotational degrees of freedom. Therefore, the impinging and outgoing rotational velocities are added to the training data. Here, the final training dataset is a 10-dimensional matrix ($v'_x, v'_y, v'_z, \omega'_1, \omega'_2, v_x, v_y, v_z, \omega_1, \omega_2$). Another parameter that directly influences the performance of the GM model and the computational cost of the training process is the number of applied Gaussian functions (K) in the model. This parameter must be specified adequately by the user to avoid the probable overfitting or underfitting. Here, a sensitivity analysis is performed to ascertain the optimal K. The details of it can be found in Appendix I. From this analysis, $K = 100$ and $K = 500$ are assigned as the number of Gaussian for the Ar-Au and H_2-Ni systems, respectively. Using these numbers of the Gaussian functions, the training of the GM model on a regular laptop computer takes around 3 and 40 minutes for the Ar-Au and H_2-Ni systems. The GM model manifests its best performance when all the components of the training data are normally distributed [34]. In the case of both gas-solid pairs considered in this work, except for the normal velocities (v'_y, v_y) following the Rayleigh distributions, the other components in the training data follow a Gaussian distribution. To obtain the same kind of distribution for the normal velocity components, initially, for each normal velocity pair (v'_y, v_y), an implicit pair ($-v'_y, -v_y$) was added to the dataset. Afterward, implementing the expression given in Equation 6.3 results in the Gaussian distributions for the normal velocity components [33].

$$T(\theta) = \sqrt{2\beta_I}\, erf^{-1}\left[1 - 2\exp\left(-\frac{\theta^2}{2\beta_I}\right)\right], \beta_I = \frac{k_B T_I}{m_g} \tag{6.3}$$

where T_I refers to the temperature of impinging gas molecules and can be calculated based on the average translational kinetic energy of the gas molecules. Applying such a transformation doubles the size of the data in the normal direction. Thus, to be consistent with the number of data points related to the other velocity components, half of the resulted new normal velocity components are added to the final training data.
After performing those mentioned above preprocessing on the MD collisional data (X), including the impinging ($\boldsymbol{x_I}$) and the outgoing ($\boldsymbol{x_O}$) velocity components, the GM model is employed to estimate the joint probability density of the collisional data. GM employs the superposition of multiple multivariate Gaussians to describe the probability density of the collisional data as:

$$P(\boldsymbol{x_O}, \boldsymbol{x_I}) = \sum_{i=1}^{K} \rho_i\, p_i(\boldsymbol{x_O}, \boldsymbol{x_I})$$

$$p_i(\boldsymbol{x_O}, \boldsymbol{x_I}) = N\left(\boldsymbol{X}|\vec{\mu}_i, \Sigma_i\right) \tag{6.4}$$

where $\rho_i, i = 1, 2, ..., K$ are the mixture component weights with the constraint that $\sum_{i=1}^{K} \rho_i = 1$, $\vec{\mu}_i$ is the mean vector, and Σ_i is the covariance matrix. These model parameters are determined using the expectation-maximization (EM) optimization algorithm [135]. Each component, p_i, of the mixture model is a multivariate Gaussian function given as:

$$N\left(\boldsymbol{X}|\vec{\mu}_i, \Sigma_i\right) = \frac{1}{(2\pi)^{M/2}|\Sigma_i|^{1/2}} exp\left[-\frac{1}{2}(\boldsymbol{X} - \vec{\mu}_i)' \Sigma_i^{-1} (\boldsymbol{X} - \vec{\mu}_i)\right] \tag{6.5}$$

where M refers to the dimensionality of the dataset. To properly incorporate the GM model results into a DSMC simulation a conditional PDF, $R^{GM}(\boldsymbol{x_O}|\boldsymbol{x_I})$, is required. To attain such Probability Density Function (PDF), in the first step, the obtained mean vector ($\vec{\mu}_i$) and covariance matrix ($\sum_i$) for each individual i, multivariate Gaussian component ($\forall\ i$ in $\{1\cdots K\}$) are partitioned as follows:

$$\vec{\mu}_i = \begin{bmatrix} \boldsymbol{\mu}_{iO} \\ \boldsymbol{\mu}_{iI} \end{bmatrix},\ \Sigma_i = \begin{bmatrix} \Sigma_{iOO}\ \Sigma_{iOI} \\ \Sigma_{iIO}\ \Sigma_{iII} \end{bmatrix} \tag{6.6}$$

where all the sub-matrices presented in $\sum_i$ are square matrices with the size of (3×3) and (5×5) for the Ar-Au and H_2-Ni systems, respectively. The GM scattering kernel is of the following form:

$$R^{GM}(\boldsymbol{X_O}|\boldsymbol{X_I}) = \sum_{i=1}^{K} \tilde{\rho}_i(\boldsymbol{X_I})\, p_i(\boldsymbol{X_O}|\boldsymbol{X_I}) \tag{6.7}$$

where the new set of weights $\tilde{\rho}_i(\boldsymbol{X_I})$ can be computed as:

$$\tilde{\rho}_i(\boldsymbol{X_I}) = \frac{\rho_i\, p_i(\boldsymbol{X_I})}{\sum_{i=1}^{K} \rho_i\, p_i(\boldsymbol{X_I})} \tag{6.8}$$

The marginal $p_i(\boldsymbol{X_I})$ and the conditional $p_i(\boldsymbol{X_O}|\boldsymbol{X_I})$ distributions demonstrated in Equations 6.8 and 6.7 are evaluated as [140]

$$p_i(\boldsymbol{X_I}) = \int p_i(\boldsymbol{X_O}, \boldsymbol{X_I})\, d\boldsymbol{X_O} = N(\boldsymbol{X_I}|\boldsymbol{\mu}_{iI}, \Sigma_{iII}) \tag{6.9}$$

and

$$\begin{aligned} p_i(\boldsymbol{X_O}|\boldsymbol{X_I}) &= \frac{p_i(\boldsymbol{X_O}, \boldsymbol{X_I})}{p_m(\boldsymbol{X_I})} = N(\boldsymbol{X_O}|\boldsymbol{\mu}_{iO|I}, \Sigma_{iO|I}) \\ \boldsymbol{\mu}_{iO|I} &= \boldsymbol{\mu}_{iO} + \Sigma_{iOI}\Sigma_{iII}^{-}1(\boldsymbol{x_I} - \boldsymbol{\mu}_{iI}) \\ \Sigma_{iO|I} &= \Sigma_{iOO} - \Sigma_{iOI}\Sigma_{iII}^{-}1\Sigma_{iIO} \end{aligned} \tag{6.10}$$

From each MD collisional dataset, 75% is applied for the training of the GM model, and the remaining data is utilized for validation purpose.
Finally, it should be pointed out that all the predicted velocity components by the GM model, in accordance with the initial distributions of the training data, follow Gaussian distributions. Therefore, to compare the GM model predictions with the original MD data, the normal velocity components are needed to be transferred back into the Rayleigh distribution using the following expression [33]:

$$Z(\theta) = \sqrt{-2\beta_O \ln\left[\frac{1}{2} - \frac{1}{2}\, erf\left(\frac{\theta}{\sqrt{2\beta_O}}\right)\right]},\ \beta_O = \frac{k_B T_O}{m_g} \tag{6.11}$$

where T_O refers to the temperature of reflected gas molecules from the surface and can be computed based on the average translational kinetic energy of the gas reflected gas molecules.

6.2.3. DSMC SIMULATION

DSMC simulation, a stochastic approach suitable for rarefied gases ($Kn > 0.1$), is considered the most popular particle-based simulation approach that has been successfully applied in a wide range of technological flow applications [1].
The DSMC method models gases using discrete particles which normally represent a large number of real molecules. A probabilistic collision method is used to solve the Boltzmann equation [11]. The DSMC particles are initially distributed randomly in the simulation box. In this work, a one-to-one correspondence between the DSMC particles and MD molecules is established. In addition to particles' locations, their initial velocities are assigned randomly from a Maxwellian distribution at the given temperature. For H_2 particles, their initial rotational energy are generated based on equipartition theorem according to the prescribed temperature. The DSMC domain is subdivided into n_{bins} in the y-direction with dimensions smaller than the particles' mean free path (see Figure 6.1). Within each cell, the macroscopic flow properties, such as density or temperature, are calculated through sampling the particle statistics over a number of independent simulation trials, presented by time averaging. Particles can freely travel across the DSMC domain and collide with other particles. One key difference between the MD and DSMC is indeed the essence of the interparticle collisions in these approaches. In an MD simulation, the collisions are deterministic, described by interatomic potential functions. Integrating these functions numerically, one can derive the molecules' exact postcollisional velocity and position. On the contrary, in a DSMC simulation, the interparticle collisions are stochastic, defined by simplified molecular interaction models. In this work, the variable hard sphere (VHS) model [11] is employed to describe the gas-gas collisions in the studied systems. The viscosity index, δ, for Ar and H_2 gases is set to 0.81 and 0.67, respectively [11]. Besides, in the case of H_2 molecules the Borgnakke-Larsen (BL) model [53] is applied to deal with the possible energy exchange between the translational and rotational energy modes. The rotational collision number [11], Z_r, required in the BL model computed using the empirical relation ($Z_r = 10480/T_a^{\delta}$) proposed by Boyd et al. [148] for H_2 gas. Assuming stochastic nature for the intermolecular collisions is one of the substantial advantages of DSMC over MD simulation that makes it significantly less intensive from a computational point of view.
In general, DSMC is a promising tool for obtaining accurate gas transport properties outside the gas-solid surface interaction layer, defined by the cutoff distance of the interatomic potential [145]. In fact, near surface effects, such as the presence of the adsorption layer caused by the attractive part of the interaction potential, can not be captured by the common DSMC tools. To overcome this shortcoming, as it was suggested in [145], the DSMC domain length, L_{DSMC}, is shortened a bit and covers only the distance considered between the virtual borders in the MD simulation as $L_{DSMC} = L_y\text{-}2r_c$ (see Figure 6.1). As a result, the exact number of DSMC particles, N_{DSMC}, is a bit less than the number of particles used in the MD simulation since the gas particles adsorbed on the surface are not included in the DSMC simulation. This issue will be elaborated on more in the following section.
In DSMC, at the gas-solid interface, the velocity vectors for the reflected particles are typically determined by sampling from a specific distribution function known as a scattering kernel. Here, the CLL and GM scattering kernels are employed at the virtual borders

(see Figure 6.1). Periodic boundary conditions are applied in the lateral directions (*x*,*z*). Another important issue is that in the case of H_2 molecules the rotational velocity vectors ($\boldsymbol{\omega'}, \boldsymbol{\omega}$) are used for the training of the GM model. However, in the currently available DSMC solvers particles are treated as the point-particles and they do not explicitly model rotational velocities but rather accounts for a scalar value for the rotational energy. To deal with this issue, as it was suggested by K. Yamamoto et al. [21], the impinging rotational energy is equally decomposed into two rotational velocities in the rotational directions ($\omega'_1 = \omega'_2 = \sqrt{E_{rot-I}/I_g}$).
In this work, DSMC simulations are carried out using dsmcFoam+ [149] solver, a DSMC solver implemented within the Open- FOAM [150] software framework. More details about integrating the GM scattering kernel into the DSMC solver are given in Appendix J.

6.3. RESULTS AND DISCUSSION

The performance of the GM and CLL scattering kernels employed as boundary conditions in DSMC simulations are investigated using the MD results as the reference solutions. For each gas-solid pair, two benchmarking case studies are considered: (i) isothermal Fourier thermal problem, in which the temperature of the bottom and top walls are set to 300 K; (ii) non-isothermal Fourier thermal problem, in which the temperature of the bottom wall is maintained at 300 K, while that of the top wall is fixed at 500 K. For each case study after performing the MD simulations, the required ACs in the CLL scattering kernel are computed. These ACs for the Ar-Au and H_2-Ni systems are listed in Tables 6.2 and 6.3 , respectively. Different physical and statistical criteria are employed to assess the performance of the applied scattering kernels, including the number density and temperature profiles between two walls, as well as the predicted velocity distributions at the gas-wall interfaces. In the case of the predicted number density and temperature by DSMC, the accuracy of the simulation results is determined by measuring the deviations of the DSMC results from the pure MD results. Considering x_{y-DSMC} as the predicted number density or temperature in a spacial bin by the DSMC approach coupled to the *y* scattering kernel, the deviations are computed by $\sqrt{(x_{y-DSMC} - x_{MD})^2}/x_{MD}$, where x_{MD} refers to the corresponding property obtained from the MD simulation.

Table 6.2: Tangential momentum (α_x) and normal energy (α_n) accommodation coefficients of the Fourier thermal problem for the Ar-Au system obtained from MD simulations. These values are from Table 3.2 (isothermal wall system, $S_w = 0$) and Table 3.3 (non-isothermal wall system, $S_w = 0$).

System	Wall	α_x	α_n
Isothermal	B/ T	0.876	0.910
Non-isothermal	B	0.886	0.910
	T	0.775	0.856

6.3.1. ISOTHERMAL AR-AU SYSTEM

Figure. 6.2 shows the density and temperature profiles of the isothermal Fourier problem for Ar-Au system obtained from the DSMC simulations combined with GM and CLL

Table 6.3: Tangential momentum (α_x), normal energy (α_n), and rotational energy (α_{rot}) accommodation coefficients of the Fourier thermal problem for the H_2-Ni system obtained from MD simulations. These values are from Table 4.2.

System	Wall	α_x	α_n	α_{rot}
Isothermal	B/ T	0.958	0.758	0.666
Non-isothermal	B	0.928	0.724	0.600
	T	0.923	0.730	0.600

scattering models alongside the pure MD results. As indicated in the previous section, gas-wall interaction zones need to be excluded to have a fair comparison between the MD and DSMC results. The reason is that DSMC cannot predict gas transport properties in these areas. In all the case studies in this work, looking at the number density profile obtained from the MD simulations, the number of the adsorbed gas molecules on the surfaces, N_{ads}, are excluded from the total number of the MD particles, N_{MD}, resulted in $N_{DSMC} = N_{MD} - N_{ads}$. As an example, in the present case study $N_{MD} = 800$. Based on the number density profile (see Figure 6.2a), the total number of Ar atoms adsorbed at the bottom and top walls are about $N_{ads} = 225$, which leads to $N_{DSMC} = 575$.

From Figure 6.2 it is deduced that the DSMC results using both the CLL and GM scattering models are in good agreement with the original MD results. To be more specific, the average deviations of the predicted number densities by the GM-DSMC and CLL-DSMC are 0.4% and 0.5%, respectively. While for the temperature, the deviations are 0.2% on average.

The correlation between the impinging and outgoing velocity components, as well as the PDF of the outgoing velocity components obtained from the applied simulation approaches for the isothermal Ar-Au system, are depicted in Figure 6.3. It is seen that in this case study, which resembles the fully equilibrium condition in the system, the velocity distributions predicted by the GM-DSMC and CLL-DSMC approaches match well the MD data.

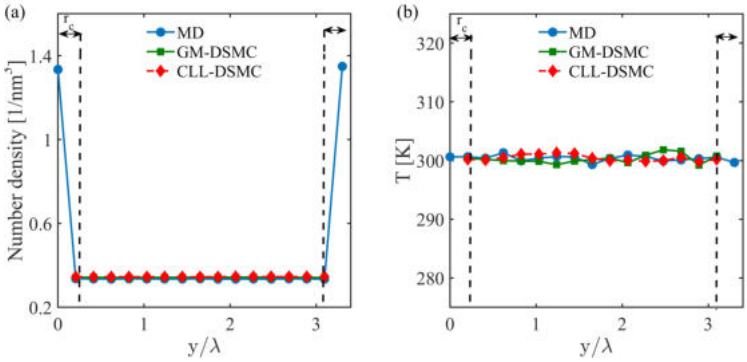

Figure 6.2: Variation of the macroscopic quantities of the isothermal Fourier problem for Ar-Au system obtained from the pure MD simulation and DSMC simulations combined with the GM and CLL scattering models. (a) number density, (b) temperature. r_c=12 Å.

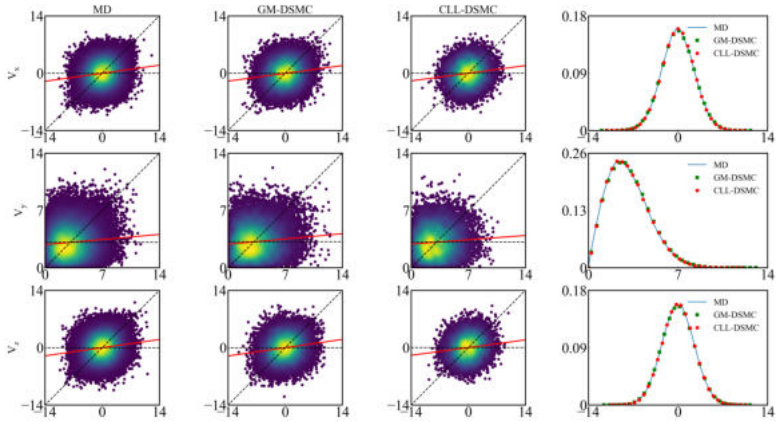

Figure 6.3: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in [Å/ps] of the isothermal Fourier problem for Ar-Au system at the bottom wall. The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data, its slope infers: 1-AC. In the last column the corresponding PDF for the reflecting particles are shown.

6.3.2. NON-ISOTHERMAL AR-AU SYSTEM

In Figure 6.4, the variation of the local number density and temperature observed in the MD simulation and the predicted trends by the GM-DSMC and CLL-DSMC of the non-isothermal Fourier problem for Ar-Au case study are presented. In general, having a higher temperature at the top wall causes less number of atoms to be adsorbed at the surfaces in this case (i.e., N_{ads}= 222, N_{DSMC}=578). Regarding the number density (see Figure 6.4a), similar to the previous case study, the predicted results by DSMC incorporating both scattering models are consistent with the MD data. Here, the average deviations of the predicted results are 0.6% for both GM-DSMC and CLL-DSMC approaches. In the case of the temperature profile (see Figure 6.4b), the trend predicted by the GM-DSMC in the bulk of the domain matches well the pure MD results. Here, the deviations are 0.2%. On the other hand, in most parts of the simulation domain, the results predicted by the CLL-DSMC method deviate from the MD results, observing the highest deviation close to the top wall, which is 4%.

Another observation is the noticeable temperature jump between the consecutive bins located at the beginning and end of the simulation domain. This observation is in line with the previously observed temperature jump adjacent to the solid surface induced by the strong gas-wall interactions within the potential cutoff distance [83].

The scattering results obtained from the MD, GM-DSMC, and CLL-DSMC approaches at the bottom and top walls are presented in Figures 6.5 and 6.6, respectively. It is seen that the velocity clouds predicted by the GM-DSMC and CLL-DSMC approaches are very similar to the MD results. Looking at the PDFs of the outgoing velocities in the x and z-directions at both walls, some discrepancies between the CLL-DSMC and MD results around the peak of the graph are seen. Nevertheless, the results from the GM-DSMC approach are always in good agreement with the MD results.

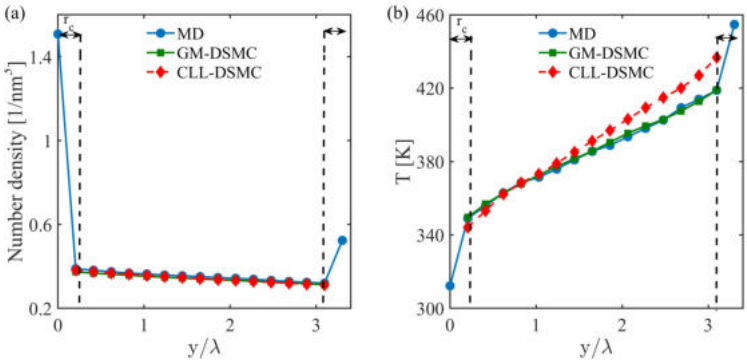

Figure 6.4: Variation of the macroscopic quantities of the non-isothermal Fourier problem for Ar-Au system obtained from the pure MD simulation and DSMC simulations combined with the GM and CLL scattering models. (a) number density, (b) temperature. r_c= 12 Å.

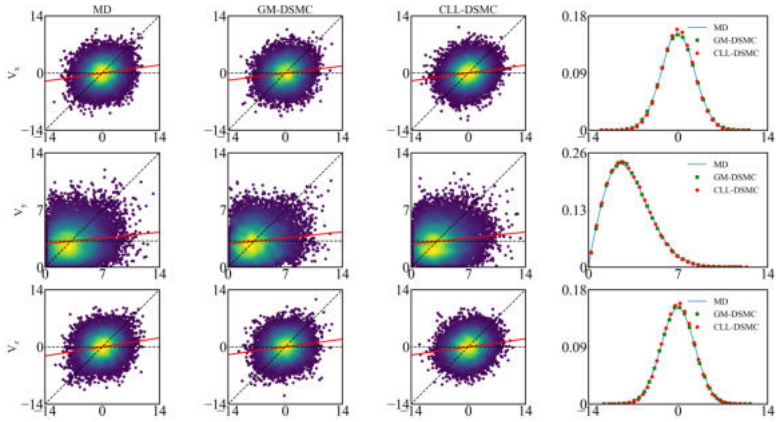

Figure 6.5: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in [Å/ps] of the non-isothermal Fourier problem for Ar-Au system at the bottom wall. The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data, its slope infers: 1-AC. In the last column the corresponding PDF for the reflecting particles are shown.

6.3.3. ISOTHERMAL H_2-NI SYSTEM

Variations of the number density and temperature in the case of the isothermal Fourier problem for the H_2-Ni system are shown in Figure 6.7. Comparing the number density profile obtained from the MD simulation with the one for the isothermal Ar-Au system (see Figure 6.2a), presence of the weaker adsorption layer is seen in the current case study. This is manifested through the relatively smaller difference between the number densities of the consecutive bins near the walls in the current case than the isothermal Ar-Au system. It is seen that for the isothermal H_2-Ni system the measured number densities in all bins are in the same order of magnitudes, while for the isothermal Ar-Au

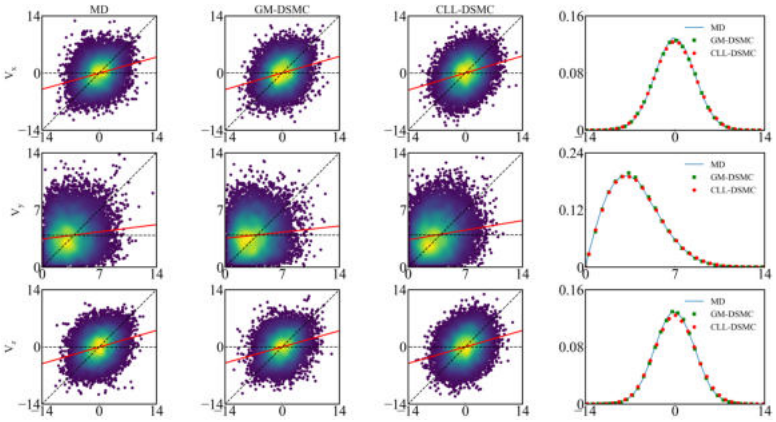

Figure 6.6: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in [Å/ps] of the non-isothermal Fourier problem for Ar-Au system at the top wall. The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data, its slope infers: 1-AC. In the last column the corresponding PDF for the reflecting particles are shown.

system the number densities near the walls are one order of magnitude higher than the ones in the bulk. This outcome is caused by significantly higher mass of the Ar atom compared to the H_2 ($m_{Ar} \approx 20 m_{H2}$). Regarding the number of particles used in the MD and DSMC simulations of the H_2-Ni system, N_{MD}=900 particles were used in the MD simulation. It was observed that around 100 particles were trapped in the gas-surface interaction zones during the MD simulation. Therefore, N_{DSMC} = 800 particles were considered in the DSMC simulation.

Going back to the predicted trends of the number density and temperature obtained based on different approaches in this case study, the results of GM-DSMC and CLL-DSMC agree with MD data. Here, the deviations of the predicted number densities by the GM-DSMC and CLL-DSMC are 0.6% and 0.4%, respectively. On the other hand, the temperature results of the DSMC simulations coupled with the GM and CLL scattering models on average deviate around 0.2% from the MD results.

The correlation plots and PDFs for different translations velocity components and energy modes of the isothermal Fourier problem for H_2-Ni system are presented in Figure 6.8. It is shown that in this case study, there is a good agreement between the correlation plots and PDFs of the partial translational velocity components, and the rotational energy mode. However, E_{tr} and E_{tot} correlation clouds of the pure MD and GM-DSMC are narrower than the CLL-DSMC approach, which since $E_{tot} = E_{tr} + E_{rot}$ the mismatch in E_{tr} induced the mismatch in the results for E_{tot}.

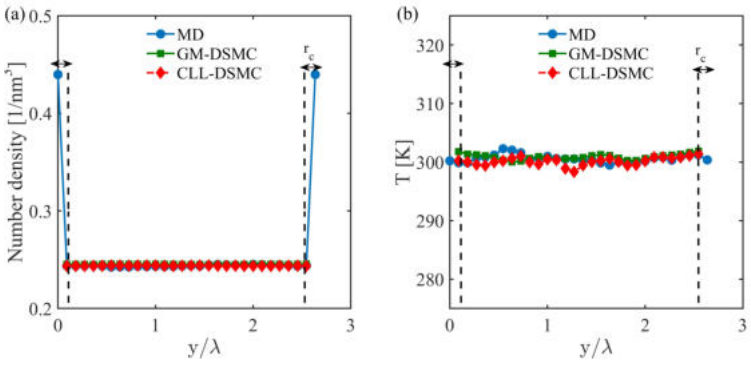

Figure 6.7: Variation of the macroscopic quantities of the isothermal Fourier problem for the H_2-Ni system obtained from the pure MD simulation and DSMC simulations combined with the GM and CLL scattering models. (a) number density, (b) temperature. r_c=10 Å.

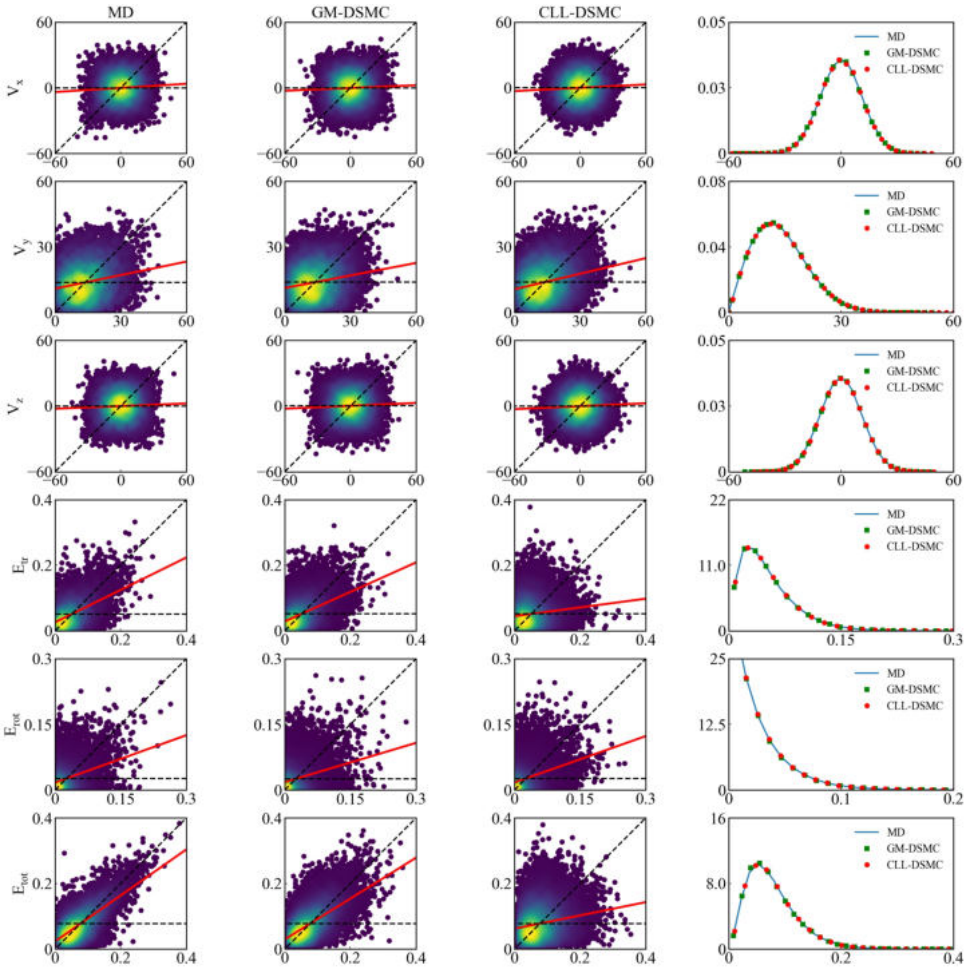

Figure 6.8: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the isothermal Fourier problem for the H_2-Ni system at the bottom wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data, its slope infers: 1-AC. In the last column the corresponding PDF of translational velocity components and energy modes for the reflecting particles are presented.

6.3.4. NON-ISOTHERMAL H_2-NI SYSTEM

Figure 6.9 shows the number density and temperature profile in the case of the non-isothermal Fourier problem for the H_2-Ni system. First of all, from the MD simulation it is realized that N_{ads}= 84 H_2 are adsorbed on the surfaces. Therefore, N_{DSMC}=806 particles are used in the DSMC simulation. Regarding the number density variation, it is observed that within the bulk of the simulation domain, the GM-DSMC approach results match well the MD data. Here, the deviations of the GM-DSMC results are around 0.6% on aver-

age. However, unlike all the previously investigated ones, a notable discrepancy between the density profiles obtained from the MD and CLL-DSMC is observed in the current case study. In this case, the highest deviation is measured near the top wall, which is 8%. In Figure 6.9b, it is depicted that the predicted temperature profiles based on both GM-DSMC and CLL-DSMC approaches deviate from the reference MD results. Nevertheless, the GM-DSMC approach still outperforms the CLL-DSMC approach. Herein, the highest deviations based on the GM-DSMC approach are measured at the first and last bins, which are 2% and 1%, respectively. On the other hand, using the CLL-DSMC approach, the deviations on the same bins are 9% and 10%, respectively.
The scattering results at the bottom wall of the non-isothermal Fourier problem for the H_2-Ni system are shown in Figure. 6.10. It is seen that the correlations plots of the translation velocity components obtained from the GM-DSMC and CLL-DSMC are in good agreement with the MD results. However, while the correlation graphs for V_x and V_z components obtained from the CLL-DSMC resemble perfect ellipsoid, the MD results look more skewed around the diagonal line. This observation, which can be perfectly captured by the GM-DSMC approach, indicates that many gas molecules with high velocity experience almost specular reflection. In addition, the PDFs of the outgoing velocity components predicted by the GM-DSMC approach are a perfect match, while the CLL-DSMC predictions deviate from the MD results around the peak value of the PDF plots. Looking at the scattering results related to the different energy modes, except for the rotational energy mode, in the other energy modes, the results from the CLL-DSMC deviate from the MD results. Nevertheless, the GM-DSMC results are in better agreement with the MD results. Except for the PDF of rotational energy mode, the peak values of all PDFs predicted by the CLL-DSMC approach are higher than MD and GM-DSMC results. A higher peak implies a lower temperature. Therefore, the CLL-DSMC overestimates the degree of H_2 accommodation to the surface. This issue is also seen in the correlation graphs for E_{tr} and E_{tot}, in which the slopes of the red lines, indicating 1-AC, are smaller for CLL-DSMC in comparison with the MD and GM-DSMC results. This leads to a higher AC for the CLL-DSMC method, and therefore to a lower temperature jump near the walls. The relatively large mismatch between the reference MD and the CLL-DSMC results in this case study is caused by the small molecular weight of H_2 molecules and the small size of the channel. In the MD simulation, the light H_2 molecules move very fast between the two plates causing the temperature in different parts of the bulk to converge towards the average value of the bottom and top plates temperatures ($T_a = 400K$). Based on the incoming and outgoing collisional data recorded at each wall, we can compute the incoming and outgoing gas temperatures. For the non-isothermal H_2-Ni system based on the MD results, the incoming and outgoing temperatures at the bottom wall are $T_{\text{in-MD-b}}$ = 403 K and $T_{\text{out-MD-b}}$ = 378 K, respectively. On the top wall, the incoming and outgoing gas temperatures are $T_{\text{in-MD-t}}$ = 385 K and $T_{\text{out-MD-t}}$ = 411 K. These temperatures are closer to the MD values. Therefore, we expect that mismatch between the MD and CLL-DSMC would decrease as we increase the system size. Our hybrid GM-DSMC model can deal with this, while the CLL scattering model can not anticipate such behavior. Next, the performance of the CLL model also highly depends on the values of applied ACs. As an example, Spijker et al. [66] computed ACs for the Ar-Pt system based on the classical and correlation approaches. The obtained ACs from the correlation approach

were slightly higher than classical approach. In their study, they showed a significant difference in the CLL scattering model results, based on the computed ACs. In this work, the obtained ACs (see Table 6.3), especially the tangential momentum AC, are high, indicating a large degree of accommodating the gas molecules to the neighboring surfaces. The scattering results at the top wall of the non-isothermal H_2-Ni system are shown in Figure 6.11. Comparing the results obtained from the MD, GM-DSMC, and CLL-DSMC, similar trends to the bottom wall can also be seen here. For E_{rot}, the results of GM-DSMC and CLL-DSMC are consistent with the MD data. For the other energy modes and translational velocity components, the GM-DSMC results are always in better agreement with MD results than the CLL-DSMC results. Looking at the PDFs, except for E_{rot}, in the other components, the peak value of the PDF predicted by the CLL-DSMC is lower than MD and GM-DSMC results. It means CLL-DSMC predicts a higher temperature for reflected gases. This observation, alongside underestimating the outgoing temperature at the bottom wall, indicates that the CLL-DSMC is not able to predict the actual temperature jump happening at the gas-solid interfaces. Such a conclusion can also be derived from Figure 6.9b, in which the predicted temperatures by the CLL-DSMC near the cold and hot walls are respectively lower and higher than the reference MD results.

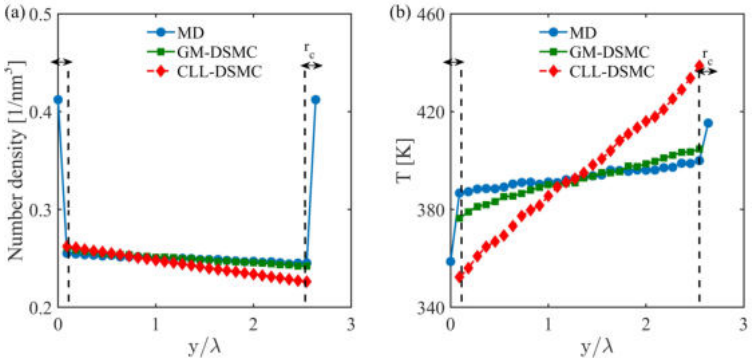

Figure 6.9: Variation of the macroscopic quantities of the non-isothermal Fourier problem for H_2-Ni system obtained from the pure MD simulation and DSMC simulations combined with the GM and CLL scattering models. (a) number density, (b) temperature. r_c= 10 Å.

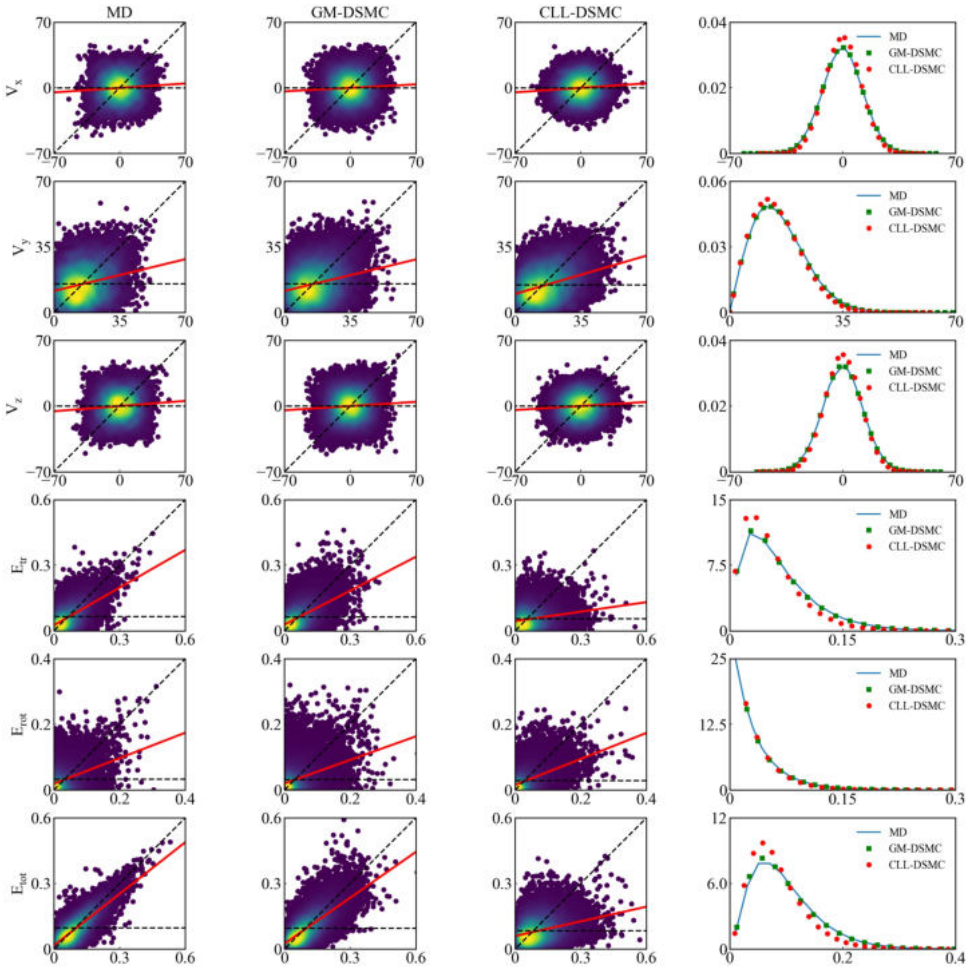

Figure 6.10: Correlations between incoming (horizontal-axis) and outgoing (verical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the non-isothermal Fourier problem for H_2-Ni system at the bottom wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data, its slope infers: 1-AC. In the last column the corresponding PDF of translational velocity components and energy modes for the reflecting particles are presented.

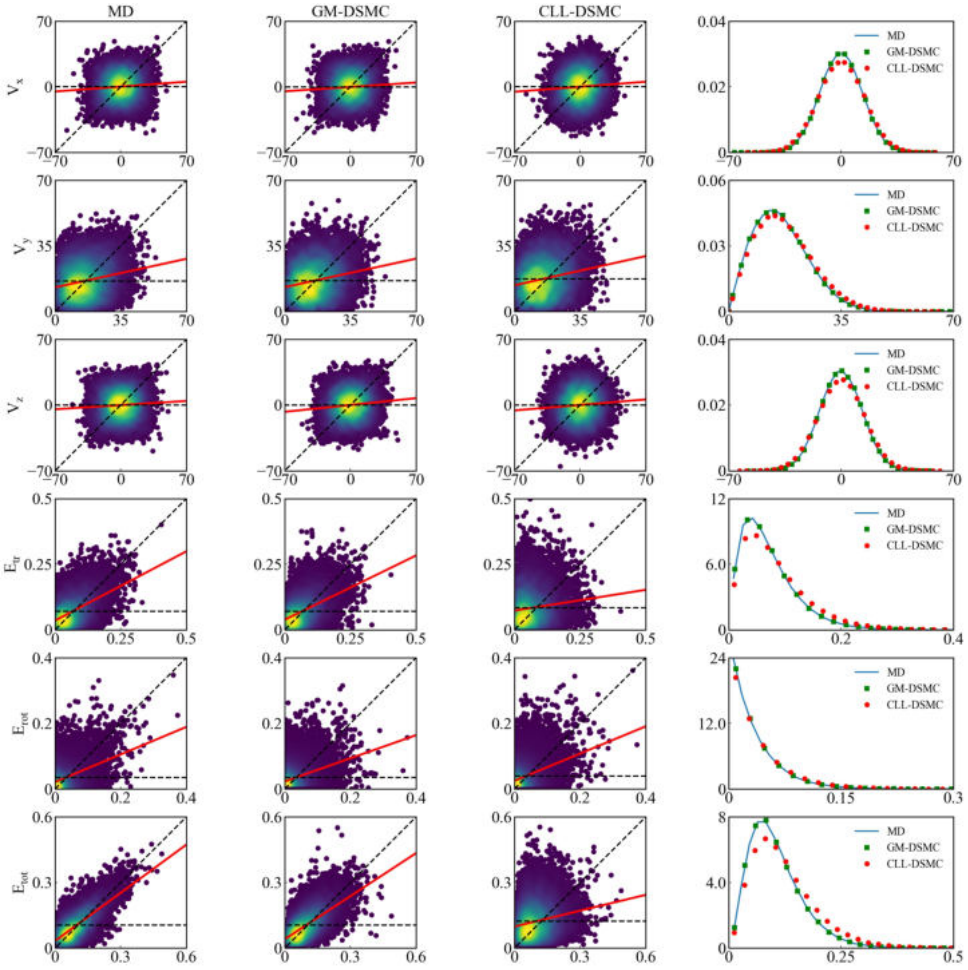

Figure 6.11: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the non-isothermal Fourier problem for the H_2-Ni system at the top wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data, its slope infers: 1-AC. In the last column the corresponding PDF of translational velocity components and energy modes for the reflecting particles are presented.

6.4. CONCLUSIONS

DSMC simulation, as the most common particle-based simulation technique, can be applied to derive precise solutions of gas transport properties outside the gas-wall interaction layer based on the accurate velocity distribution function of the reflected gas molecules provided by MD simulation. In this work, MD simulations on monoatomic and diatomic gases-surface interactions under thermal equilibrium and non-equilibrium conditions are carried out. Using the MD collisional data, including the pre and postscattered molecular gas velocities from each case study, an unsupervised machine learning approach, called the GM model, is employed to construct a gas-surface boundary model. The GM boundary model is in the form of a conditional multivariate PDF that can reproduce the postscattered gas molecular velocities based on their initial imposed velocities. The GM scattering kernel is incorporated in a pure DSMC simulation as a boundary condition to study isothermal and non-isothermal Fourier problems. In each case study, the performance of the proposed model (GM-DSMC) is assessed against the DSMC simulations based on the CLL scattering kernel (CLL-DSMC) using MD results as the reference solution. Comparing different physical (e.g., temperature field) and stochastic (e.g., velocity and energy distributions) parameters obtained from the applied simulation approaches confirm the superiority of the GM-DSMC approach.

These results pave the way to develop a multiscale simulation scheme that can accurately measure different flow field properties by efficiently including the microscopic non-continuum phenomena into the DSMC simulation.

7

CONCLUSIONS AND OUTLOOK

7.1. CONCLUSIONS

In this work, Molecular Dynamics (MD) simulation, as the most accurate particle-based simulation technique, is employed to achieve a fundamental understanding of the physics of the scattering process of rarefied gases from solid surfaces. In such simulations, the pair potential used to describe the gas-solid interactions is among the most important parameters affecting the MD results. Performing Quantum Mechanics (QM) calculations is the most reliable approach for deriving such pair potential functions. To investigate this issue, the energy and momentum accommodation coefficients of noble gases (Ar and He) interacting with the Au surface were computed using MD simulations based on pair potential functions obtained from different approaches. Different mixing rules, as well as existing ab-initio computations, which is a QM calculation technique, were employed to derive the coefficients of the pair potential function applied in MD simulations. Comparing the obtained MD results for accommodation coefficients with the empirical and numerical values in the literature reveals that the interaction potential based on ab-initio calculations is the most reliable one for computing accommodation coefficients. The study also reveals that the gas–gas interaction is an important aspect that needs to be taken into account in the transient Knudsen regime since it leads to an enhancement in obtained accommodation coefficients compared to values obtained via the molecular beam approach. Therefore "full MD"simulations are preferred for deriving accommodation coefficients.

Although MD simulations can provide a detailed understanding of rarefied gas flow field properties, pure MD simulations cannot be employed for macroscopic engineering problems. This is due to the considerably high computational cost of MD simulations. In the most common approach, MD simulation is used to derive different accommodation coefficients. These coefficients are fed into various empirical scattering kernels, such as the Cercignani–Lampis–Lord (CLL) model, that can be employed as boundary conditions for coarse-grained particle-based simulation techniques such as DSMC to compute flow field properties. However, in the systems encountered with highly nonequilibrium or complex gas flow conditions, empirical scattering models that are based on a limited number of constant parameters cannot completely describe the reflection of gas molecules at the gas-solid interface.

Therefore, the Gaussian Mixture (GM) approach, an unsupervised machine learning technique, is utilized to construct a robust gas scattering model based on the MD simulations results. The GM approach is a parametric approach. However, the number of parameters in the model is not limited, and it can be adjusted to get the best performance from the model. Based on a two parallel walls MD simulation setup, the GM approach is used to derive the scattering kernels for monoatomic gases (Ar, He) interacting with Au surfaces and diatomic gases (H_2, N_2) interacting with Ni surfaces. The entire MD collisional data is applied to train the GM scattering model. In the case of monoatomic gases, this means the pre and postcollisional translational velocity components of the gas atoms. For diatomic gas molecules, this refers to the pre and postcollisional translational and rotational velocity components of gas molecules. Such consideration guarantees the capability of the model to predict the possible interplay between the translational velocity components as well as the possible energy exchange between the translational and rotational energy modes that can happen in the case of

highly nonequilibrium systems. In addition, including gas-wall and gas-gas interactions in the performed MD simulations makes the GM scattering kernel a suitable model for dealing with adsorption-related problems. Benchmarked with the original MD results, the performance of the GM scattering kernel is compared with the CLL scattering model in different thermal problems that are commonly faced in rarefied gas flow systems, such as the Fourier thermal problem, the Couette flow problem, and the combined Fourier-Couette flow problem. The assessment is carried out in terms of different physical and statistical parameters of the system and shows the superiority of the GM scattering kernel.

The observed promising performance of the GM scattering kernel for both monoatomic and diatomic gases confirms its potential to derive a generalized boundary condition for rarefied gas systems. To investigate this matter, the surface temperature, one of the important factors determining the behavior of gas molecules after reflecting from the surface, is added to the training data used to construct the GM scattering kernel. Here, two parallel walls systems: Ar-Au and H_2-Ni are chosen as the case studies. For each gas-solid pair, while the temperature of the bottom wall (T_b) is kept constant, the temperature of the top wall changes in the range of $200 < T_t < 600$ with steps of 100 K. The collisional data obtained at the top wall of independent MD simulations are gathered and utilized to train the GM scattering model at once. Using the MD results as the reference, the performance of the generalized GM scattering kernel is compared with the CLL scattering kernel in terms of predicting different accommodation coefficients, as well as the distributions of different velocity and energy modes. It is observed the results predicted by the GM scattering kernel are always in better agreement with MD results.

In the final step, the GM scattering kernel is incorporated in a Direct Simulation Monte Carlo (DSMC) solver to study isothermal and non-isothermal rarefied gas flows in a system with two parallel walls. Two gas-solid pairs are studied: Ar-Au and H_2-Ni systems. To fully investigate the coupling mechanism between the GM scattering kernel and the DSMC approach, a one-to-one mapping between the corresponding MD and DSMC simulation setups is considered. The DSMC simulations coupled to the GM scattering kernel (GM-DSMC) results are assessed against the DSMC simulations coupled to the CLL scattering kernel (CLL-DSMC) and the original MD solutions. The assessment is carried out based on physical (e.g., temperature field) and stochastic (e.g., velocity and energy distributions) parameters and reveals the superiority of the GM-DSMC approach.

7.2. OUTLOOK

In this work, MD simulation combined with machine learning technique is employed to develop a computational tool to study rarefied gas flow systems. Next to the experimental tools, the proposed computational tool in this work can be applied to gain a fundamental understanding of the complex physics of rarefied gases-surface interactions.

The main focus of this work is on Extreme Ultraviolet Light (EUV) Photolithography machines used in the semiconductor industry. In these machines H_2 gas is typically applied for cooling and cleaning purposes [151]. Therefore, the H_2-Ni system was one of the case studies in this work to exploit the interaction between H_2 and solid surfaces in terms of momentum and energy exchange. To describe the interactions between H and

Ni atoms in this work, the sixth-power Waldman-Hagler combining rules [128] is employed. However, as addressed in Chapter 2, QM computations is the most reliable technique to derive the pair potential functions. Therefore, the first recommendation would be employing QM techniques, such as ab-initio or density functional theory (DFT), to derive more accurate force fields between H atoms of H_2 molecules and desired solid surfaces. However, this task is very challenging for H_2 molecules. Such complexity arises from the low polarizability of H_2 molecules caused by the presence of only one valence electron [152], as well as the nuclear quantum effects, such as tunneling and nuclear delocalization [153].

In this work, the surface temperature is used as the only system parameter to construct a generalized scattering kernel. However, other important parameters, such as gas temperature, surface roughness, and surface speed can also affect the scattering process of gas molecules from a surface. These parameters can also be added to the training data of the GM scattering kernel to derive a more comprehensive scattering kernel.

This work proposes a direct coupling between the GM scattering kernel and DSMC. This has been carried out to represent the fundamental capability of the GM scattering for being integrated into a DSMC solver. The next step would be to exploit the capabilities of the methodology for indirect coupling where the coarse-graining capability of the DSMC technique is employed. Doing such results in a multiscale simulation scheme that can accurately measure different flow field properties by efficiently including the microscopic non-continuum phenomena into the DSMC simulation.

The GM scattering kernel is also considered a promising boundary model for implementation in other common numerical techniques, such as LBM and MoM. Nevertheless, further studies are required to define the suitable schemes for coupling the GM scattering kernel into the aforementioned numerical methods.

A

Appendix - Comparison between velocity correlations, PDFs, and ACs obtained from training, testing, and the GM model for the Ar-Au system

A

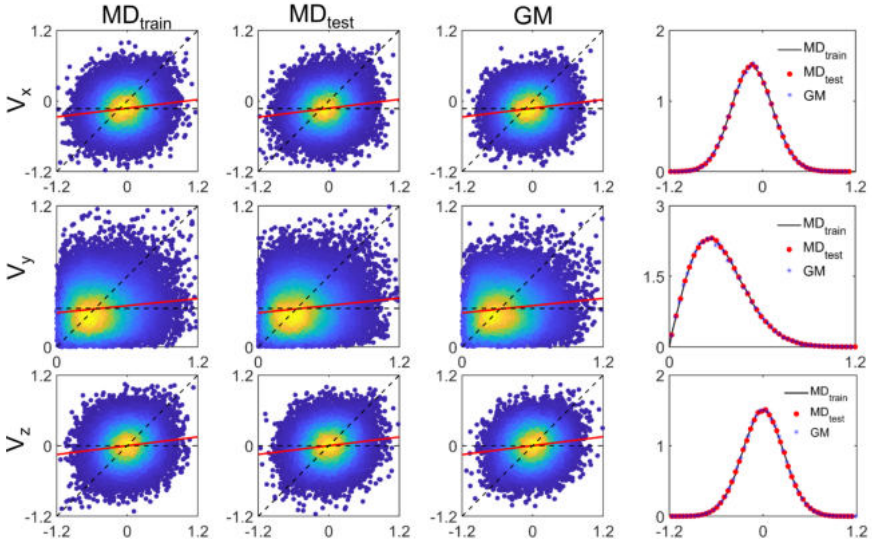

Figure A.1: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for the Ar-Au system at the bottom wall (T_b = 300 K, T_t = 500 K, S_w = 0.5). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

Table A.1: Tangential momentum (α_x,α_z), normal energy (α_{NE}), and energy (α_E) accommodation coefficients for two parallel walls Ar-Au system at the bottom wall (T_b = 300 K, T_t = 500 K, S_w = 0.5), using MD training and test data sets, as well as the GM model.

Model	α_x	α_z	α_{NE}	α_E
MD_{train}	0.876	0.875	0.897	0.869
MD_{test}	0.874	0.876	0.895	0.866
GM	0.877	0.873	0.897	0.877

B

APPENDIX-VELOCITY CORRELATIONS AND PDFS FOR THE AR-AU SYSTEM THAT ARE NOT DISCUSSED IN THE MAIN TEXT

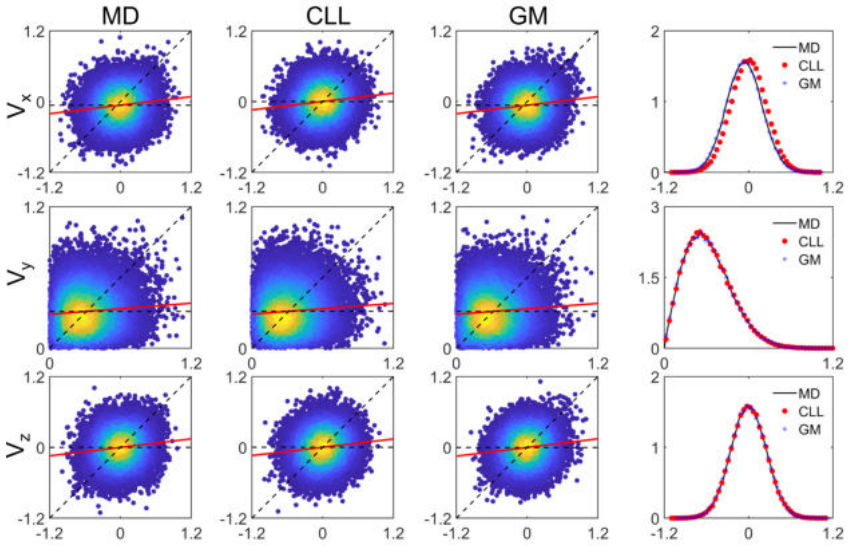

Figure B.1: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for the Ar-Au system at the bottom wall (T_b = 300 K, T_t = 300 K S_w = 0.25). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

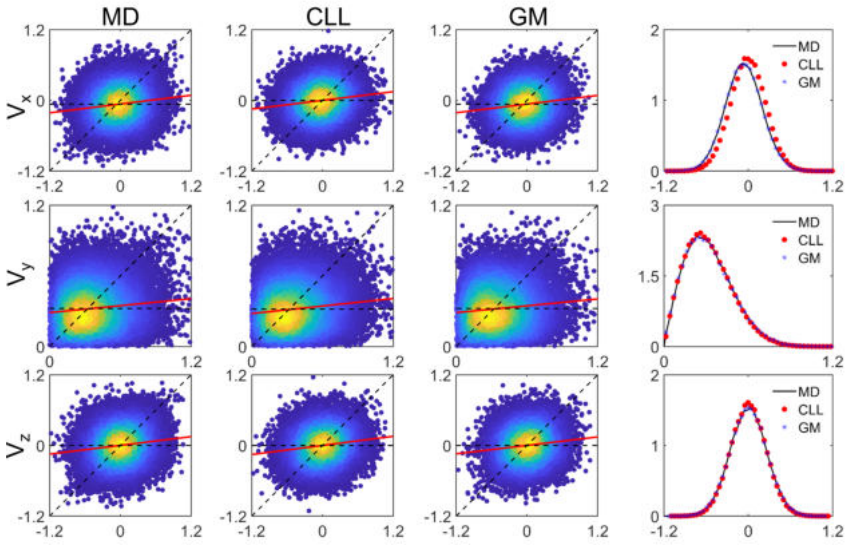

Figure B.2: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for the Ar-Au system at the bottom wall (T_b = 300 K, T_t = 500 K S_w = 0.25). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

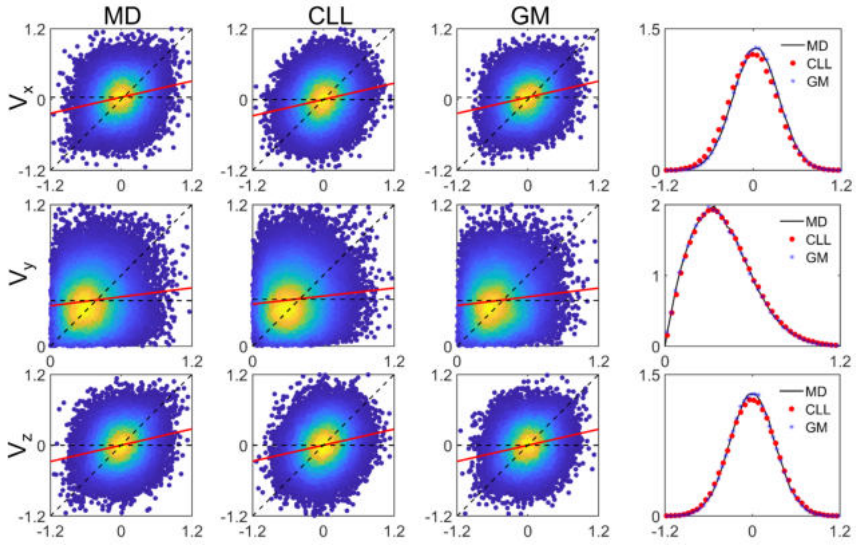

Figure B.3: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for the Ar-Au system at the top wall ($T_b = 300$ K, $T_t = 500$ K $S_w = 0.25$). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

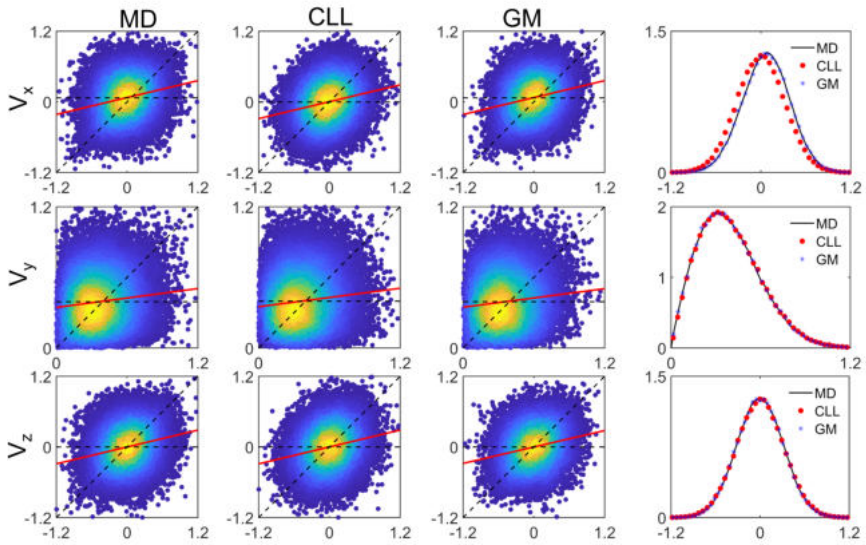

Figure B.4: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for the Ar-Au system at the top wall ($T_b = 300$ K, $T_t = 500$ K $S_w = 0.5$). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

C

Appendix-Velocity correlations and PDFs for the He-Au system that are not discussed in the main text

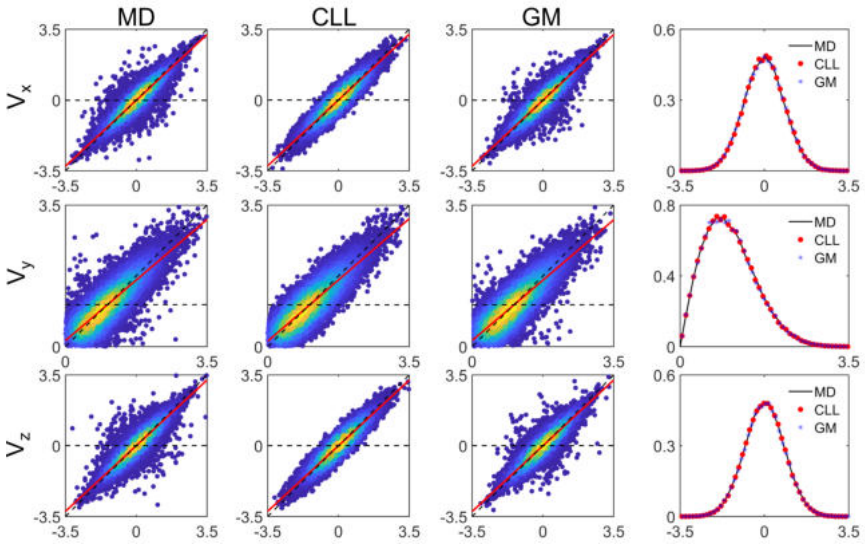

Figure C.1: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for the He-Au system at the bottom wall (T_b = 300 K, T_t = 300 K S_w = 0.25). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

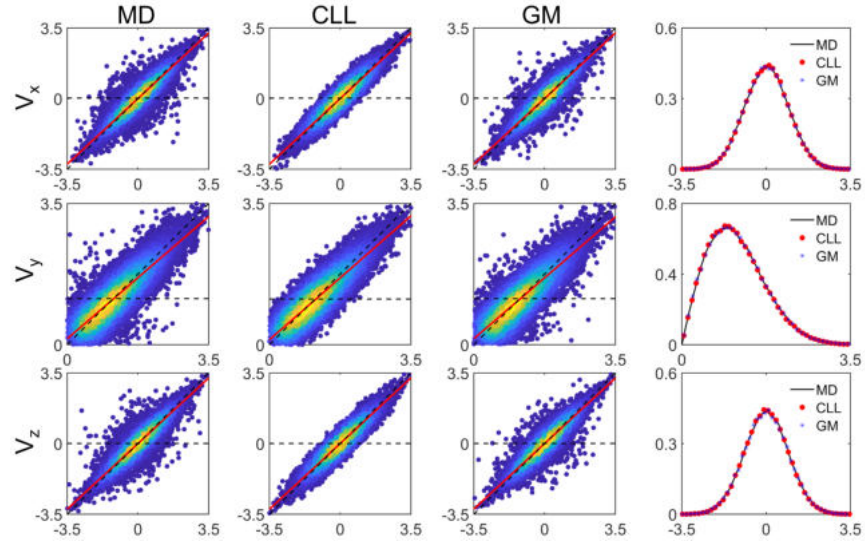

Figure C.2: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for the He-Au system at the bottom wall (T_b = 300 K, T_t = 500 K S_w = 0). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

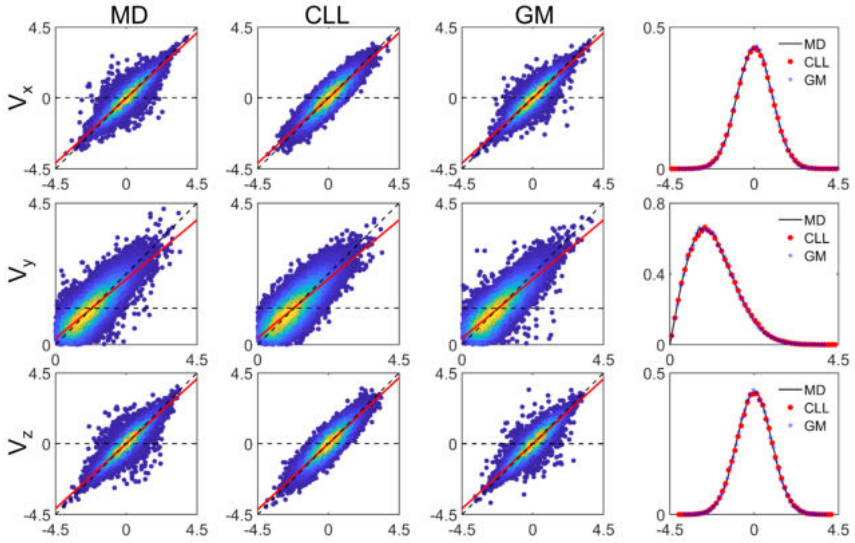

Figure C.3: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for the He-Au system at the top wall (T_b = 300 K, T_t = 500 K S_w = 0). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

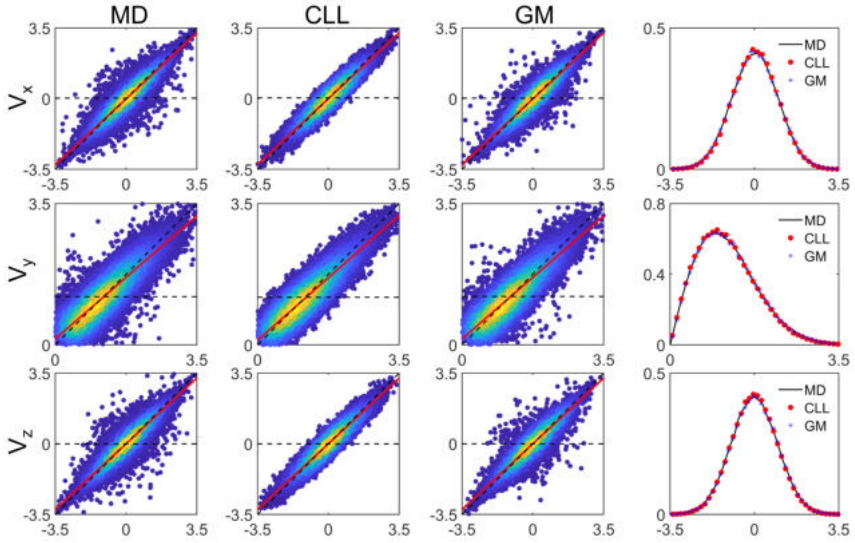

Figure C.4: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for the He-Au system at the bottom wall (T_b = 300 K, T_t = 500 K S_w = 0.25). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

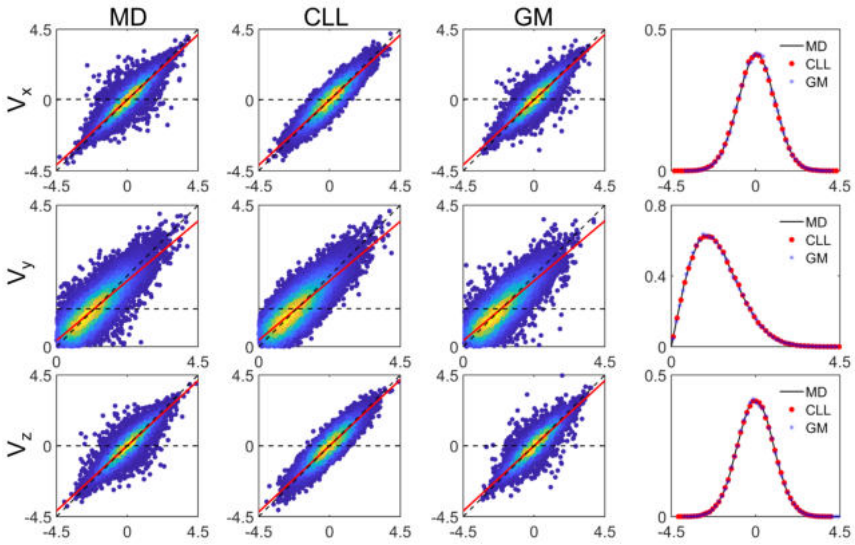

Figure C.5: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for the He-Au system at the top wall ($T_b = 300$ K, $T_t = 500$ K $S_w = 0.25$). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

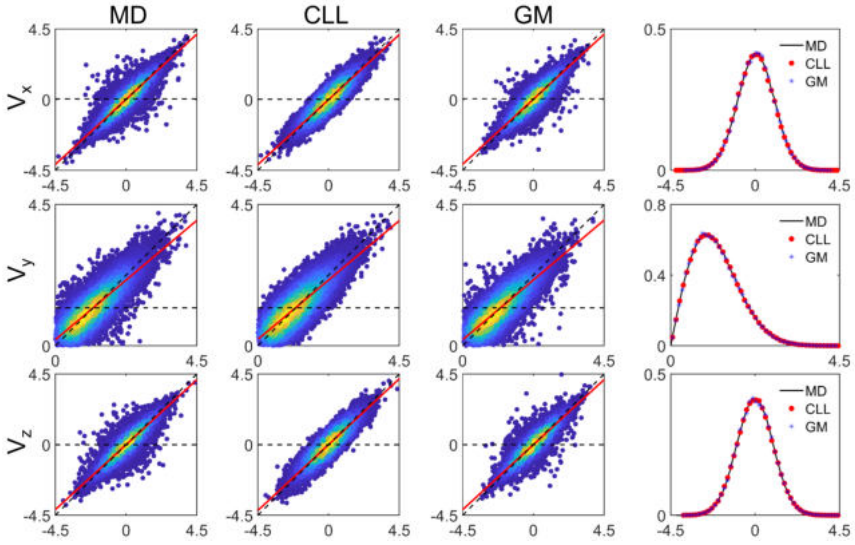

Figure C.6: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in nm/ps for the He-Au system at the bottom wall ($T_b = 300$ K, $T_t = 500$ K $S_w = 0.5$). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density function for the reflecting particles is shown.

D

Appendix- Sensitivity analysis to decide the number of Gaussians used in the GM model for diatomic gas molecules

In the case of the GM model a sensitivity analysis is needed to select the optimal number of the Gaussian functions. Here, for both gas-solid pairs the isothermal Fourier thermal problem without any heat flux($T_b = T_t$ =300 K) is chosen as the benchmark system. Afterwards, the number of Gaussians K is varied in the range of $K = 1$ to 1000, and the ACs for different gas molecule's energy modes consist of the translational energy AC (α_{tr}), the rotational energy AC (α_{rot}), and the total energy AC (α_{tot}) driven from the GM model are compared with the corresponding values obtained from the original MD data. From Figure D.1 it can be understood that for H_2-Ni system there is no considerable change in the values of the ACs for $K \geqslant 500$. Therefore, $K = 500$ is used as the number of Gaussian functions in the GM model in the case of H_2-Ni system. The results for N_2-Ni system are illustrated in Figure D.1. In this case study for $K \geqslant 100$ already the difference between all the computed ACs based on the GM model and the MD results is less than 5%. However, considering more importance for the total energy AC, the smallest K, at which the difference between the GM model and MD results is the lowest ($K = 450$) is used as the number of Gaussians for training the GM model based on the results from N_2-Ni system.

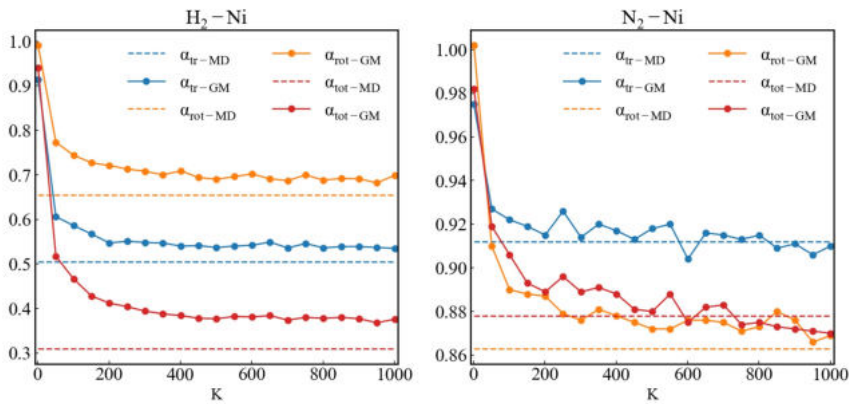

Figure D.1: Accommodation coefficients related to the translational energy (α_{tr}), rotational energy (α_{rot}), and total energy (α_{tot}) computed using atomistic simulations and the GM scattering model using different number of Gaussian functions K

E

Appendix-Correlation graphs and PDFs for H_2-Ni case studies that are not discussed in the main text

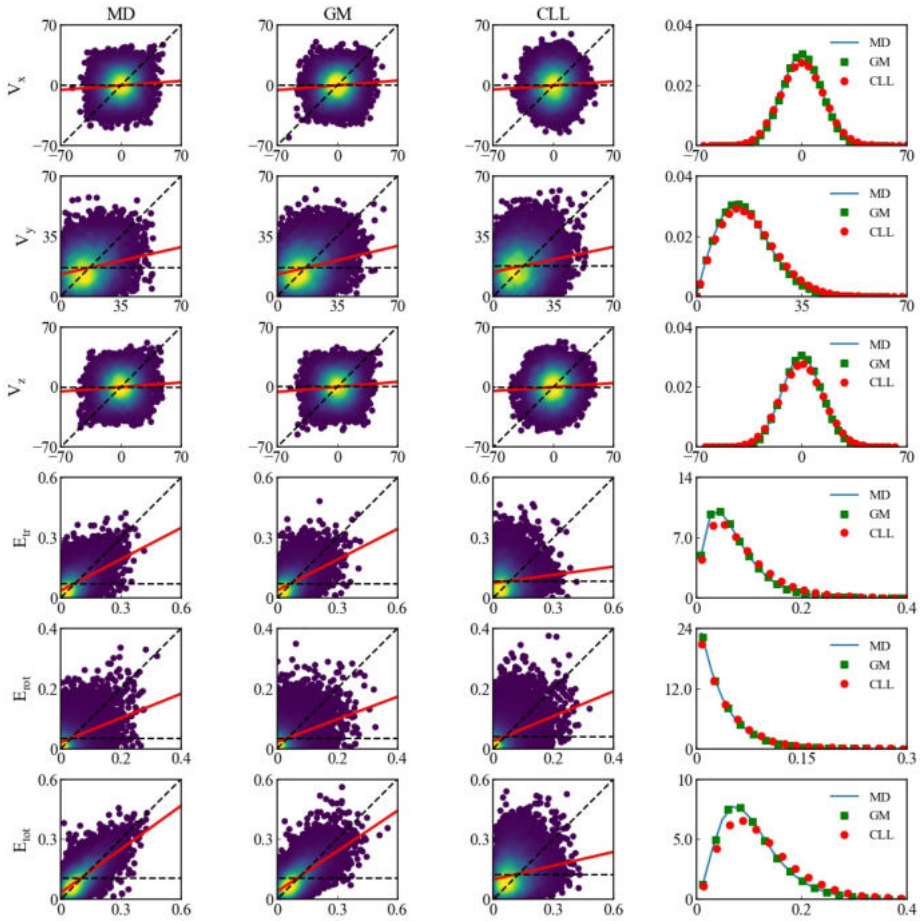

Figure E.1: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the non-isothermal Fourier thermal problem for H_2-Ni system at the top wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

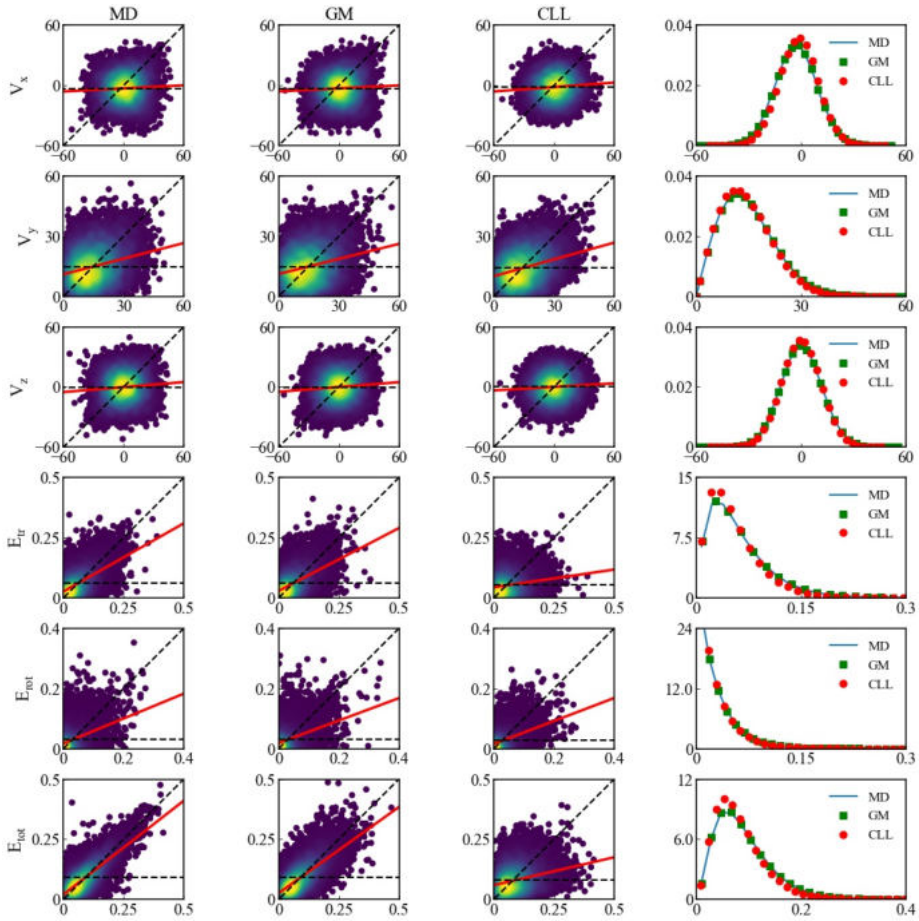

Figure E.2: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of Couette flow problem (S_U =0.2) for H_2-Ni system at the bottom wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

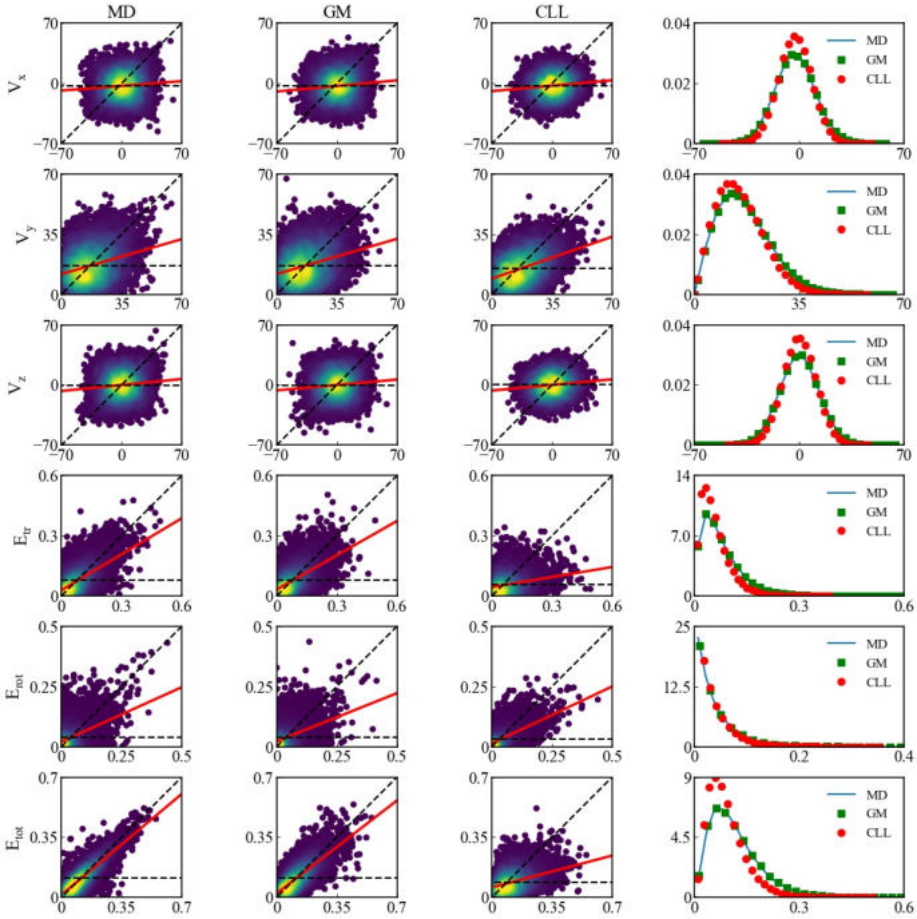

Figure E.3: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of combined Fourier-Couette flow problem for H_2-Ni system at the bottom wall ($S_v = 0.2$). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

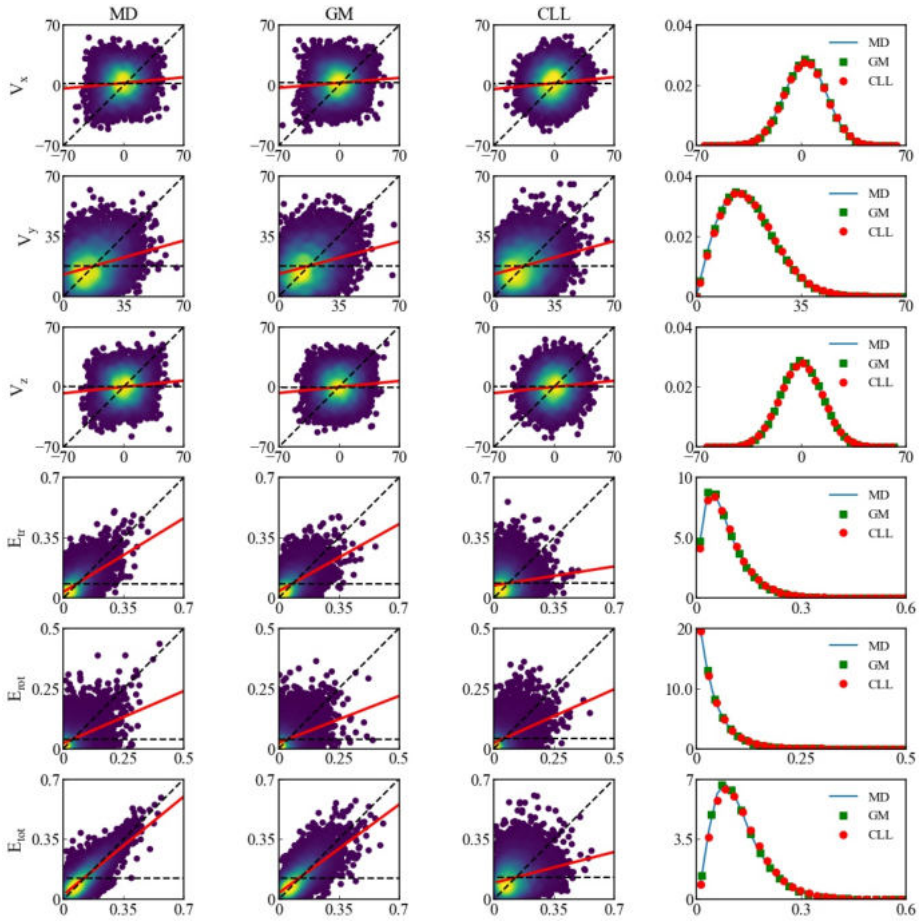

Figure E.4: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of combined Fourier-Couette flow problem for H_2-Ni system at the top wall ($S_v = 0.2$). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

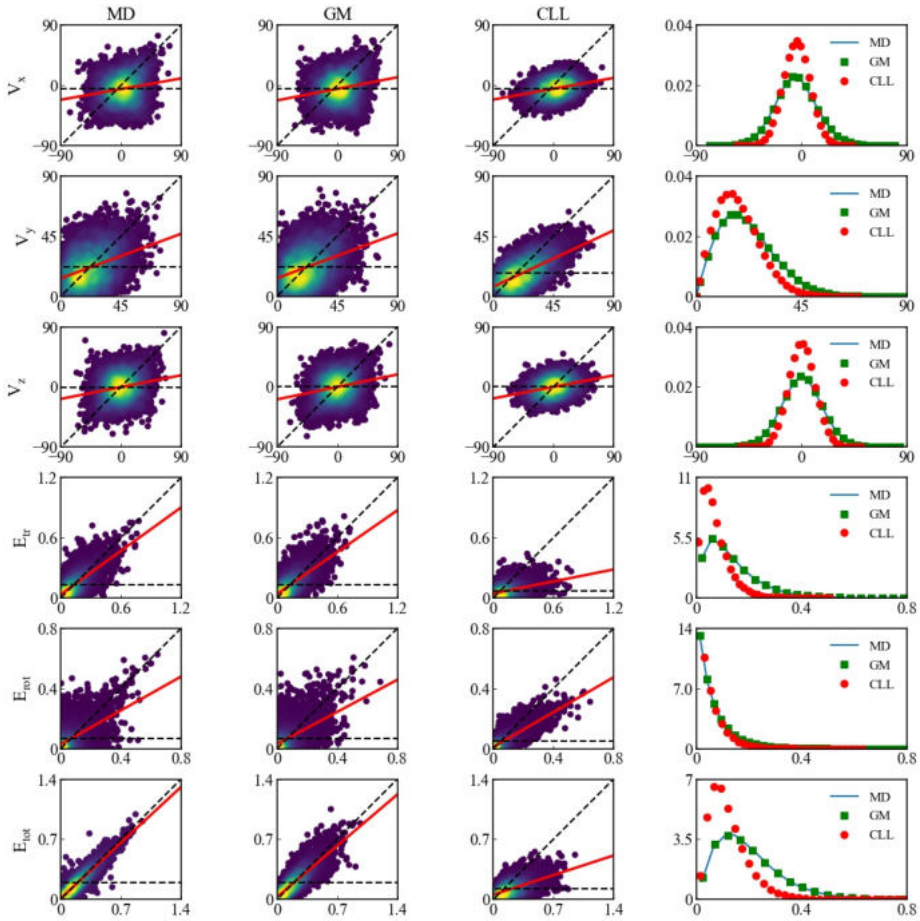

Figure E.5: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of combined Fourier-Couette flow problem for H_2-Ni system at the bottom wall ($S_v = 0.4$). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

F

Appendix-Correlation graphs and PDFs for N_2-Ni case studies that are not discussed in the main text

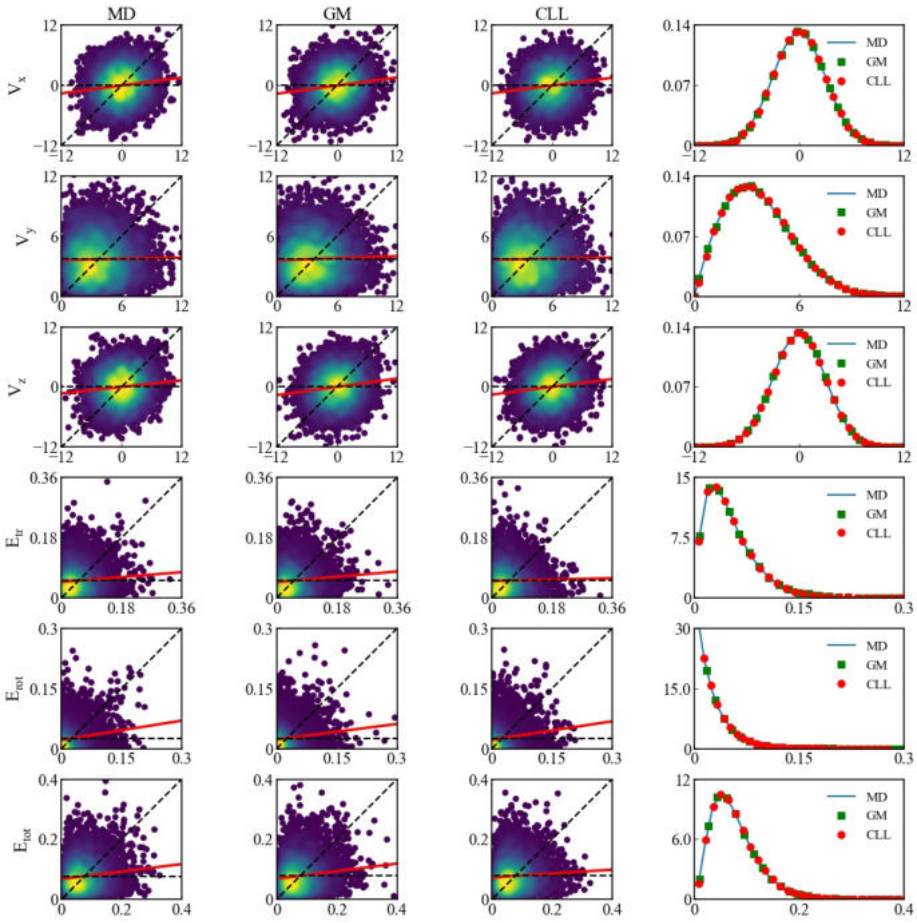

Figure F.1: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the isothermal Fourier thermal problem for N_2-Ni system at the bottom wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of translational velocity components and energy modes for the reflecting particles are presented

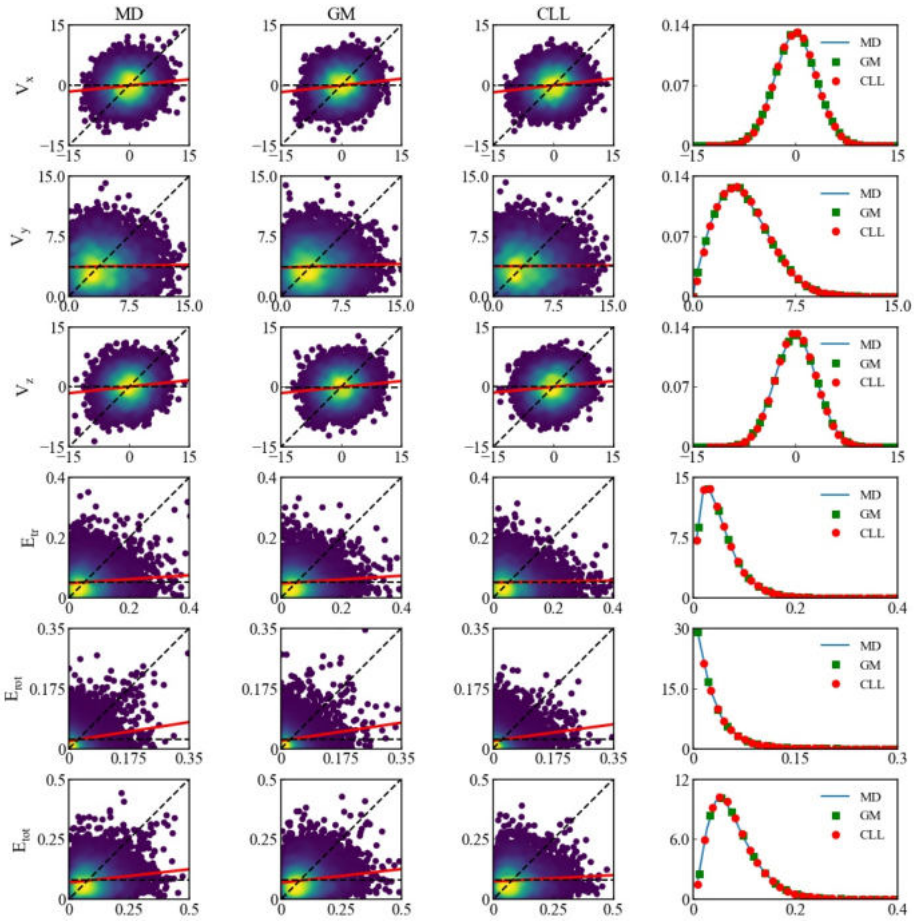

Figure F.2: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of the non-isothermal Fourier thermal problem for N_2-Ni system at the bottom wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

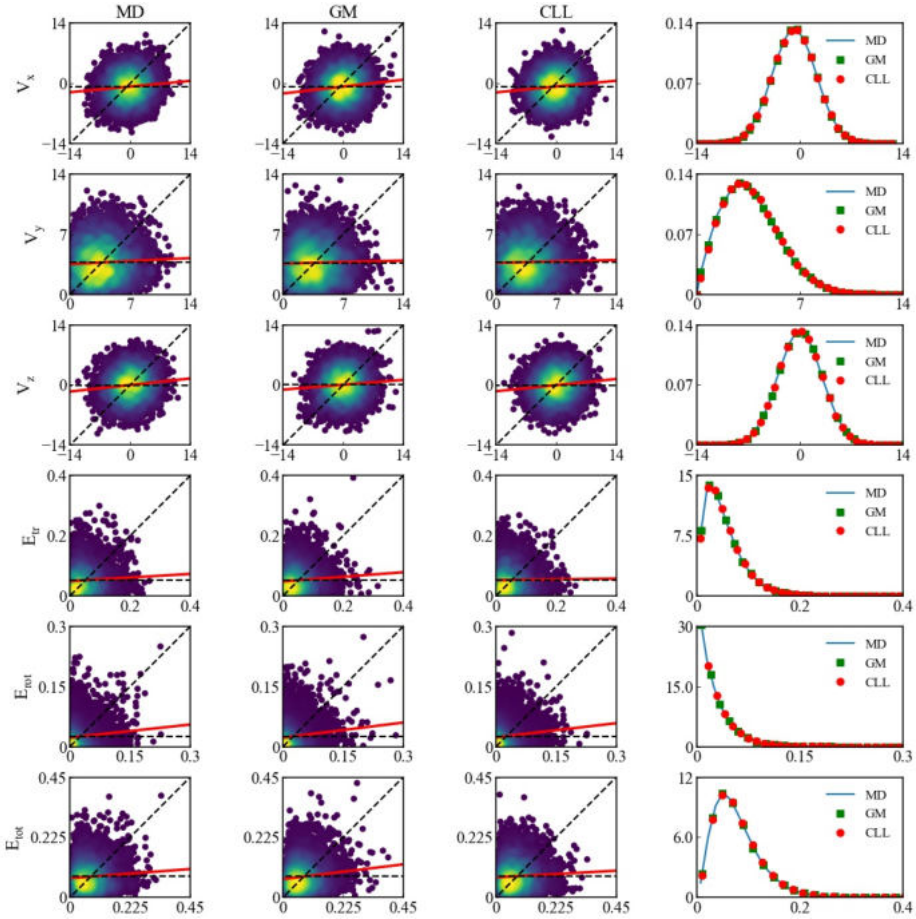

Figure F.3: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of Couette flow problem (S_v =0.2) for N_2-Ni system at the bottom wall. The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

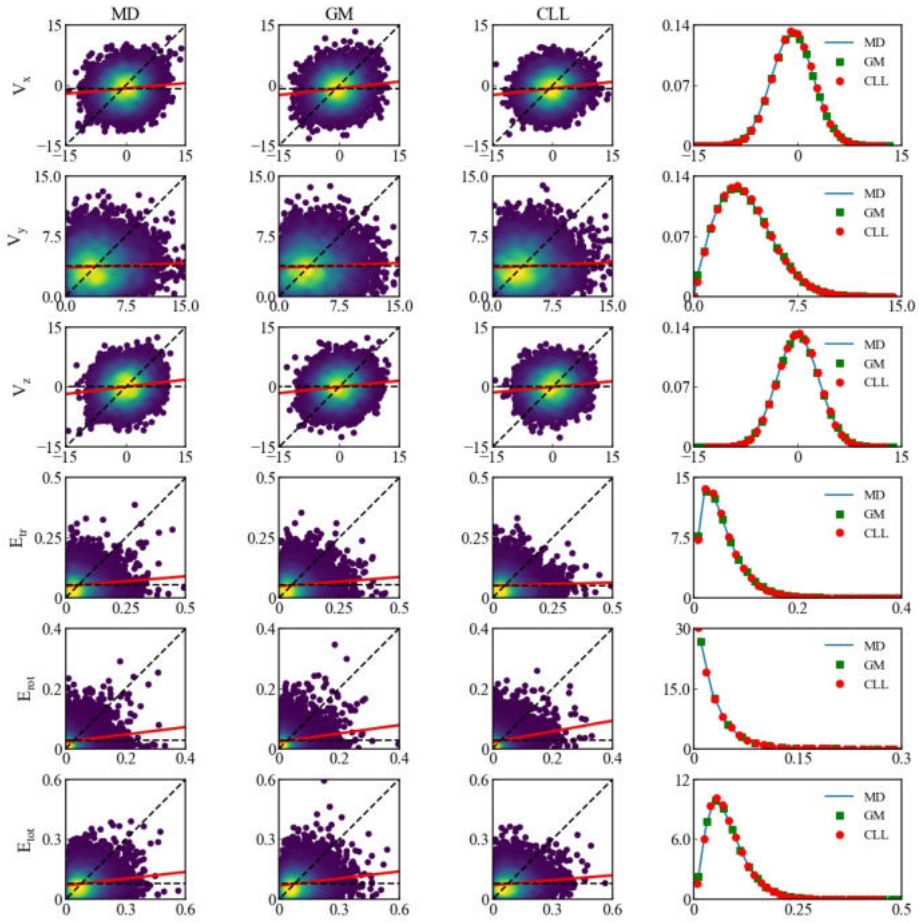

Figure F.4: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of combined Fourier-Couette flow problem for N_2-Ni system at the bottom wall ($S_U = 0.2$). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

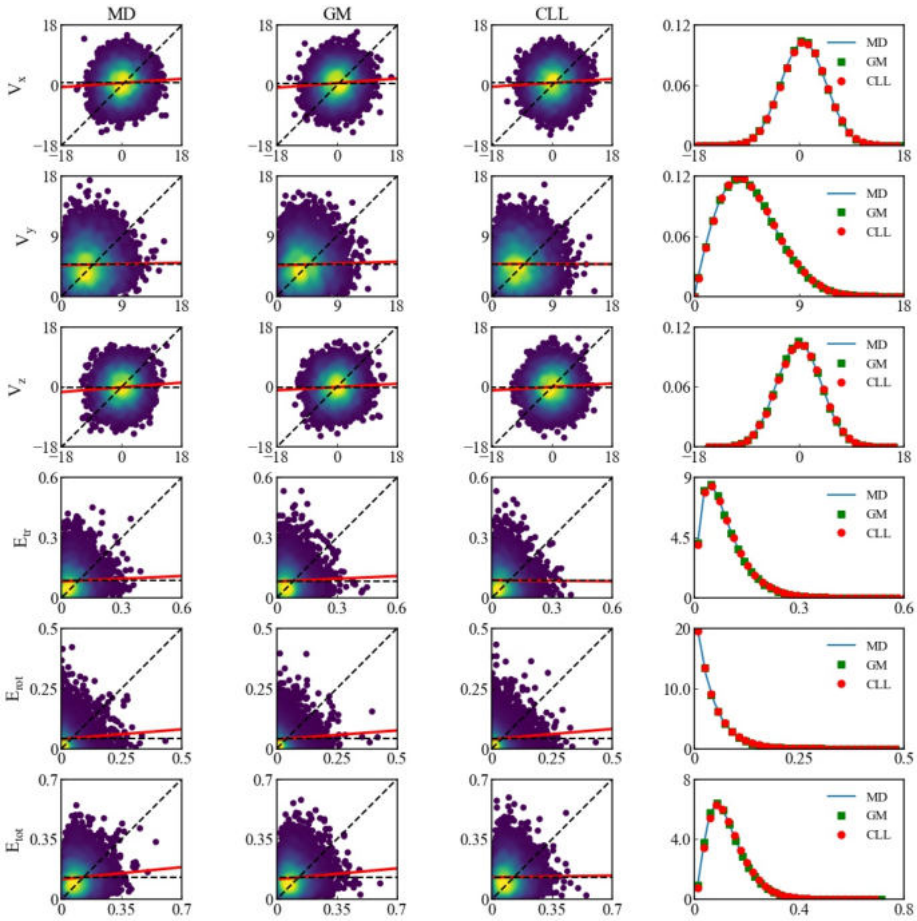

Figure F.5: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of combined Fourier-Couette flow problem for N_2-Ni system at the top wall ($S_v = 0.2$). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

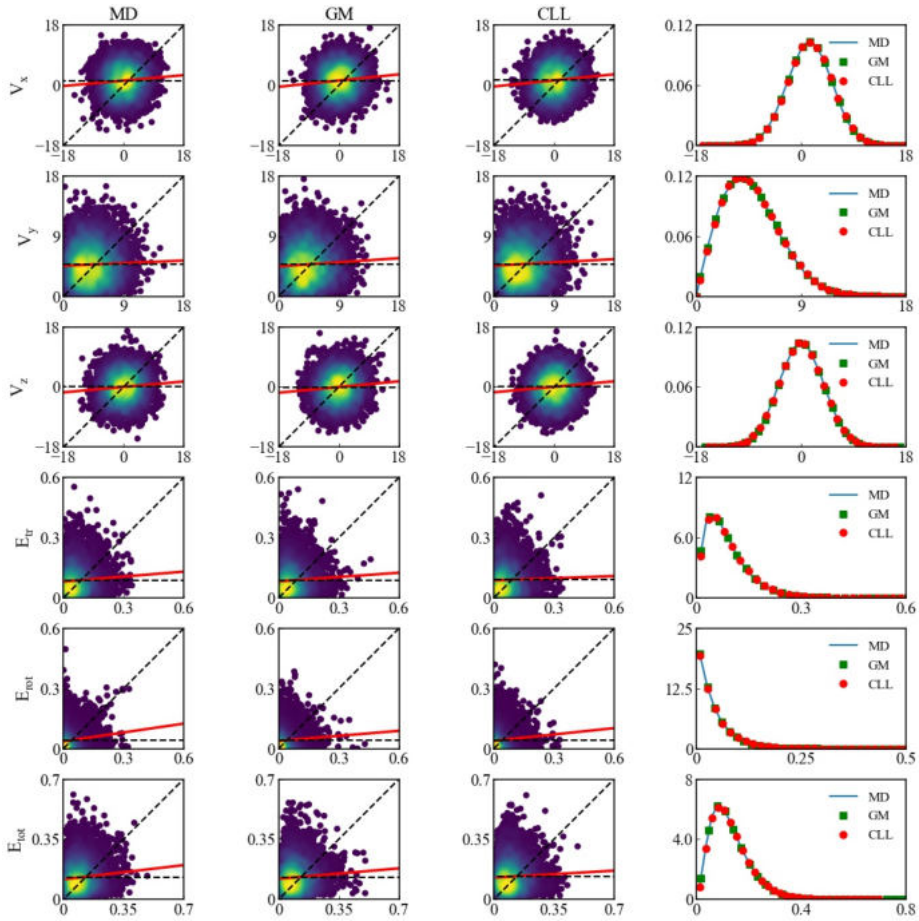

Figure F.6: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of combined Fourier-Couette flow problem for N_2-Ni system at the top wall ($S_U = 0.4$). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

G

Appendix-Correlation graphs and PDFs of the generalized GM model for the Ar-Au system

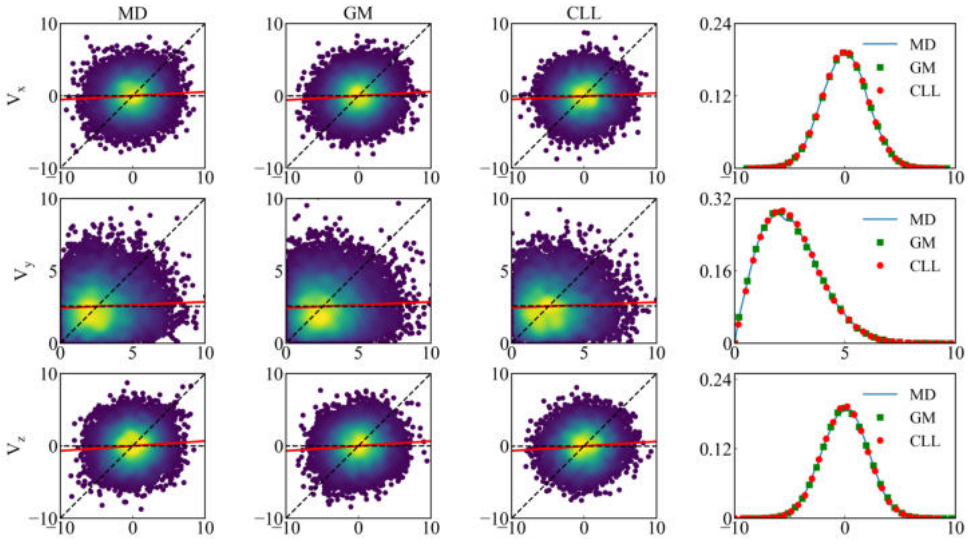

Figure G.1: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in [Å/ps] for Ar-Au system at the top wall(T_b=300 K,T_t=200 K). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density functions for the reflecting particles are shown.

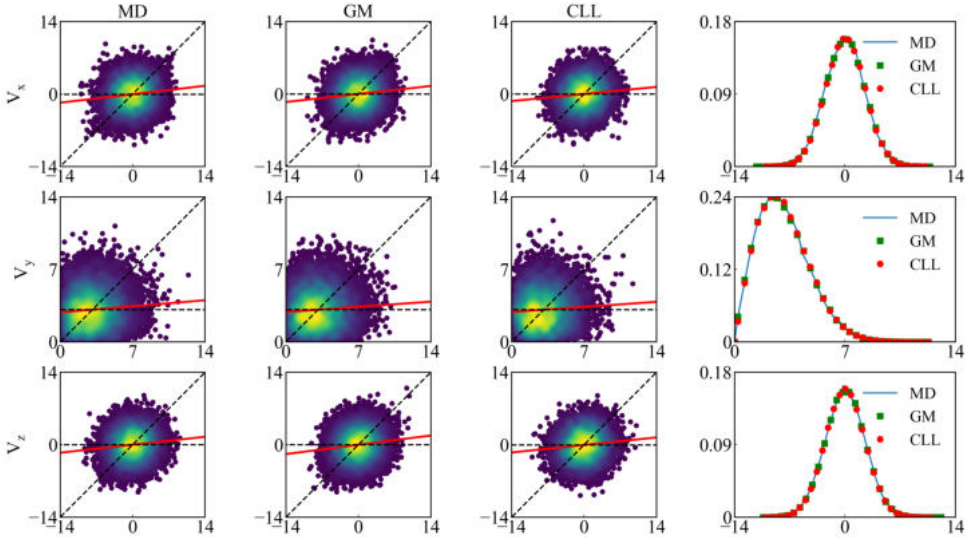

Figure G.2: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in [Å/ps] for Ar-Au system at the top wall(T_b=300 K,T_t=300 K). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density functions for the reflecting particles are shown.

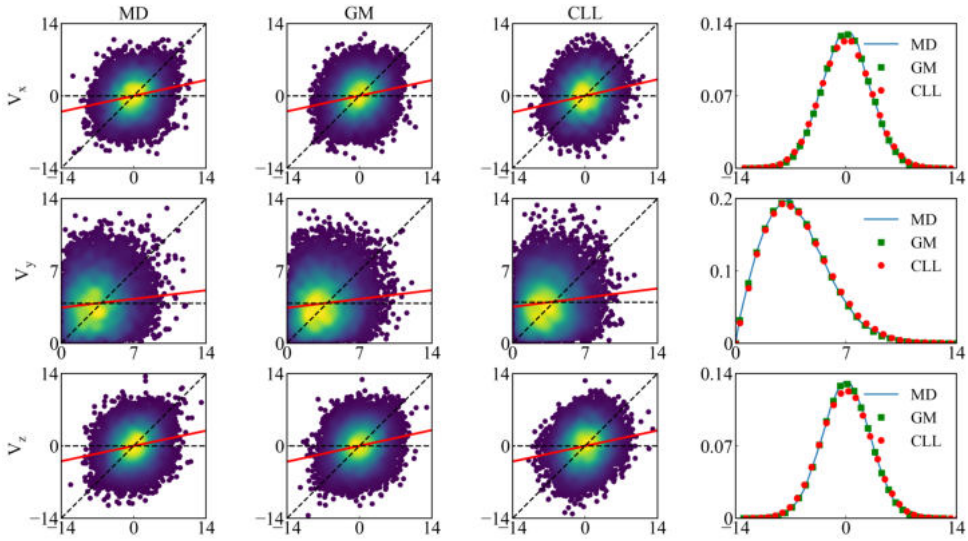

Figure G.3: Velocity correlations of impinging (horizontal-axis) and reflected (vertical-axis) velocity components in [Å/ps] for Ar-Au system at the top wall(T_b=300 K,T_t=500 K). The dashed horizontal and diagonal lines indicate fully diffusive and specular conditions, respectively. Red lines indicate the least-square linear fit of the data. In the last column the corresponding probability density functions for the reflecting particles are shown.

H

Appendix-Correlation graphs and PDFs of the generalized GM model for the H_2-Ni system

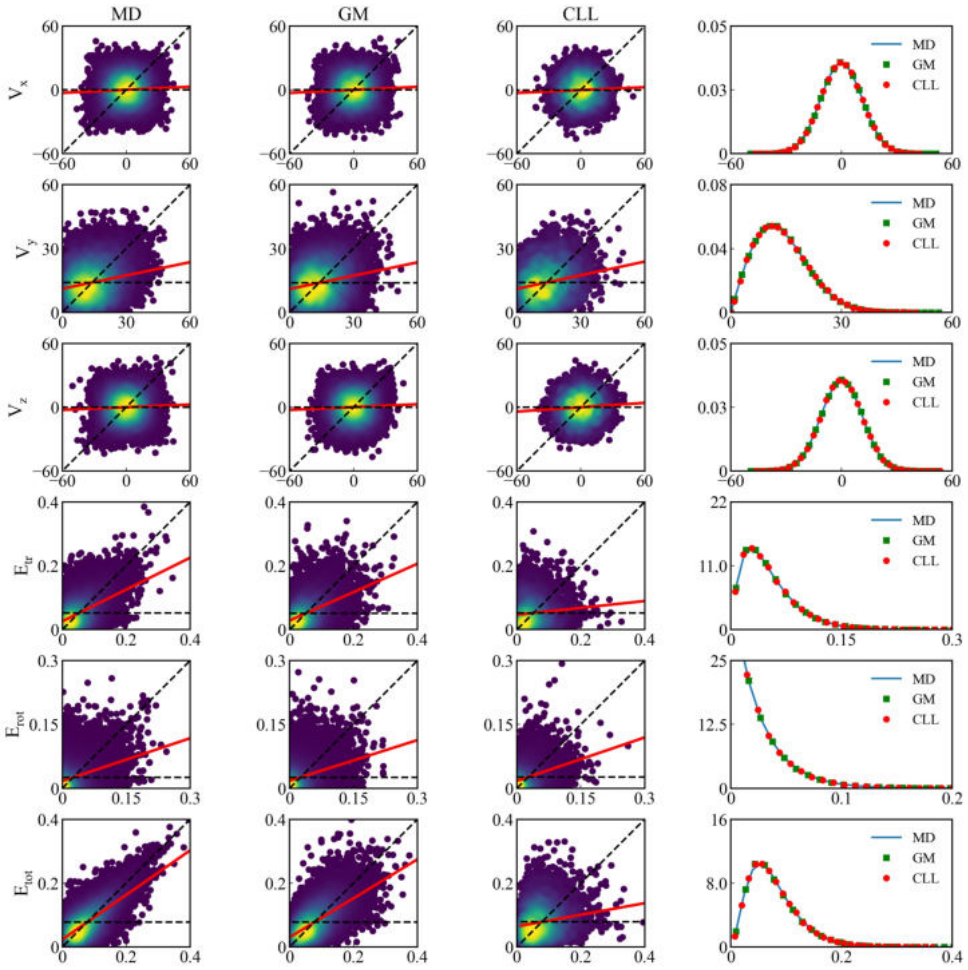

Figure H.1: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of H_2-Ni system at the top wall (T_b = 300 K, T_t = 300 K). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

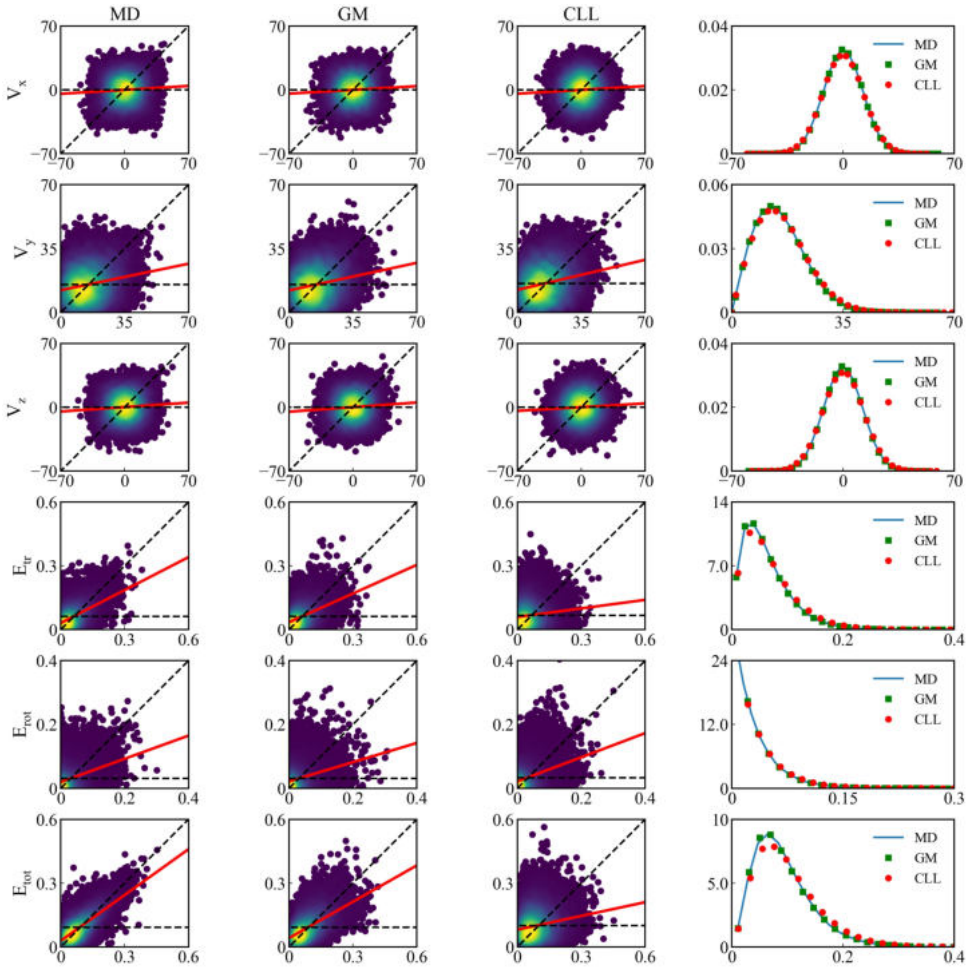

Figure H.2: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of H_2-Ni system at the top wall (T_b = 300 K, T_t = 400 K). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

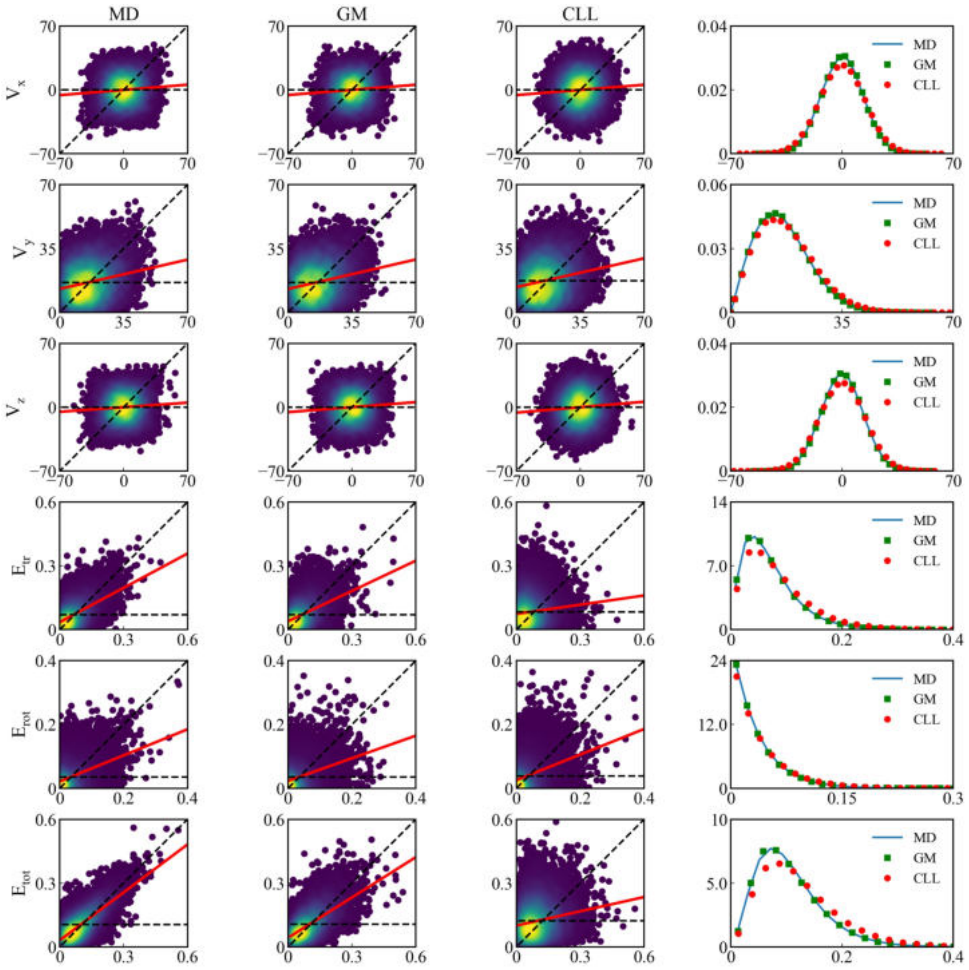

Figure H.3: Correlations between incoming (horizontal-axis) and outgoing (vertical-axis) translational velocity components in [Å/ps] and energy modes in [eV] of H_2-Ni system at the top wall (T_b = 300 K, T_t = 500 K). The dashed horizontal and diagonal lines demonstrate fully diffusive and specular reflection, respectively. Solid red lines demonstrate the least-square linear fit of the kinetic data. In the last column the corresponding probability density functions of the translational velocity components and energy modes for the reflecting particles are presented

I

Appendix-Sensitivity analysis to decide the number of Gaussians used in the GM model for mono/diatomic gas molecules

In the case of the GM model a sensitivity analysis is required to choose the optimal number of the Gaussian functions. In this work, for both gas-solid pairs the isothermal Fourier thermal problem ($T_b = T_t$ =300 K) is chosen as the benchmark system. The number of Gaussians, K, is varied in the range of K = 1 to 1000.
For the Ar-Au system, including only the translational degrees of freedom, the translational kinetic energy AC (α_{tr}) driven from the GM model is compared with the corresponding value obtained from MD simulation. On the other hand, for the H_2-Ni system the ACs for different gas molecule's energy modes consist of the translational energy AC (α_{tr}), the rotational energy AC (α_{rot}), and the total energy AC (α_{tot}) driven from the GM model are compared with the reference MD results. The results of sensitivity analysis for both gas-solid pairs are shown in Figure I.1. In the case of the Ar-Au system, for $K \geqslant 50$ already the difference between α_{tr} based on the GM model and the MD data is less than 5%. Therefore, K = 100 is utilized as the number of Gaussians for training the GM model based on the results from the Ar-Au system. For the H_2-Ni system for $K \geqslant 500$ the computed ACs become stable and no significant change in their values are observed. Therefore, K = 500 is employed as the number of Gaussian functions in the GM model in the case of H_2-Ni system.

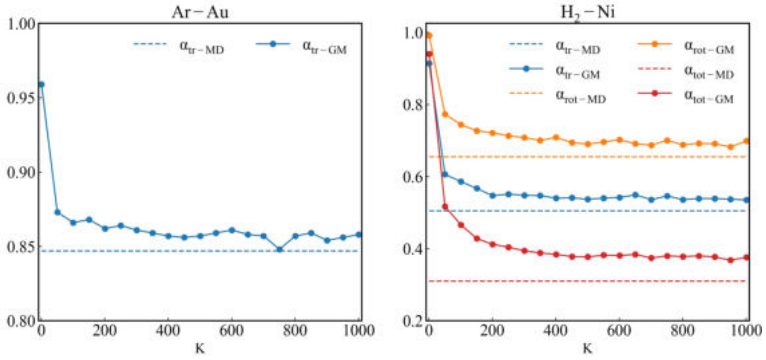

Figure I.1: Accommodation coefficients related to the translational energy (α_{tr}), rotational energy (α_{rot}), and total energy (α_{tot}) computed based on MD simulations and the GM scattering model using different number of Gaussian functions K

J

Appendix-Integration of the GM scattering kernel into the DSMC solver

This section elaborates further upon integrating of the GM scattering kernel with the applied DSMC solver in the case of H_2-Ni system. The integration procedure for the Ar-Au system is not discussed here as it is similar except that it excludes the angular velocities.

First of all, considering $\boldsymbol{x_I} = (v'_x, v'_y, v'_z, \omega'_1, \omega'_2)$ as the impinging, and $\boldsymbol{x_O} = (v_x, v_y, v_z, \omega_1, \omega_2)$ as the outgoing velocity components obtained from the MD simulation the corresponding translational temperatures (T_I, T_O) are computed. As addressed in Section 6.2.2.2 (Equation 6.3), T_I is applied to transform the distribution of the impinging normal velocity component from Rayleigh to Gaussian distribution. On the other hand, T_O is utilized to transform the distribution of the outgoing normal velocity component produced by the GM scattering kernel from Gaussian to Rayleigh distribution (see Section 6.2.2.2 (Equation 6.11)).

After training the GM model using the the MD collisional data $X = (\boldsymbol{x_I}, \boldsymbol{x_O})$ at a specific wall, the model parameters $(\{\rho_i, \vec{\mu}_i, \Sigma_i\} \ \forall \ i \text{ in } \{1 \cdots K\})$ are employed to calculate the following constants for each multivariate Gaussian function i:

$$S_{A,i} = \frac{1}{(2\pi)^{M/4} |\Sigma_{iII}|^{1/2}} \tag{J.1}$$

$$S_{B,i} = \Sigma_{iII}^{-1} \tag{J.2}$$

$$S_{C,i} = \Sigma_{iOI} \Sigma_{iII}^{-1} \tag{J.3}$$

$$S_{D,i} = Chol\left[\Sigma_{iO|I} + \epsilon I\right] \tag{J.4}$$

$S_{A,i}$ and $S_{B,i}$ are used in Equation 6.9 to compute $p_i(X_I)$. S_{Ci} and $S_{D,i}$ are applied to sample from a multivariate Gaussian with arguments $\boldsymbol{\mu}_{iO|I}$ and $\Sigma_{iO|I}$. The procedure proposed in Ref [140] is followed for sampling purpose. Assuming $\Sigma_{iO|I}$ is a positive definitive matrix, the Cholesky decomposition (*Chol*) is used to decompose the covariance matrix into the product of a lower triangular matrix and its transpose. In practice, adding a small multiple of the identity matrix I to the covariance matrix may be required for numerical stability. Here, $\epsilon = 0.0001$ is used as a small perturbation. Generating data based on the trained GM scattering kernel is one of the most crucial steps in the GM-DSMC model. Therefore, the steps followed in sampling from a GM model are elaborated further in this appendix and in Figure J.1. The sampling procedure starts with computing the probability, p_i, of a given transformed impinging velocity, $X^*_{I,j}$ belonging to the Gaussian component i (block C1 in Figure J.1). Doing such for all K components of the GM model, the accumulated weight, α, is computed. Using p_is and α, the new set of weights, $\tilde{\rho}_i$, is computed in block C2. These new weights, alongside selecting a random number, R_1, uniformly distributed between 0 and 1, are applied to choose the specific component i from the mixture model to generate the final sample (block C4 in Figure J.1). The arguments of the chosen Gaussian component (μ_{iI}, μ_{iO}, S_{Ci}, $S_{D,i}$) are used to generate the final sample (block C6 in Figure J.1) based on a random vector, u, drawn from a normal distribution with the mean value of 0 and variance of 1 (block C5 in Figure J.1). This procedure is repeated M times to generate the corresponding outgoing velocity of all the impinging velocities.

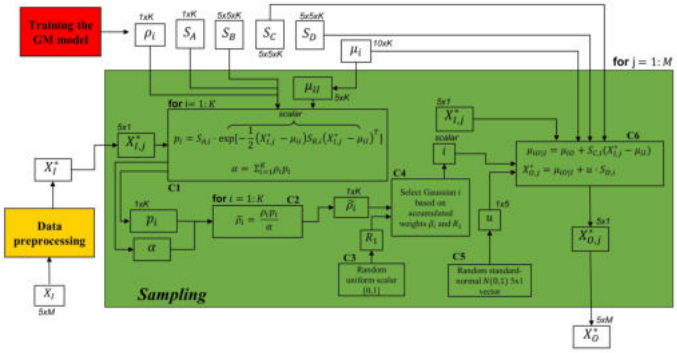

Figure J.1: The steps (C1:C6) followed to generate M samples from a GM model with K components; X_I^* and X_O^* are the transformed impinging and outgoing velocity vectors.

The scheme followed to couple the GM scattering kernel into the DSMC solver is shown in Figure J.2. A DSMC particle approaching to the system boundary has the translational velocity (v'_x, v'_y, v'_z) and the rotational energy E_{rot-I}. Since the GM mode is trained based on the rotational velocities, in the first step (block E1 in Figure J.2) E_{rot-I} is applied to derive the corresponding rotational velocities (ω'_1, ω'_2). In the next step (block E2 in Figure J.2), the data preprocessing is carried out, which includes transforming the normal velocity distribution from the Rayleigh into the Gaussian distribution and normalizing the translational and rotational velocity vectors. The preprocessed impinging velocity, X_I^*, is used to generate the outgoing velocity vector, X_O^*, in block E3. Afterward, X_O^* is converted back into the initial units, and the normal velocity distribution is also transformed back into the Rayleigh distribution (block E4 in Figure J.2). Finally, the rotational velocities are used to compute the corresponding outgoing rotational energy value, E_{rot-O}, that alongside the outgoing translational velocity components (v'_x, v'_y, v'_z) are assigned to the reflected DSMC particle.

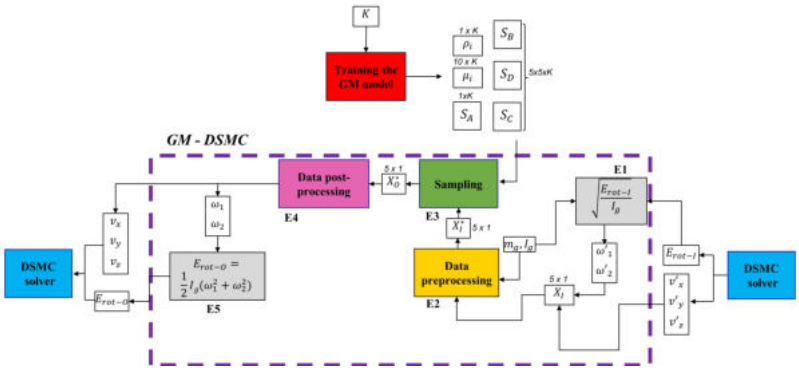

Figure J.2: The scheme showing the integration of the GM scattering kernel into DSMC solver.

BIBLIOGRAPHY

[1] George Karniadakis, Ali Beskok, and Narayan Aluru. *Microflows and nanoflows: fundamentals and simulation*, volume 29. Springer Science & Business Media, 2006.

[2] Michael Maurice Rudolph Williams and Sudarshan K Loyalka. Aerosol science: Theory and practice. 1991.

[3] Wei Chen, Shuyu Chen, Tengfei Liang, Qiang Zhang, Zhongli Fan, Hang Yin, Kuo-Wei Huang, Xixiang Zhang, Zhiping Lai, and Ping Sheng. High-flux water desalination with interfacial salt sieving effect in nanoporous carbon composite membranes. *Nature nanotechnology*, 13(4):345–350, 2018.

[4] Ching Shen. *Rarefied gas dynamics: fundamentals, simulations and micro flows*. Springer Science & Business Media, 2006.

[5] Kerson Huang. *Statistical mechanics*. John Wiley & Sons, 2008.

[6] Zi-Xiang Tong, Ya-Ling He, and Wen-Quan Tao. A review of current progress in multiscale simulations for fluid flow and heat transfer problems: The frameworks, coupling techniques and future perspectives. *International Journal of Heat and Mass Transfer*, 137:1263–1289, 2019.

[7] Joel H Ferziger, Milovan Perić, and Robert L Street. *Computational methods for fluid dynamics*, volume 3. Springer, 2002.

[8] Suhas V Patankar. *Numerical heat transfer and fluid flow*. CRC press, 2018.

[9] C David Levermore. Moment closure hierarchies for kinetic theories. *Journal of statistical Physics*, 83(5):1021–1065, 1996.

[10] Shiyi Chen and Gary D Doolen. Lattice boltzmann method for fluid flows. *Annual review of fluid mechanics*, 30(1):329–364, 1998.

[11] Graeme A Bird and JM Brady. *Molecular gas dynamics and the direct simulation of gas flows*, volume 5. Clarendon press Oxford, 1994.

[12] Michael P Allen and Dominic J Tildesley. *Computer simulation of liquids*. Oxford university press, 2017.

[13] G Di Staso. Hybrid discretizations of the boltzmann equation for the dilute gas flow regime. 2018.

[14] Earle H Keddard. *Kinetic theory of gases: with an introduction to statistical mechanics*. McGraw-Hill, 1938.

[15] James Clerk Maxwell. Iii. on stresses in rarefied gases arising from inequalities of temperature. *Proceedings of the Royal Society of London*, 27(185-189):304–308, 1878.

[16] Carlo Cercignani and Maria Lampis. Kinetic models for gas-surface interactions. *Transport theory and statistical physics*, 1(2):101–114, 1971.

[17] RG Lord. Some extensions to the cercignani–lampis gas–surface scattering kernel. *Physics of Fluids A: Fluid Dynamics*, 3(4):706–710, 1991.

[18] Melvin Epstein. A model of the wall boundary condition in kinetic theory. *AIAA Journal*, 5(10):1797–1800, 1967.

[19] S Kokou Dadzie and J Gilbert Méolans. Temperature jump and slip velocity calculations from an anisotropic scattering kernel. *Physica A: Statistical Mechanics and its Applications*, 358(2-4):328–346, 2005.

[20] Quy-Dong To, Van-Huyen Vu, Guy Lauriat, and Céline Léonard. Boundary conditions for gas flow problems from anisotropic scattering kernels. *Journal of Mathematical Physics*, 56(10):103101, 2015.

[21] Kyoji Yamamoto, Hideki Takeuchi, and Toru Hyakutake. Characteristics of reflected gas molecules at a solid surface. *Physics of Fluids*, 18(4):046103, 2006.

[22] Henning Struchtrup. Maxwell boundary condition and velocity dependent accommodation coefficient. *Physics of Fluids*, 25(11):112001, 2013.

[23] AN Yakunchikov, VL Kovalev, and SV Utyuzhnikov. Analysis of gas-surface scattering models based on computational molecular dynamics. *Chemical Physics Letters*, 554:225–230, 2012.

[24] M Hossein Gorji and Patrick Jenny. A gas-surface interaction kernel for diatomic rarefied gas flows based on the cercignani-lampis-lord model. *Physics of fluids*, 26(12):122004, 2014.

[25] Aldo Frezzotti and Livio Gibelli. A kinetic model for fluid—wall interaction. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 222(5):787–795, 2008.

[26] Roderick Matheson Logan and RE Stickney. Simple classical model for the scattering of gas atoms from a solid surface. *The Journal of Chemical Physics*, 44(1):195–201, 1966.

[27] Roderick Matheson Logan and James C Keck. Classical theory for the interaction of gas atoms with solid surfaces. *The Journal of Chemical Physics*, 49(2):860–876, 1968.

[28] John C Tully. Washboard model of gas–surface scattering. *The Journal of chemical physics*, 92(1):680–686, 1990.

[29] Tengfei Liang, Qi Li, and Wenjing Ye. A physical-based gas–surface interaction model for rarefied gas flow simulation. *Journal of Computational Physics*, 352:105–122, 2018.

[30] Neil A Mehta and Deborah A Levin. Molecular-dynamics-derived gas–surface models for use in direct-simulation monte carlo. *Journal of Thermophysics and Heat Transfer*, 31(4):757–771, 2017.

[31] Meng Liao, Quy-Dong To, Céline Léonard, and Vincent Monchiet. Non-parametric wall model and methods of identifying boundary conditions for moments in gas flow equations. *Physics of Fluids*, 30(3):032008, 2018.

[32] Nemanja Andric, Daniel W Meyer, and Patrick Jenny. Data-based modeling of gas-surface interaction in rarefied gas flow simulations. *Physics of Fluids*, 31(6):067109, 2019.

[33] Meng Liao, Quy-Dong To, Céline Léonard, and Wenlu Yang. Prediction of thermal conductance and friction coefficients at a solid-gas interface from statistical learning of collisions. *Physical Review E*, 98(4):042104, 2018.

[34] Shahin Mohammad Nejad, Eldhose Iype, Silvia Nedea, Arjan Frijns, and David Smeulders. Modeling rarefied gas-solid surface interactions for couette flow with different wall temperatures using an unsupervised machine learning technique. *Physical Review E*, 104(1):015309, 2021.

[35] Zijing Wang, Chengqian Song, Fenghua Qin, and Xisheng Luo. Establishing a data-based scattering kernel model for gas–solid interaction by molecular dynamics simulation. *Journal of Fluid Mechanics*, 928, 2021.

[36] O Sazhin, A Kulev, S Borisov, and S Gimelshein. Numerical analysis of gas–surface scattering effect on thermal transpiration in the free molecular regime. *Vacuum*, 82(1):20–29, 2007.

[37] Tengfei Liang, Qi Li, and Wenjing Ye. Performance evaluation of maxwell and cercignani-lampis gas-wall interaction models in the modeling of thermally driven rarefied gas transport. *Physical Review E*, 88(1):013009, 2013.

[38] Subhash C Saxena and Ram Kumar Joshi. Thermal accommodation and adsorption coefficients of gases. 1989.

[39] Amit Agrawal and SV Prabhu. Survey on measurement of tangential momentum accommodation coefficient. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, 26(4):634–645, 2008.

[40] Charles T Rettner. Thermal and tangential-momentum accommodation coefficients for n/sub 2/colliding with surfaces of relevance to disk-drive air bearings derived from molecular beam scattering. *IEEE transactions on magnetics*, 34(4):2387–2395, 1998.

[41] D Bruno, M Cacciatore, S Longo, and M Rutigliano. Gas-surface scattering models for particle fluid dynamics: a comparison between analytical approximate models and molecular dynamics calculations. *chemical physics letters*, 320(3-4):245–254, 2000.

[42] Wenbin Liu, Jinbai Zhang, Yazhong Jiang, Laiwen Chen, and Chun-Hian Lee. Dsmc study of hypersonic rarefied flow using the cercignani–lampis–lord model and a molecular-dynamics-based scattering database. *Physics of Fluids*, 33(7):072003, 2021.

[43] Hsi Guang Sung. *Gaussian mixture regression and classification*. Rice University, 2004.

[44] Hui Wu, Weifang Chen, and Zhongzheng Jiang. Gaussian mixture models for diatomic gas- surface interactions under thermal non-equilibrium conditions. *Physics of Fluids*, 34(8):082007, 2022.

[45] Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*, volume 4. Springer, 2006.

[46] Van P Carey et al. *Statistical thermodynamics and microscale thermophysics*. Cambridge University Press, 1999.

[47] Herman JC Berendsen, JPM van Postma, Wilfred F Van Gunsteren, ARHJ DiNola, and Jan R Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.

[48] William G Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical review A*, 31(3):1695, 1985.

[49] Peter Spijker, Albert J Markvoort, Silvia V Nedea, and Peter AJ Hilbers. Computation of accommodation coefficients and the use of velocity correlation profiles in molecular dynamics simulations. *Physical Review E*, 81(1):011203, 2010.

[50] Roland Glowinski and Alain Lichnewsky. *Computing methods in applied sciences and engineering*. Number 45. SIAM, 1990.

[51] Wolfgang Wagner. A convergence proof for bird's direct simulation monte carlo method for the boltzmann equation. *Journal of Statistical Physics*, 66(3):1011–1044, 1992.

[52] PS Prasanth and Jose K Kakkassery. Molecular models for simulation of rarefied gas flows using direct simulation monte carlo method. *Fluid dynamics research*, 40(4):233, 2008.

[53] Claus Borgnakke and Poul S Larsen. Statistical collision model for monte carlo simulation of polyatomic gas mixture. *Journal of computational Physics*, 18(4):405–420, 1975.

[54] Amit Agrawal. A comprehensive review on gas flow in microchannels. *International Journal of Micro-Nano Scale Transport*, 2(1):1–40, 2011.

[55] Bing-Yang Cao, Jun Sun, Min Chen, and Zeng-Yuan Guo. Molecular momentum transport at fluid-solid interfaces in mems/nems: a review. *International journal of molecular sciences*, 10(11):4638–4706, 2009.

[56] S Colin. Rarefaction and compressibility effects on steady or transient gas flows in microchannels. In *International Conference on Nanochannels, Microchannels, and Minichannels*, volume 41642, pages 13–24, 2004.

[57] Gert Due Billing. The dynamics of molecule-surface interaction. *Computer Physics Reports*, 12(6):383–450, 1990.

[58] John Adair Barker and Daniel J Auerbach. Gas—surface interactions and dynamics; thermal energy atomic and molecular beam studies. *Surface Science Reports*, 4(1-2):1–99, 1984.

[59] James Clerk Maxwell. Iii. on stresses in rarefied gases arising from inequalities of temperature. *Proceedings of the Royal Society of London*, 27(185-189):304–308, 1878.

[60] RG Lord. Some further extensions of the cercignani–lampis gas–surface interaction model. *Physics of Fluids*, 7(5):1159–1161, 1995.

[61] Junfeng Zhang. Lattice boltzmann method for microfluidics: models and applications. *Microfluidics and Nanofluidics*, 10(1):1–28, 2011.

[62] Harold Grad. On the kinetic theory of rarefied gases. *Communications on pure and applied mathematics*, 2(4):331–407, 1949.

[63] Sooraj K Prabha and Sarith P Sathian. Computational study of thermal dependence of accommodation coefficients in a nano-channel and the prediction of velocity profiles. *Computers & fluids*, 68:47–53, 2012.

[64] V Chirita, BA Pailthorpe, and RE Collins. Non-equilibrium energy and momentum accommodation coefficients of ar atoms scattered from ni (001) in the thermal regime: A molecular dynamics study. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 129(4):465–473, 1997.

[65] Jun Sun and Zhi-Xin Li. Three-dimensional molecular dynamic study on accommodation coefficients in rough nanochannels. *Heat transfer engineering*, 32(7-8):658–666, 2011.

[66] Peter Spijker, Albert J Markvoort, Silvia V Nedea, and Peter AJ Hilbers. Computation of accommodation coefficients and the use of velocity correlation profiles in molecular dynamics simulations. *Physical Review E*, 81(1):011203, 2010.

[67] KJ Daun. Thermal accommodation coefficients between polyatomic gas molecules and soot in laser-induced incandescence experiments. *International Journal of Heat and Mass Transfer*, 52(21-22):5081–5089, 2009.

[68] J Reinhold, Thomas Veltzke, B Wells, J Schneider, Florian Meierhofer, L Colombi Ciacchi, A Chaffee, and J Thöming. Molecular dynamics simulations on scattering of single ar, n2, and co2 molecules on realistic surfaces. *Computers & Fluids*, 97:31–39, 2014.

[69] Tejas Mane, Pooja Bhat, Vigor Yang, and Dilip Srinivas Sundaram. Energy accommodation under non-equilibrium conditions for aluminum-inert gas systems. *Surface Science*, 677:135–148, 2018.

[70] Meng Liao, Romain Grenier, Quy-Dong To, María Pilar de Lara-Castells, and Celine Leonard. Helium and argon interactions with gold surfaces: Ab initio-assisted determination of the he–au pairwise potential and its application to accommodation coefficient determination. *The Journal of Physical Chemistry C*, 122(26):14606–14614, 2018.

[71] Kyoji Yamamoto. Slightly rarefied gas flow over a smooth platinum surface. In *AIP Conference Proceedings*, volume 585, pages 339–346. American Institute of Physics, 2001.

[72] Toru Hyakutake, Kyoji Yamamoto, and Hideki Takeuchi. Flow of gas mixtures through micro channel. In *AIP Conference Proceedings*, volume 762, pages 780–788. American Institute of Physics, 2005.

[73] George W Finger, Jayanta S Kapat, and Aniket Bhattacharya. Molecular dynamics simulation of adsorbent layer effect on tangential momentum accommodation coefficient. 2007.

[74] Jun Sun and Zhi-Xin Li. Effect of gas adsorption on momentum accommodation coefficients in microgas flows using molecular dynamic simulations. *Molecular Physics*, 106(19):2325–2332, 2008.

[75] KJ Daun, TA Sipkens, JT Titantah, and M Karttunen. Thermal accommodation coefficients for laser-induced incandescence sizing of metal nanoparticles in monatomic gases. *Applied Physics B*, 112(3):409–420, 2013.

[76] Aldo Frezzotti. Monte carlo simulation of the heat flow in a dense hard sphere gas. *European Journal of Mechanics-B/Fluids*, 18(1):103–119, 1999.

[77] HW Sheng, MJ Kramer, A Cadien, T Fujita, and MW Chen. Highly optimized embedded-atom-method potentials for fourteen fcc metals. *Physical Review B*, 83(13):134118, 2011.

[78] Romain Grenier, Quy-Dong To, Maria Pilar de Lara-Castells, and Céline Léonard. Argon interaction with gold surfaces: Ab initio-assisted determination of pair ar–au potentials for molecular dynamics simulations. *The Journal of Physical Chemistry A*, 119(26):6897–6908, 2015.

[79] Hendrik Heinz, RA Vaia, BL Farmer, and RR Naik. Accurate simulation of surfaces and interfaces of face-centered cubic metals using 12- 6 and 9- 6 lennard-jones potentials. *The Journal of Physical Chemistry C*, 112(44):17281–17290, 2008.

[80] Daniel V Schroeder. Interactive molecular dynamics. *American Journal of Physics*, 83(3):210–218, 2015.

[81] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics*, 117(1):1–19, 1995.

[82] Lloyd B Thomas and Robert E Brown. The accommodation coefficients of gases on platinum as a function of pressure. *The Journal of Chemical Physics*, 18(10):1367–1372, 1950.

[83] Albert J Markvoort, PAJ Hilbers, and SV Nedea. Molecular dynamics study of the influence of wall-gas interactions on heat flow in nanochannels. *Physical review E*, 71(6):066702, 2005.

[84] Wayne M Trott, Jaime N Castañeda, John R Torczynski, Michael A Gallis, and Daniel J Rader. An experimental assembly for precise measurement of thermal accommodation coefficients. *Review of scientific instruments*, 82(3):035120, 2011.

[85] Amit Agrawal and SV Prabhu. Survey on measurement of tangential momentum accommodation coefficient. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, 26(4):634–645, 2008.

[86] WB Mann. The exchange of energy between a platinum surface and gas molecules. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 146(859):776–791, 1934.

[87] Lloyd B Thomas and Francois Olmer. The accommodation coefficients of he, ne, a, h2, d2, o2, co2, and hg on platinum as a function of temperature. *Journal of the American Chemical Society*, 65(6):1036–1043, 1943.

[88] Lloyd B Thomas and EB Schofield. Thermal accommodation coefficient of helium on a bare tungsten surface. *The Journal of Chemical Physics*, 23(5):861–866, 1955.

[89] Ching Shen. *Rarefied gas dynamics: fundamentals, simulations and micro flows*. Springer Science & Business Media, 2006.

[90] S-A Kuhlmann, J Reimann, and S Will. On heat conduction between laser-heated nanoparticles and a surrounding gas. *Journal of aerosol science*, 37(12):1696–1716, 2006.

[91] Michail A Gallis, Daniel John Rader, Jaime N Castaneda, John Robert Torczynski, Thomas W Grasser, and Wayne Merle Trott. Microscale rarefied gas dynamics and surface interactions for euvl and mems applications. Technical report, Sandia National Laboratories, 2004.

[92] Carlo Cercignani and Maria Lampis. Kinetic models for gas-surface interactions. *transport theory and statistical physics*, 1(2):101–114, 1971.

[93] RG Lord. Some extensions to the cercignani–lampis gas–surface scattering kernel. *Physics of Fluids A: Fluid Dynamics*, 3(4):706–710, 1991.

[94] William W Liou and Yichuan Fang. *Microfluid mechanics: principles and modeling*. McGraw-Hill Education, 2006.

[95] Tengfei Liang, Qi Li, and Wenjing Ye. Performance evaluation of maxwell and cercignani-lampis gas-wall interaction models in the modeling of thermally driven rarefied gas transport. *Physical Review E*, 88(1):013009, 2013.

[96] Bing-Yang Cao, Jun Sun, Min Chen, and Zeng-Yuan Guo. Molecular momentum transport at fluid-solid interfaces in mems/nems: a review. *International journal of molecular sciences*, 10(11):4638–4706, 2009.

[97] D Bruno, M Cacciatore, S Longo, and M Rutigliano. Gas-surface scattering models for particle fluid dynamics: a comparison between analytical approximate models and molecular dynamics calculations. *chemical physics letters*, 320(3-4):245–254, 2000.

[98] AN Yakunchikov, VL Kovalev, and SV Utyuzhnikov. Analysis of gas-surface scattering models based on computational molecular dynamics. *Chemical Physics Letters*, 554:225–230, 2012.

[99] Jose F Padilla and Iain D Boyd. Assessment of gas-surface interaction models for computation of rarefied hypersonic flow. *Journal of Thermophysics and Heat Transfer*, 23(1):96–105, 2009.

[100] Kyoji Yamamoto, Hideki Takeuchi, and Toru Hyakutake. Scattering properties and scattering kernel based on the molecular dynamics analysis of gas-wall interaction. *Physics of Fluids*, 19(8):087102, 2007.

[101] Nobuhiro Yamanishi, Yoichiro Matsumoto, and Kosuke Shobatake. Multistage gas–surface interaction model for the direct simulation monte carlo method. *Physics of Fluids*, 11(11):3540–3552, 1999.

[102] SV Nedea, AJH Frijns, AA van Steenhoven, Albert J Markvoort, and PAJ Hilbers. Hybrid method coupling molecular dynamics and monte carlo simulations to study the properties of gases in microchannels and nanochannels. *Physical Review E*, 72(1):016705, 2005.

[103] Tengfei Liang and Wenjing Ye. An efficient hybrid dsmc/md algorithm for accurate modeling of micro gas flows. *Communications in Computational Physics*, 15(1):246–264, 2014.

[104] Neil A Mehta and Deborah A Levin. Molecular-dynamics-derived gas–surface models for use in direct-simulation monte carlo. *Journal of Thermophysics and Heat Transfer*, 31(4):757–771, 2017.

[105] Meng Liao, Quy-Dong To, Céline Léonard, and Vincent Monchiet. Non-parametric wall model and methods of identifying boundary conditions for moments in gas flow equations. *Physics of Fluids*, 30(3):032008, 2018.

[106] Nemanja Andric, Daniel W Meyer, and Patrick Jenny. Data-based modeling of gas-surface interaction in rarefied gas flow simulations. *Physics of Fluids*, 31(6):067109, 2019.

[107] Yunus A Cengel and Michael A Boles. *Thermodynamics: An Engineering Approach 6th Editon (SI Units)*. The McGraw-Hill Companies, Inc., New York, 2007.

[108] Shahin Mohammad Nejad, Silvia Nedea, Arjan Frijns, and David Smeulders. The influence of gas–wall and gas–gas interactions on the accommodation coefficients for rarefied gases: A molecular dynamics study. *Micromachines*, 11(3):319, 2020.

[109] Haoyan Sha, Roland Faller, Gulcin Tetiker, and Peter Woytowitz. Molecular simulation study of aluminum–noble gas interfacial thermal accommodation coefficients. *AIChE Journal*, 64(1):338–345, 2018.

[110] Sai Abhishek Peddakotla, Kishore K Kammara, and Rakesh Kumar. Molecular dynamics simulation of particle trajectory for the evaluation of surface accommodation coefficients. *Microfluidics and Nanofluidics*, 23(6):79, 2019.

[111] Douglas A Reynolds and Richard C Rose. Robust text-independent speaker identification using gaussian mixture speaker models. *IEEE transactions on speech and audio processing*, 3(1):72–83, 1995.

[112] Haim Permuter, Joseph Francos, and Ian H Jermyn. Gaussian mixture models of texture and colour for image database retrieval. In *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03).*, volume 3, pages III–569. IEEE, 2003.

[113] Haim Permuter, Joseph Francos, and Ian Jermyn. A study of gaussian mixture models of color and texture features for image classification and segmentation. *Pattern Recognition*, 39(4):695–706, 2006.

[114] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22, 1977.

[115] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.

[116] Aldo Frezzotti, SV Nedea, AJ Markvoort, P Spijker, and Livio Gibelli. Comparison of molecular dynamics and kinetic modeling of gas-surface interaction. In *AIP Conference Proceedings*, volume 1084, pages 635–640. American Institute of Physics, 2008.

[117] Wen-Ming Zhang, Guang Meng, and Xueyong Wei. A review on slip models for gas microflows. *Microfluidics and nanofluidics*, 13(6):845–882, 2012.

[118] Xiao-jun Gu and David R Emerson. A high-order moment approach for capturing non-equilibrium phenomena in the transition regime. *Journal of fluid mechanics*, 636:177–216, 2009.

[119] RG Lord. Application of the cercignani-lampis scattering kernel to direct simulation monte carlo calculations. In *Rarefied Gas Dynamics: 17th International Symposium on Rarefied Gas Dynamics, 1989*, pages 1427–1433, 1989.

[120] Nobuhiro Yamanishi, Yoichiro Matsumoto, and Kosuke Shobatake. Multistage gas–surface interaction model for the direct simulation monte carlo method. *Physics of Fluids*, 11(11):3540–3552, 1999.

[121] Murat Barisik and Ali Beskok. Molecular dynamics simulations of shear-driven gas flows in nano-channels. *Microfluidics and nanofluidics*, 11(5):611–622, 2011.

[122] Shahin Mohammad Nejad, Silvia Nedea, Arjan Frijns, and David Smeulders. The influence of gas–wall and gas–gas interactions on the accommodation coefficients for rarefied gases: A molecular dynamics study. *Micromachines*, 11(3):319, 2020.

[123] Zhi Liang and Pawel Keblinski. Parametric studies of the thermal and momentum accommodation of monoatomic and diatomic gases on solid surfaces. *International Journal of Heat and Mass Transfer*, 78:161–169, 2014.

[124] Peter Atkins and Julio De Paula. *Physical chemistry for the life sciences*. Oxford University Press, USA, 2011.

[125] SM Foiles, MI Baskes, and Murray S Daw. Embedded-atom-method functions for the fcc metals cu, ag, au, ni, pd, pt, and their alloys. *Physical review B*, 33(12):7983, 1986.

[126] Huai Sun. Compass: an ab initio force-field optimized for condensed-phase applications overview with details on alkane and benzene compounds. *The Journal of Physical Chemistry B*, 102(38):7338–7364, 1998.

[127] Jie Yang, Yi Ren, A m Tian, and Huai Sun. Compass force field for 14 inorganic molecules, he, ne, ar, kr, xe, h2, o2, n2, no, co, co2, no2, cs2, and so2, in liquid phases. *The Journal of Physical Chemistry B*, 104(20):4951–4957, 2000.

[128] Marvin Waldman and Arnold T Hagler. New combining rules for rare gas van der waals parameters. *Journal of computational chemistry*, 14(9):1077–1084, 1993.

[129] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics*, 117(1):1–19, 1995.

[130] V Chirita, BA Pailthorpe, and RE Collins. Molecular dynamics study of low-energy ar scattering by the ni (001) surface. *Journal of Physics D: Applied Physics*, 26(1):133, 1993.

[131] Geoffrey J McLachlan and Kaye E Basford. *Mixture models: Inference and applications to clustering*, volume 38. M. Dekker New York, 1988.

[132] Douglas A Reynolds and Richard C Rose. Robust text-independent speaker identification using gaussian mixture speaker models. *IEEE transactions on speech and audio processing*, 3(1):72–83, 1995.

[133] Matti Raitoharju, Ángel F García-Fernández, Roland Hostettler, Robert Piché, and Simo Särkkä. Gaussian mixture models for signal mapping and positioning. *Signal Processing*, 168:107330, 2020.

[134] Lishuai Li, R John Hansman, Rafael Palacios, and Roy Welsch. Anomaly detection via a gaussian mixture model for flight operation and safety monitoring. *Transportation Research Part C: Emerging Technologies*, 64:45–57, 2016.

[135] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22, 1977.

[136] Wayne Trott, Daniel Rader, Jaime Castaneda, John Torczynski, and Michael Gallis. Experimental measurements of thermal accommodation coefficients for microscale gas-phase heat transfer. In *39th AIAA Thermophysics Conference*, page 4039, 2007.

[137] I Amdur and LA Guildner. Thermal accommodation coefficients on gas-covered tungsten, nickel and platinum1. *Journal of the American Chemical Society*, 79(2):311–315, 1957.

[138] Qing Wang, Sanjeev R Kulkarni, and Sergio Verdú. Divergence estimation of continuous distributions based on data-dependent partitions. *IEEE Transactions on Information Theory*, 51(9):3064–3074, 2005.

[139] Naoya Uene, Hideki Takeuchi, Yasutaka Hayamizu, and Takashi Tokumasu. Study of reflection models of gas molecules on water adsorbed surfaces in high-speed flows. *Journal of Fluid Science and Technology*, 15(1):JFST0005–JFST0005, 2020.

[140] Christopher KI Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA, 2006.

[141] FM Devienne, J Souquet, and JC Roustan. Study of the scattering of high energy molecules by various surfaces. *Rarefied Gas Dynamics, Volume 2*, 2:584, 1965.

[142] John C Gregory and Palmer N Peters. A measurement of the angular distribution of 5 ev atomic oxygen scattered off a solid surface in earth orbit. In *International Symposium on Rarefied Gas Dynamics*, 1986.

[143] Deepak S Watvisave, Bhalchandra P Puranik, and Upendra V Bhandarkar. A hybrid md-dsmc coupling method to investigate flow characteristics of micro-devices. *Journal of Computational Physics*, 302:603–617, 2015.

[144] SV Nedea, AJH Frijns, AA Van Steenhoven, Albert J Markvoort, and PAJ Hilbers. Hybrid method coupling molecular dynamics and monte carlo simulations to study the properties of gases in microchannels and nanochannels. *Physical Review E*, 72(1):016705, 2005.

[145] Tengfei Liang and Wenjing Ye. An efficient hybrid dsmc/md algorithm for accurate modeling of micro gas flows. *Communications in Computational Physics*, 15(1):246–264, 2014.

[146] SM Longshaw, R Pillai, L Gibelli, DR Emerson, and DA Lockerby. Coupling molecular dynamics and direct simulation monte carlo using a general and high-performance code coupling library. *Computers & Fluids*, 213:104726, 2020.

[147] Shahin Mohammad Nejad, Silvia Nedea, Arjan Frijns, and David Smeulders. Development of a scattering model for diatomic gas–solid surface interactions by an unsupervised machine learning approach. *Physics of Fluids*, 34(11):117122, 2022.

[148] ID Boyd, DR Beattie, and MA Cappelli. Numerical and experimental investigations of low-density supersonic jets of hydrogen. *Journal of Fluid Mechanics*, 280:41–67, 1994.

[149] Craig White, Matthew K Borg, Thomas J Scanlon, Stephen M Longshaw, Benzi John, David R Emerson, and Jason M Reese. dsmcfoam+: An openfoam based direct simulation monte carlo solver. *Computer Physics Communications*, 224:22–43, 2018.

[150] Henry G Weller, Gavin Tabor, Hrvoje Jasak, and Christer Fureby. A tensorial approach to computational continuum mechanics using object-oriented techniques. *Computers in physics*, 12(6):620–631, 1998.

[151] Bruce W Smith and Kazuaki Suzuki. *Microlithography: science and technology*. CRC press, 2018.

[152] Shiyi Wang, Kaiyi Hou, and Hendrik Heinz. Accurate and compatible force fields for molecular oxygen, nitrogen, and hydrogen to simulate gases, electrolytes, and heterogeneous interfaces. *Journal of Chemical Theory and Computation*, 17(8):5198–5213, 2021.

[153] Huziel E Sauceda, Luis E Gálvez-González, Stefan Chmiela, Lauro Oliver Paz-Borbón, Klaus-Robert Müller, and Alexandre Tkatchenko. Bigdml—towards accurate quantum machine learning force fields for materials. *Nature communications*, 13(1):1–16, 2022.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my wife, Samereh, who has been my soulmate and a very important part of my being. Since we left our home country, like newborn children, we have grown together. We have experienced a lot of ups and downs in our new life abroad. But she was always there for me, and her presence has been a great support throughout my life. I would like to thank my parents for their unconditional love and support during all these years. Without them, this day would not have been possible. I would like to thank Janan, a beautiful small creature, that having him in our family has brought us a lot of joy and calmness.

I would like to thank my supervisors, Arjan Frijns and Silvia Gaasta-Nedea, for their guidance, support, and commitment during my PhD years. Besides receiving a lot of insightful suggestions in our weekly meetings, I always felt trusted by them. This issue motivated me to put more effort into what I am doing and accomplish as much as possible. Furthermore, I really liked the friendly atmosphere during our meetings and all the non-work related discussions we had in our meetings. I thank my promotor, prof. David Smeulders for giving me the opportunity as a PhD candidate. Next, I would like to thank Erik Arlemark for all the interesting discussions we had during my PhD. Such discussions helped me to adapt my simulations so my results can be used in real engineering applications. I would like to thank all members of the Raretrans project, from whom I received numerous constructive and helpful suggestions during the Raretrans project progress meetings. I would like to thank Eldhose Iype for his help and guidance at the beginning of my PhD in becoming familiar with the Machine Learning concept. I would also like to thank all the supervised students who contributed in their way to my project. Especially Frank Peters for his thorough and dedicated work. I would like to express my great appreciation to Linda Essink, who has been very kind and helpful concerning administrative tasks. I sincerely thank my committee members for their assessment and valuable comments on my thesis.

Since moving to Eindhoven in 2018, I have gotten to know many beautiful and amazing people in and outside my job. Siavash, you have become a friend whom I have been missing in my life since I left Iran. I had a lot of memorable times with you these years, and I am delighted to have you in my life. Anastasija, thank you for your friendship and all the wonderful movie and steak nights, as well as those amazing parties. Namdar, I am thankful for your friendship during all these years. Especially for accommodating and helping me during the first couple of weeks when I arrived in the Netherlands. Mauro and Luca, you were not just my good colleagues, you always reminded me of my second home Italy, and I feel very lucky to have you as friends. Koen, I never told you that Samereh and I call you in Farsi (our mother tongue :) "Koene Mehrabun", which means the kind Koen. In fact, you are one of the kindest and most chilled people I have ever met. Thank you for all your kindness and friendship during these years. Max and Ruud, thank you for all those coffee and lunch breaks we had these years, in which you patiently lis-

tened to me nagging about my struggles during my PhD. Menting and Sebastian, thank you for all our nice times in Eindhoven and Rotterdam. Prakhyat and Jessica, you are very dear to me and thank you for all your help and advice when I wanted to start looking for a job. Rick, thank you for all the nights we played tennis during the last couple of years. Although I never won a full set against you, I still enjoyed it a lot during those nights. Dena, thank you for your friendship and all your helpful suggestions and hints while I was writing my thesis.
As a final word, I should mention that during the last few months, people in my home country have been fighting for their freedom. This thesis is heartily dedicated to all those brave souls, especially women of my country, who have initiated and are leading this movement. Women, Life, Freedom.

CURRICULUM VITÆ

Shahin MOHAMMAD NEJAD

Shahin Mohammad Nejad was born on 8 January 1988 in Rasht, Iran. He received his Bachelor's degree in Mechanical Engineering in 2012 from K. N. Toosi University of Technology in Iran. Directly after, he pursued a Master's degree in Mechanical Engineering at Politecnico di Torino in Italy. For his Master's thesis, he went to RWTH Aachen University, Germany, where he worked on Modeling and controlling the thermoacoustic instabilities in a combustion chamber with Laminar flame. He finished his Master's in 2016. After that, for almost two years, he worked as a graduate research assistant in the energy department at Politecnico di Torino. His research topic was modeling the thermal properties of polymers reinforced by carbon nanofibers. In 2018, Shahin was accepted for a PhD position in the Energy Technology group at the Technical University of Eindhoven. Under the supervision of Arjan Frijns and Silvia Gaastra-Nedea, he was involved in a project named Raretrans. In this project, combining the Molecular dynamics simulations with an unsupervised Machine Learning approach named Gaussian Mixture approach, he proposed a new generalized scattering model, which can be used to study the scattering process of rarefied gases interacting with solid surfaces.

LIST OF PUBLICATIONS

- **Mohammad Nejad, S.**, Nedea, S., Frijns, A. and Smeulders, D. (2020) The influence of gas–wall and gas–gas interactions on the accommodation coefficients for rarefied gases: A molecular dynamics study. *Micromachines*, 11(3), p.319.
- **Mohammad Nejad, S.**, Iype, E., Nedea, S., Frijns, A. and Smeulders, D. (2021) Modeling rarefied gas-solid surface interactions for Couette flow with different wall temperatures using an unsupervised machine learning technique. *PHYSICAL REVIEW E*, 104(1), p.015309.
- **Mohammad Nejad, S.**, Nedea, S., Frijns, A. and Smeulders, D. (2022) Development of a scattering model for diatomic gas-solid surface interactions by an unsupervised machine learning approach. *Physics of Fluids*, 34(11), 117122
- **Mohammad Nejad, S.**, Peter, F., Nedea, S., Frijns, A. and Smeulders, D. A hybrid Gaussian Mixture/DSMC approach to study Fourier thermal problem. *to be submitted*